Meaningful Alignments

Agnès Desolneux*, Lionel Moisan*, Jean-Michel Morel*

Abstract

We propose a method for detecting geometric structures in an image, without any *a priori* information. Roughly speaking, we say that an observed geometric event is "meaningful" if the expectation of its occurences would be very small in a random image. We discuss the apories of this definition, solve several of them by introducing "maximal meaningful events" and analyzing their structure. This methodology is applied to the detection of alignments in images.

1 Introduction

Most theories of image analysis tend to find in a given image structures (regions, contours, etc.). These theories assume generally that the images contain such structures and then try to compute their best estimate. The variational framework is quite well adapted to such a viewpoint (for a complete review, see e.g. [14]). The general idea is to minimize a functional of the kind

$$F(u, u_0) + R(u) \ ,$$

where u_0 is the given image defined on a domain $\Omega \subset \mathbb{R}^2$, $F(u, u_0)$ is a fidelity term and $R(u)$ is a regularity term. F and R define an *a priori* model. Let us give two examples:

- The Mumford-Shah model (see [14]), where the energy functional to be minimized is

$$E(u, K) = \lambda^2 \int_{\Omega - K} |\nabla u|^2 \, dx + \mu \lambda^2 \text{length}(K) + \int_{\Omega - K} (u - u_0)^2 \, dx,$$

where u is the estimated image, K its discontinuity set, and the result (u, K) is called a "segmentation" of u_0.

- The Bayesian model (see [5] and [6]): let us denote by $\vec{y} = (y_s)_{s \in S}$ the observation (the degraded image). The aim is to find the "real" image $\vec{x} = (x_s)_{s \in S}$ knowing that the degradation model is given by a conditionnal probability $\Pi(\vec{y}|\vec{x})$, and that the *a priori* law of $\vec{x}$ is given by a Gibbs distribution $\Pi(\vec{x}) = Z^{-1} \exp(-U(\vec{x}))$ (for binary images, the main example is the Ising model). We then have to find the M.A.P. (Maximum A Posteriori) of

$$\Pi(\vec{x}|\vec{y}) = \frac{\Pi(\vec{y}|\vec{x}) \Pi(\vec{x})}{\Pi(\vec{y})} \ .$$

Assume that $\Pi(\vec{y}|\vec{x}) = C \exp(-V(\vec{x}, \vec{y}))$. For example, in the case of a Gaussian noise,

$$\Pi(\vec{y}|\vec{x}) = (\frac{1}{2\pi\sigma^2})^{\frac{|S|}{2}} \exp(-\frac{1}{2\sigma^2} \sum_{s \in S} (y_s - x_s)^2),$$

finding the MAP is equivalent to seeking for the minimum of the functional

$$V(\vec{x}, \vec{y}) + U(\vec{x}) \ .$$

*CMLA, ENS Cachan, 61 avenue du président Wilson 94235 Cachan cedex - France

A main drawback of all the variational methods is that they introduce normalization constants(λ, μ, ...) and the resulting segmentation depends a lot upon the value of these constants. The other point is that they will always deliver a minimum for their functional and so they assume that any image may be segmented (even a white noise). Indeed, they do not yield any criterion to decide whether segmentation is relevant or not. Of course, the probabilistic framework leading to variational methods should in principle give a way to estimate the parameters of the segmentation functional. In the deterministic framework, these parameters can sometimes be estimated as Lagrange multipliers when (e.g.) a noise model is at hand, like in the Rudin-Osher-Fatemi method (see [16]). It is nonetheless easy to check that, first, most variational methods propose a very rough and inaccurate model for the image, second, their parameters are generally not correctly estimated anyway, yielding to supervised methods.

Another drawback of most segmentation methods is their locality. Despite the Gestaltists theories, they look rather for local structure. Let us mention some nonlocal theories of image analysis : the Hough Transform (see [13]), the detection of globally salient structures by Sha'Ashua and Ullman (see [17]), the Extension Field of Guy and Medioni (see [8]) and the Parent and Zucker curve detector (see [15]). These methods have the same drawback as the variational models of segmentation described above. The main point is that they *a priori* suppose that what they want to find (lines, circles, curves...) is in the image. They may find too many or too little such structures in the image and do not yield an **existence proof** for the found structures.

According to Gestalt theory, "grouping" is the law of visual perception (see [11]). Its main idea is that whenever points have a characteristic in common, they get grouped and form the same visual object, a "Gestalt". The main grouping characteristics are alignment, parallelism, common orientation, convexity and closedness (for a curve), ... In addition, this grouping principle is recursive. For example, if points have been grouped into lines, then these lines may again be grouped according (e.g.) to parallelism.

The main aim of this paper is *not* to propose a new segmentation method. We rather propose a computational method to decide wether a given Gestalt (obtained by any segmentation or grouping method) is sure or not. Although most of what we write here can be generalized to other geometric structures, we shall focus on alignments, as one of the most basic Gestalt (see [20]). The above mentioned Hough transform yields a fast algorithm to compute all groups of aligned points in an image. Now, this method assumes the existence of thresholds according to which a peak in the Hough transform means an existing alignment. As one of our main aims, we attempt to give *absolute thresholds*, that is, thresholds permitting to decide when a peak in the Hough transform is significant or not.

In this paper, we only consider the detection of alignments, but we will first give a general definition of what we will call "a meaningful event". The main idea is that a meaningful event is an event that, according to probabilistic estimates, should not happen in an image and therefore is significant. And so we can say that it is a "proven event".

Definition 1 (ε-meaningful event) *We say that an event of type "such configuration of points has such property" is ε-meaningful, if the expectation in a image of the number of occurences of this event is less than ε.*

When $\varepsilon << 1$, we talk about meaningful events. This seems to contradict our notion of a parameter-less theory. Now, it does not, since the ε-dependency of meaningfulness will be low (it will be in fact a $\log \varepsilon$-dependency). The probability that a meaningful event is observed by accident will be very small. In such a case, our perception is liable to see the event, no matter whether it is "true" or not.

The program we state here has been proposed several times in Computer Vision. We know of at least two instances: David Lowe [12] and Witkin-Tenenbaum [21]. Let us quote extensively David Lowe's program, whose mathematical consequences we shall try to develop in this paper: "we need to determine the probability that each relation in the image could have arisen by accident, $P(a)$. Naturally, the smaller that this value is, the more likely the relation is to have a causal interpretation. If we had completely accurate image measurements, the probability of accidental occurence could become vanishingly small. For example, the probability of two image lines being exactly parallel by accident of viewpoint and position is zero. However, in real images there are many factors contributing to limit the accuracy of measurements. Even more important is the fact that we do not want to limit ourselves to perfect instances of each relation in the scene - we want to be able to use the information available from even approximate instances of a relation. Given an image relation that holds within some degree of accuracy, we wish to calculate the probability that it could have arisen by accident to within that level of accuracy. This can only be done in the context of some assumption regarding the surrounding distribution of objects, which serves as the null hypothesis against which we judge significance. One of the most general and obvious assumptions we can make is to assume that a background of independently positioned objects in three-space, which in turn implies independently positioned projections of the objects in the image. This null hypothesis has much to recommend it. (...) Given the assumption of independence in three-space position and orientation, it is easy to calculate the probability that a relation would have arisen to within a given degree of accuracy by accident. For example if two straight lines are parallel to within 5 degrees, we can calculate that the chance is only $5/180 = 1/36$ that the relation would have arisen by accident from two independent objects."

Before proceeding to the mathematical theory, let us give some example and discuss the utility of "ε-meaningfulness".

Example : Let us consider an image of size 100×100 pixels. We assume that the grey-level at each pixel is 0 or 1, which means that we work on a binary image. Our main asumption is that if two points do not belong to the same object, then their grey-levels are independent (and equally distributed if the image is equalized). Now, imagine that we observe the following event: a black 10×10 square. The expectation of the number of 10×10 black squares in the image is simply the number of 10×10 squares in the 100×100 image times the probability that each pixel of a 10×10 square is black. And so the expectation is

$$90 \cdot 90 \cdot \left(\frac{1}{2}\right)^{100} ,$$

which is much less than 1. We conclude that this event is meaningful.
Remarks :
1) Subsquares (large enough) are also meaningful, and so are also candidates to be "Gestalt".
2) Interaction of Gestalts : if we take into account that we observe a 10×10 black square on a 30×30 white background, then the expectation of the number of occurences of this square-on-background event is

$$70 \cdot 70 \cdot \left(\frac{1}{2}\right)^{100} \cdot \left(\frac{1}{2}\right)^{800} ,$$

and so we get a "much more meaningful" event. This is rather a toy example, but it shows immediatly which kind of difficulties and apories are associated with "meaningfulness" :

1. Too many meaningful events : by the same argument as above, all large enough parts of the black square are meaningful. If (e.g.) we take all parts of this square with cardinality larger than 50, they are all meaningful and their number is larger than 2^{50} ! We will see how to solve the problem of having too many meaningful event by defining the notion of "maximal meaningful event".

2. Problem of the *a priori/a posteriori* definition of the event : if we take an arbitrary 10×10 pattern in a 100×100 random binary image, then the expectation of the number of occurences of this event is $90 \cdot 90 \cdot \left(\frac{1}{2}\right)^{100}$ which is much less than 1. The answer is that we need an *a priori* geometric definition of the event, as done in Gestaltism. The event cannot be defined from the observed image itself !

3. Moreover, we can remark that the definition of the geometric event changes its "meaningfulness". For example if we consider our 10×10 black square as a convex set with area 100, then the expectation becomes $\left(\frac{1}{2}\right)^{100}$ times the number of convex sets with area 100. And so the event may loose its meaningfulness.

4. Abstract geometrical character of the information, lack of localization.
ex.1: if we observe a meaningful black patch, all what we can say is: "there is a black patch and the indicated dots may belong to it". We do not know which points belong "for sure" to the event.
ex.2: if we observe a meaningful alignment of points, then we can say "on that line, there are aligned points" but we are not able to define the endpoints.

Our plan is as follows. In Section 2, we explain our definition of meaningful alignments. Section 3 is devoted to the structure properties of the "number of false alarms". In Section 4, we give asymptotic (as $l \to \infty$) and non-asymptotic estimates about the meaningfulness of the following observation : "k well-aligned points in a segment of length l". Section 5 introduces maximal meaningfulness as a mean to reduce the number of events and localize them. Section 6 gives strong arguments in favour of our main conjecture : two maximal meaningful segments on the same line are disjoints. In the experimental Section 7, we compute meaningful and maximal meaningful alignments in several images.

2 Definition of meaningful segments

2.1 Very local computation of the direction of the level lines

Let us consider a gray image of size N (that is N^2 pixels). At each point, we compute a direction, which is the direction of the level line passing by the point calculated on a $q \times q$ pixels neighbourhood (generally $q = 3$ or $q = 2$). Because of the discretization of the image, there are only a finite number n of possible values for the direction at a point. No previous smoothing on the image will be performed and no restoration: such processes would loose the *a priori* independence of directions which is required for the detection method.
We shall use two different methods to compute the direction at a point.

- The first method, that we call the morphological one, is such that $q = 3$ and $n = 16$. At each pixel (i, j), we look at its neighbours in a 3×3 window and at the level lines passing through this window. For example, if

$$\min_{x,y \in \{-1,0,1\}, x+y \geqslant 0} u(i + x, j + y) > \max_{x,y \in \{-1,0,1\}, x+y<0} u(i + x, j + y),$$

  then we say that $\frac{3\pi}{4}$ is a direction of the pixel (i, j). Notice that with this method a pixel may have several directions.

- The second method is based on an interpolation (we have $q = 2$ and we will take $n = 16$). We define the direction at pixel (i, j) by rotating by $\frac{\pi}{2}$ the direction of the gradient of the order 2 interpolation at the center of the 2×2 window made of pixels (i, j), $(i+1, j)$, $(i, j+1)$ and $(i + 1, j + 1)$. We get

$$\mathrm{dir}(i, j) = \frac{1}{||\vec{D}||}\vec{D} \quad \text{where} \quad \vec{D} = \begin{pmatrix} -[u(i, j + 1) + u(i + 1, j + 1)] + [u(i, j) + u(i + 1, j)] \\ [u(i + 1, j) + u(i + 1, j + 1)] - [u(i, j) + u(i, j + 1)] \end{pmatrix}.$$

  Then we say that two points X and Y have the same direction with precision $\frac{1}{n}$ if

$$\mathrm{Angle}(\mathrm{dir}(X), \mathrm{dir}(Y)) \leqslant \frac{2\pi}{n}.$$

2.2 Probabilistic model

Our main assumption is following: we shall assume that in order to have no alignments, the direction at all points in an image must be a random variable, and that each direction value has uniform probability $\frac{1}{n} = p$. And in all the following, we assume that $n > 2$ so that $p < \frac{1}{2}$. In a structureless image, when two pixels are such that their distance is more than 3 (or 2, it depends on the method used to compute the directions), the directions computed at the two considered pixels should be independent random variables. We assume that every deviation from this randomness assumption will lead to the detection of a structure (Gestalt) in the image. Thus, it is no surprise that all of our estimates below will be done as if the image was a random process. There is another way to understand these estimates, however. We know (by experience) that images have contours and therefore meaningful alignments. This is mainly due to the smoothness of contours of solid objects and the generation of geometric structure by most physical and biological laws. Now, it can be assumed that in a first approximation, the relative positions of objects are independent. This means that whenever two points x and y belong to the same contour, their directions are likely to be higly correlated, while if they belong to two different objects, their directions should be independent. We stick to the idea that perception simply gives a privilege to any deviation from uniform randomness.

From now on, the computations will be performed on any image presenting at each pixel a direction which is uniformly distributed, two points at a distance larger than $q = 3$ having independent directions. Let S be a segment in the image made of l independent pixels (it means that the distance between two consecutive points of S is 3 and so, the real length of S is $3l$). We are interested in the number of points of S which have the property of having their direction aligned with the direction of S. Such points of S will simply be called *aligned points of* S.

The question is to know what is the minimal number $k(l)$ of aligned points that we must observe on a length l segment so that this event becomes meaningful when it is observed in a real image.

2.3 Definition of meaning

Let S be a straight segment with length l and $x_1, x_2, \dots, x_l$ be the l (independent) points of S. Let X_i be the random variable whose value is 1 when the direction at pixel x_i is aligned with the direction of S, and 0 otherwise. We then have the following distribution for X_i :

$$\mathrm{P}[X_i = 1] = p \qquad \text{and} \qquad \mathrm{P}[X_i = 0] = 1 - p \; .$$

The random variable representing the number of x_i having the "good" direction is

$$S_l = X_1 + X_2 + \dots + X_l \; .$$

Because of the independence of the X_i, the law of S_l is given by the binomial distribution

$$\mathrm{P}[S_l = k] = \binom{l}{k} p^k (1-p)^{l-k} \; .$$

When we consider a length l segment, we want to know whether it is ε-meaningful or not among all the segments of the image (and not only among the segments having the same length l). Let $m(l)$ be the number of oriented segments of length l in a $N \times N$ image. We define the total number of oriented segments in a $N \times N$ image as the number of pairs (x, y) of points in the image (an oriented segment is given by its starting point and its ending point) and so we have

$$\sum_{l=0}^{l_{max}} m(l) = N^4 \; .$$

Definition 2 (ε-meaningful segment) *A length l segment is ε-meaningful in a $N \times N$ image if it contains at least $k(l)$ points having their direction aligned with the one of the segment, where $k(l)$ is given by*

$$k(l) \;=\; \min\left\{k \in \mathbb{N},\; \mathrm{P}\left[S_l \geqslant k\right] \leqslant \frac{\varepsilon}{N^4}\right\} \; .$$

Let us develop and explain this definition. For $1 \leqslant i \leqslant N^4$, let A_i be the following event: "the i-th segment is ε-meaningful" and χ_{A_i} denote the characteristic function of the event A_i. We have

$$\mathrm{P}\left[\chi_{A_i} = 1\right] \;=\; \mathrm{P}\left[S_{l_i} \geqslant k(l_i)\right]$$

where l_i is the length of the i-th segment. Notice that if l_i is small we may have $\mathrm{P}\left[S_{l_i} \geqslant k(l_i)\right] = 0$. Let R be the random variable representing the exact number of A_i occuring simultaneously in a trial. The law of probability of R depends a lot upon the relations of dependence between the segments, but not its expectation (see [3]) : since $R = \chi_{A_1} + \chi_{A_2} + \ldots + \chi_{A_{N^4}}$, the expectation of R is

$$E(R) \;=\; E(\chi_{A_1}) + E(\chi_{A_2}) + \;\ldots\; + E(\chi_{A_{N^4}}) \;=\; \sum_{l=0}^{l_{max}} m(l)\mathrm{P}\left[S_l \geqslant k(l)\right] \; .$$

By definition we have

$$\mathrm{P}\left[S_l \geqslant k(l)\right] \leqslant \frac{\varepsilon}{N^4}, \quad \text{so that} \quad E(R) \;\leqslant\; \frac{\varepsilon}{N^4} \cdot N^4 \leqslant \varepsilon.$$

This means that the expectation of the number of ε-meaningful segments in an image is less than ε.

Remark : We could have defined a ε-meaningful length l segment as a segment ε-meaningful only among the set of the length l segments. It would have been a segment with at least $k'(l)$ points having the "good" direction where $k'(l)$ is defined by $m(l) \cdot \mathrm{P}\left[S_l \geqslant k'(l)\right] \leqslant \varepsilon$. Notice that $m(l) \simeq N^3$ because there are approximately N^2 possible discrete straight lines in a $N \times N$ image and on each discrete line, about N choices for the starting point of the segment. But we did not keep this definition because when looking for alignments we cannot *a priori* know the length of the segment we look for. In the same way, we never consider events like : "a segment has exactly k aligned points", but rather "a segment has at least k aligned points", and k must be given, as we do, by a detectability criterion and not *a priori* fixed.

3 Number of false alarms

3.1 Definition

Definition 3 (Number of false alarms) *let S be a segment of length l_0 with at least k_0 points having their direction aligned with the direction of S. We define the number of false alarms of S as*

$$NF(k_0, l_0) \;=\; N^4 \cdot \mathrm{P}\left[S_{l_0} \geqslant k_0\right] \;=\; N^4 \cdot \sum_{k=k_0}^{l_0} \binom{l_0}{k} p^k (1-p)^{l_0-k} \; .$$

Interpretation of this definition : the number $NF(k_0, l_0)$ of false alarms of the segment S represents an upper-bound of the expectation in an image of the number of segments of probability less than the one of the considered segment.

Remark : (relative notion) Let S be a segment and $NF(k_0, l_0)$ its number of false alarms. Then S is ε-meaningful if and only if $NF(k_0, l_0) \;\leqslant\; \varepsilon$, but it is worth noticing that we could have compared $NF(k_0, l_0)$ not to ε but to the real number of segments with probability less than the one of S, observed in the image. For example, if we observe 100 segments of probability less than α, and if the expected value R of the number of segments of probability less than α was 10, we are able to say that this 100-segments event could happen with probability less than 1/10, since $10 = E(R) \geqslant 100 \cdot \mathrm{P}[R = 100]$. Now, each of these 100 segments only is 10-meaningful !

3.2 Properties of the number of false alarms

Proposition 1 *The number of false alarms $NF(k_0, l_0)$ has the following properties :*

1. $NF(0, l_0) = N^4$, *which proves that the event for a segment to have more than zero aligned points is never meaningful !*
2. $NF(l_0, l_0) = N^4 \cdot p^{l_0}$, *which shows that a segment such that all of its points have the "good" direction is ε-meaningful if its length is larger than $(-4 \ln N + \ln \varepsilon)/\ln p$.*
3. $NF(k_0+1, l_0) < NF(k_0, l_0)$. *This can be interpreted by saying that if two segments have the same length l_0, the "more meaningful" is the one which has the more "aligned" points.*
4. $NF(k_0, l_0) < NF(k_0, l_0 + 1)$. *This property can be illustrated by the following figure of a segment (where a $\bullet$ represents a misaligned point, and a $\rightarrow$ represents an aligned point) :*

   $$\rightarrow\rightarrow \bullet \rightarrow\rightarrow \bullet\bullet \rightarrow\rightarrow\rightarrow\rightarrow\rightarrow \bullet$$

   If we remove the last point (on the right), which is misaligned, the new segment is less probable and therefore more meaningful than the considered one.
5. $NF(k_0+1, l_0+1) < NF(k_0, l_0)$. *Again, we can illustrate this property :*

   $$\rightarrow\rightarrow \bullet \rightarrow\rightarrow \bullet\bullet \rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow$$

   If we remove the last point (on the right), which is aligned, the new segment is more probable and therefore less meaningful than the considered one.

Proof : The four first properties present no difficulty. Concerning the fifth one, we prove that

$$\mathrm{P}[S_{l_0+1} \geqslant k_0 + 1] = \mathrm{P}[S_{l_0} \geqslant k_0] - (1-p) \cdot \binom{l_0}{k_0} p^{k_0}(1-p)^{l_0-k_0},$$

which is a simple consequence of relations between binomial coefficients.

$$\begin{aligned}
\mathrm{P}[S_{l_0+1} \geqslant k_0 + 1] &= \sum_{k=k_0}^{l_0} \binom{l_0+1}{k+1} p^{k+1}(1-p)^{l_0-k} \\
&= \sum_{k=k_0}^{l_0} \binom{l_0}{k} p^{k+1}(1-p)^{l_0-k} + \sum_{k=k_0}^{l_0-1} \binom{l_0}{k+1} p^{k+1}(1-p)^{l_0-k} \\
&= p\mathrm{P}[S_{l_0} \geqslant k_0] + (1-p)\left(\mathrm{P}[S_{l_0} \geqslant k_0] - \binom{l_0}{k_0} p^{k_0}(1-p)^{l_0-k_0}\right) \\
&= \mathrm{P}[S_{l_0} \geqslant k_0] - (1-p) \cdot \binom{l_0}{k_0} p^{k_0}(1-p)^{l_0-k_0}\ .
\end{aligned}$$

□

If we consider a length l segment (made of l independent pixels), then the expectation of the number of points of the segment having the same direction as the one of the segment is simply the expectation of the random variable S_l, that is

$$\mathrm{E}(S_l) \;=\; \sum_{i=1}^{l} E(X_i) \;=\; \sum_{i=1}^{l} \mathrm{P}\,[X_i = 1] \;=\; p \cdot l\ .$$

We are interested in ε-meaningful segments, which are the segments such that their number of false alarms is less than ε. These segments have a small probability (less than ε/N^4), and since they represent alignments (deviation from randomness), they should contain more aligned points than the expected number computed above. That is the main point of the following proposition.

Proposition 2 *Let S be a segment of length $l_0 \geqslant 1$, containing at least k_0 points having the same direction as the one of S. If $NF(k_0, l_0) \leqslant p \cdot N^4$, (which is the case when S is meaningful), then*

$$k_0 \;\geqslant\; pl_0 \,+\, (1-p)\ .$$

This proposition will be proved by Lemma 4, where we will extend the discrete function $P(k,l) = \mathrm{P}\,[S_l \geqslant k]$ to a continuous domain.

4 Thresholds and asymptotic estimations

In the following, ε and p are fixed numbers smaller than 1, and we use the notation

$$P(k,l) = \mathrm{P}\,[S_l \geqslant k] = \sum_{i=k}^{l} \binom{l}{i} p^i (1-p)^{l-i}\ .$$

We recall that a segment of length l is ε-meaningful as soon as it contains at least $k(l)$ points having the "right" direction, where $k(l)$ is defined by

$$k(l) \;=\; \min\left\{k \in \mathbb{N},\; P(k,l) \;\leqslant\; \frac{\varepsilon}{N^4}\right\}. \tag{1}$$

In this section, we shall give precise asymptotic and non-asymptotic estimates of the thresholds $k(l)$, which roughly say that

$$k(l) \simeq pl + \sqrt{C \cdot l \cdot \ln \frac{N^4}{\varepsilon}},$$

where $2p(1-p) \leqslant C \leqslant \frac{1}{2}$. Some of these results are illustrated by Figure 1

4.1 Sufficient condition of meaningfulness

In this part, we will see how the theory of large deviations and other inequalities concerning the tail of the binomial distribution can provide us a sufficient condition of meaningfulness. The key point is the following result due to Hoeffding (see [9]).

Theorem 1 (Hoeffding's inequality) *If k, l are positive integers with $k \leqslant l$, and if p is a real number such that $0 < p < 1$.*
Then if $r = k/l \geqslant p$, we have the inequalities

$$P(k,l) \leqslant \exp\left(lr \ln\frac{p}{r} + l(1-r)\ln\frac{1-p}{1-r}\right) \leqslant \exp(-l(r-p)^2 h(p)) \leqslant \exp(-2l(r-p)^2)\ ,$$

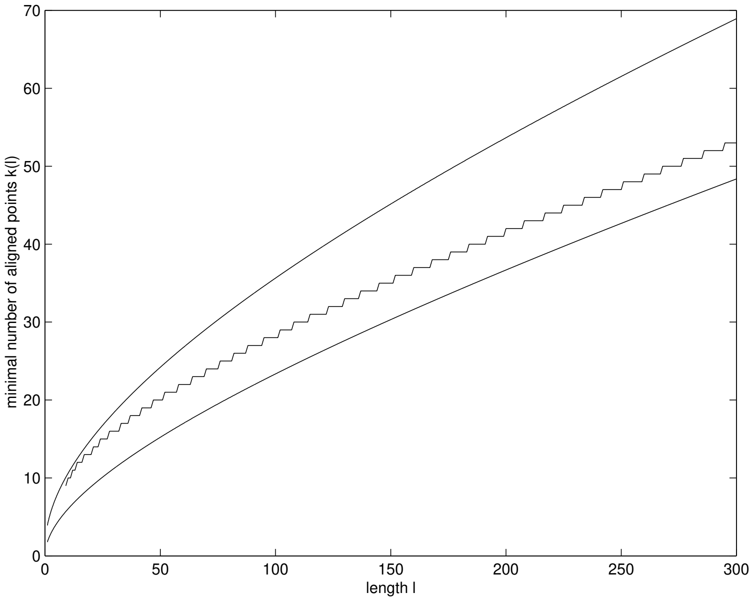

Figure 1: Estimates for the threshold of meaningfulness $k(l)$

The middle (stepcase) curve represents the exact value of the minimal number of aligned points $k(l)$ to be observed on a 1-meaningful segment of length l in an image of size 512, for a direction precision of $1/16$. The upper and lower curves represent estimates of this threshold obtained by Corollary 1 and Proposition 4.

Where h is the function defined on $]0,1[$ by

$$h(p) = \frac{1}{1-2p} \ln \frac{1-p}{p} \qquad \text{for } 0 < p < \frac{1}{2} ,$$

$$h(p) = \frac{1}{2p(1-p)} \qquad \text{for } \frac{1}{2} \leqslant p < 1 .$$

Using this theorem, we deduce a sufficient condition for a segment to be meaningful. The size N of the image, and the probability $p < 1/2$ of a given direction are fixed.

Proposition 3 (sufficient condition of ε-meaningfulness) *Let S be a length l segment, containing at least k aligned points. If*

$$k \geqslant pl + \sqrt{\frac{4 \ln N - \ln \varepsilon}{h(p)}} \sqrt{l} ,$$

then S is ε-meaningful.

Proof : Let S be a length l segment, containing at least k aligned points, where k and l are such that

$$k \geqslant pl + \sqrt{\frac{4 \ln N - \ln \varepsilon}{h(p)}} \sqrt{l} .$$

If we denote $r = k/l$, then $r \geqslant p$ and

$$l(r-p)^2 \geqslant \frac{4 \ln N - \ln \varepsilon}{h(p)} .$$

And so finally, by Theorem 1 we deduce that

$$P(k,l) \leqslant \exp(-l(r-p)^2 h(p)) \leqslant \exp(-4\ln N + \ln\varepsilon) = \frac{\varepsilon}{N^4} ,$$

which means, by definition, that the segment S is ε-meaningful.

□

Corollary 1 *Let S be a length l segment, containing at least k aligned points. If*

$$k \geqslant pl + \sqrt{\frac{l}{2}(4\ln N - \ln\varepsilon)} ,$$

then S is ε-meaningful.

Proof : This result is a simple consequence of Proposition 3 and of the fact that for p in $]0,1[$, $h(p) \geqslant 2$ (see Hoeffding [9]).

□

Notice that Corollary 1 is interesting only when

$$l \geqslant pl + \sqrt{\frac{l}{2}(4\ln N - \ln\varepsilon)} ,$$

that is when

$$l \geqslant \frac{4\ln N - \ln\varepsilon}{2(1-p)^2} .$$

Numerical example: for $\varepsilon = 1$, $N = 512$ and $p = 1/16$, we obtain $l \geqslant 15$.

4.2 Necessary conditions for meaningfulness

The first simple necessary condition we can get is a threshold on the length l. For an ε-meaningful segment, we have

$$p^l \leqslant \mathrm{P}\left[S_l \geqslant k(l)\right] \leqslant \frac{\varepsilon}{N^4} ,$$

so that

$$l \geqslant \frac{-4\ln N + \ln\varepsilon}{\ln p} .$$

Let us give a numerical example : if the size of the image is $N = 512$, and if $p = 1/16$ (which corresponds to 16 possible directions), the minimal length of a 1-meaningful segment is $l_{min} = 9$.

This necessary condition is only on l, so we now look for more precise conditions involving both k and l.

Lemma 1 *Let $0 < r < 1$ be a real number, and g_r the function defined on $]0,1[$ by*

$$g_r(x) = r\ln x + (1-r)\ln(1-x) ,$$

then g_r is concave and has its maximum at point $x = r$. Moreover if $0 < p \leqslant r$ then

$$2(r-p)^2 \leqslant g_r(r) - g_r(p) \leqslant \frac{(r-p)^2}{p(1-p)} .$$

Proof : By simple differentiation of g_r we get

$$g_r'(x) \;=\; \frac{r}{x} - \frac{1-r}{1-x} \;=\; \frac{r-x}{x(1-x)} \quad \text{and} \quad g_r''(x) \;=\; -\frac{r}{x^2} - \frac{1-r}{(1-x)^2}\;,$$

which proves that g_r is concave, and is increasing on $]0,r]$ and decreasing on $[r,1[$.
Let $0 < p \leqslant r$, then the left inequality is a consequence of the property that if x is a real number in $[0,1]$, then $x(1-x) \leqslant 1/4$. Then we have

$$g_r(r) - g_r(p) = \int_p^r g_r'(x)\,dx = \int_p^r \frac{r-x}{x(1-x)}\,dx \geqslant 4\int_p^r (r-x)\,dx = 2(r-p)^2\,.$$

The right inequality is a direct consequence of the concavity of g_r:

$$g_r(r) - g_r(p) \leqslant (r-p)g_r'(p) = \frac{(r-p)^2}{p(1-p)}\;.$$

□

Lemma 2 *If $N \geqslant 3$ and if $S=(k,l)$ is a ε-meaningful segment with $1 \leqslant k < l$, then if we denote $r = k/l$,*

$$g_r(r) - g_r(p) \;>\; \frac{\ln N - \ln\varepsilon}{l} \;>\; \frac{1}{l}.$$

Proof : Let $S=(k,l)$ be an ε- meaningful segment, then

$$\binom{l}{k} p^k(1-p)^{l-k} \;\leqslant\; P(k,l) \;\leqslant\; \frac{\varepsilon}{N^4}\;.$$

If n is an integer larger than 1, we have the Stirling's formula refined to (see [3] for example)

$$n^n e^{-n}\sqrt{2\pi n}e^{1/(12n+1)} \;\leqslant\; n! \;\leqslant\; n^n e^{-n}\sqrt{2\pi n}e^{1/12n}\;.$$

We then deduce that

$$\binom{l}{k} \geqslant \frac{l^l\sqrt{2\pi l}}{k^k\sqrt{2\pi k}(l-k)^{(l-k)}\sqrt{2\pi(l-k)}}e^{\frac{1}{12l+1}-\frac{1}{12k}-\frac{1}{12(l-k)}}\;.$$

We assumed that $1 \leqslant k \leqslant l-1$ and so we get

$$\frac{1}{12k(l-k)} - \frac{1}{12l+1} \leqslant \frac{1}{12(l-1)} - \frac{1}{12l+1} = \frac{13}{12(l-1)(12l+1)} \leqslant \frac{13}{6l\cdot 12l} \leqslant \frac{1}{5l^2}\;.$$

On the other hand, we have $k(l-k) = l^2 r(1-r) \leqslant l^2/4$, and we then obtain

$$\frac{\varepsilon}{N^4} \geqslant P(k,l) \geqslant \binom{l}{k} p^k(1-p)^{l-k} \geqslant \frac{2}{\sqrt{2\pi l}}e^{-1/5l^2}e^{l(g_r(p)-g_r(r))}\;.$$

And consequently

$$l(g_r(r)-g_r(p)) \geqslant 4\ln N - \ln\varepsilon - \frac{1}{2}\ln l + \ln\frac{2}{\sqrt{2\pi}} - \frac{1}{5l^2}\;.$$

Since the size of the considered image is $N\times N$ and l is a length of a segment of the image, we have $l \leqslant \sqrt{2}N$. And so

$$l(g_r(r)-g_r(p)) \geqslant \left(4-\frac{1}{2}\right)\ln N - \ln\varepsilon - \frac{1}{2}\ln\frac{\pi}{\sqrt{2}} - \frac{1}{5l^2}\;.$$

This inequality permits to conclude since the hypothesis $N \geqslant 3$ ensures that

$$\frac{7}{2}\ln N - \frac{1}{2}\ln\frac{\pi}{\sqrt{2}} - \frac{1}{5l^2} > \ln N.$$

$\square$

Corollary 2 (necessary condition) *If $S=(k,l)$ is a 1-meaningful segment, then*

$$k \geqslant pl + \sqrt{p(1-p)\cdot l \ln N}.$$

Proof : This is a direct consequence of Lemma 1 and Lemma 2.

$\square$

We shall refine this estimation in Section 4.4 by using a large deviation refinment of the Central Limit Theorem.

4.3 Asymptotics for the meaningfulness threshold $k(l)$

In this section, we still consider that ε and p are fixed. We will work on asymptotic estimations of $P(k,l)$ when l is "large". We first recall a version of the Central limit theorem in the particular case of the binomial distribution (see [3]).

Theorem 2 (De Moivre-Laplace limit theorem) *If α is a fixed positive number, then as l tends to $+\infty$,*

$$\mathrm{P}\left[S_l \geqslant pl + \alpha\sqrt{l\cdot p(1-p)}\right] \longrightarrow \frac{1}{\sqrt{2\pi}}\int_{\alpha}^{+\infty} e^{-x^2/2}\,dx\ .$$

Our aim is to get the asymptotic behaviour when l is large of the threshold $k(l)$ defined by (1). The problem is that if l gets to infinity, we also have to consider that N tends to infinity (because, since l is the length of a segment in a $N\times N$ image, necessarily $l \leqslant \sqrt{2}N$). And so the α used in the De Moivre-Laplace theorem will depend on N. This is the reason why we use the following stronger version of the previous theorem (see [3]).

Theorem 3 (Feller) *If $\alpha(l)\to+\infty$ and $\alpha(l)^6/l\to 0$ as $l\to+\infty$, then*

$$\mathrm{P}\left[S_l \geqslant pl + \alpha(l)\sqrt{l\cdot p(1-p)}\right] \sim \frac{1}{\sqrt{2\pi}}\int_{\alpha(l)}^{+\infty} e^{-x^2/2}\,dx\ .$$

Proposition 4 (asymptotic behaviour of $k(l)$) *When $N\to+\infty$ and $l\to+\infty$ in such a way that $l/(\ln N)^3\to+\infty$, one has*

$$k(l) = pl + \sqrt{2p(1-p)\cdot l\cdot\left(\ln\frac{N^4}{\varepsilon} + O(\ln\ln N)\right)}.$$

Proof : We define, for $i\in\{0,1\}$,

$$\alpha_i(l,N) = \frac{k(l)-i-pl}{\sqrt{lp(1-p)}}.$$

Corollary 2 implies that

$$\alpha_0(l,N) \geqslant \sqrt{\ln N},$$

so that $\alpha_i(l,N) \to +\infty$ as $l \to \infty$. Conversely, Corollary 1 says that

$$k(l) \leqslant pl + \sqrt{\frac{l}{2}(4\ln N - \ln\varepsilon)},$$

from which we deduce that

$$\frac{\alpha_i^6(l,N)}{l} \leqslant C\frac{(4\ln N - \ln\varepsilon)^3}{l} ,$$

where C is a constant depending only on p. Since ε is fixed and $l/(\ln N)^3 \to +\infty$, we get that $\alpha_i^6(l,N)/l \to 0$. Hence, we can apply Feller's Theorem to obtain

$$\forall i \in \{0,1\}, \quad \mathrm{P}\left[S_l \geqslant pl + \alpha_i(l,N)\sqrt{l\cdot p(1-p)}\right] \sim \frac{1}{\sqrt{2\pi}}\int_{\alpha_i(l,N)}^{+\infty} e^{-x^2/2}\,dx. \tag{2}$$

For $i = 0$ (resp. for $i = 1$), the left hand term of (2) is smaller (resp. larger) than ε/N^4. Besides, the right hand term is equivalent to

$$\frac{1}{\sqrt{2\pi}\alpha_i(l,N)}e^{-\alpha_i^2(l,N)/2}.$$

For $i = 0$, we deduce that

$$\frac{1}{\sqrt{2\pi}}\frac{1}{\alpha_0(l,N)}e^{-\alpha_0^2(l,N)/2}(1+o(1)) \leqslant \frac{\varepsilon}{N^4},$$

which implies

$$O(1) + O(\ln(\alpha_0(l,N))) - \frac{\alpha_0^2(l,N)}{2} + o(1) \leqslant \ln\frac{N^4}{\varepsilon},$$

and finally

$$\alpha_0(l,N)^2 \geqslant 2\ln\frac{N^4}{\varepsilon} + O(\ln\ln N),$$

that is

$$k(l) \geqslant pl + \sqrt{2p(1-p)\cdot l\cdot\left(\ln\frac{N^4}{\varepsilon} + O(\ln\ln N)\right)}. \tag{3}$$

The case $i = 1$ gives in a similar way

$$k(l) - 1 \leqslant pl + \sqrt{2p(1-p)\cdot l\cdot\left(\ln\frac{N^4}{\varepsilon} + O(\ln\ln N)\right)}. \tag{4}$$

Finally (3) and (4) yield the estimation of $k(l)$ announced in Proposition 4.

□

4.4 Lower bound for the meaningfulness threshold $k(l)$

In this part, we refine the necessary condition of ε-meaningfulness obtained in Section 4.2 by using a comparison between the Binomial and the Gaussian laws given by the following

Theorem 4 (Slud 1977) *If $0 < p \leqslant 1/4$ and $pl \leqslant k \leqslant l$, then*

$$\mathrm{P}\left[S_l \geqslant k\right] \geqslant \frac{1}{\sqrt{2\pi}}\int_{\alpha(k,l)}^{+\infty} e^{-x^2/2}\,dx \qquad \text{where} \quad \alpha(k,l) = \frac{k-pl}{\sqrt{lp(1-p)}} .$$

Proposition 5 (necessary condition of meaningfulness) *We assume that $0 < p \leqslant 1/4$ and N are fixed. If a segment $S = (k, l)$ is ε-meaningful then*

$$k \geqslant pl + \alpha(N)\sqrt{lp(1-p)} \ ,$$

where $\alpha(N)$ is uniquely defined by

$$\frac{1}{\sqrt{2\pi}} \int_{\alpha(N)}^{+\infty} e^{-x^2/2}\, dx \ = \ \frac{\varepsilon}{N^4} \ .$$

This proposition is a direct consequence of Slud's Theorem. The assumption $0 < p \leqslant 1/4$ is not a strong condition since it is equivalent to consider that the number of possible oriented directions is larger than 4.

5 Maximal meaningful segments

5.1 Definition

Suppose that on a straight line we have found a meaningful segment S with a very small number of false alarms (i.e. $NF(S) << 1$). Then if we add some "spurious" points at the end of the segment we obtain another segment with probability higher than the one of S and having still a number of false alarms less than 1, which means that this new segment is still meaningful (see figure).

$$\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow \bullet \bullet \bullet\bullet$$

In the same way, it is likely to happen in general that many subsegments of S having a probability higher than the one of S will still be meaningful (see experimental Section, where this problem obviously occurs for the DNA image). These remarks justify the introduction of the following notion of "maximal segment".

Definition 4 (Maximal segment) *A segment A is maximal if*

1. *it does not contain a strictly more meaningful segment : $\forall B \subset A,\ NF(B) \geqslant NF(A)$,*
2. *it is not contained in a more meaningful segment : $\forall B \supset A,\ NF(B) > NF(A)$,*

Then we say that a segment is *maximal meaningful* if it is both maximal and meaningful. This notion of "maximal meaningful segment" is linked to what Gestaltists called the "masking phenomenon". According to this phenomenon, most parts of an object are "masked" by the object itself except the parts which are significant from the point of view of the construction of the whole object. For example, if one considers a square, the only significant segments of this square are the four sides, and not large parts of the sides. With our definition, long enough parts of a side may be meaningful segments, but only the whole side itself will be a maximal meaningful segment.

5.2 Elementary properties

Proposition 6 (Properties of maximal segments) *Let A be a maximal segment, then*

1. *the two endpoints of A have their direction aligned with the direction of A,*
2. *the two points next to A (one on each side) do not have their direction aligned with the direction of A.*

Proof : Let A be a maximal segment with length l_0 and containing k_0 points having their direction aligned with the direction of A.

1. Let us suppose that one endpoint $\{x\}$ of A has its direction not aligned with the direction of A. Then,

$$NF(A - \{x\}) \;=\; NF(k_0, l_0 - 1) \;<\; NF(k_0, l_0) \;=\; NF(A),$$

which contradicts the fact that A is maximal.

2. If we assume that a point $\{y\}$ next to A has its direction aligned with the one of A, then

$$NF(A \cup \{y\}) \;=\; NF(k_0 + 1, l_0 + 1) \;<\; NF(k_0, l_0) \;=\; NF(A),$$

which also contradicts the fact that A is maximal. □

5.3 Continuous extension of the binomial tail

We first extend the discrete function $P(k,l)$ to a continuous domain.

Lemma 3 *The map*

$$\tilde{P} : (k,l) \;\mapsto\; \frac{\displaystyle\int_0^p x^{k-1}(1-x)^{l-k}\,dx}{\displaystyle\int_0^1 x^{k-1}(1-x)^{l-k}\,dx} \tag{5}$$

is

1. *continuous on the domain* $\left\{(k,l) \in \mathbb{R}^2, \quad 0 \leqslant k \leqslant l < +\infty\right\}$,
2. *decreasing with respect with* k,
3. *increasing with respect with* l,

and for all integer values of k *and* l *one has* $\tilde{P}(k,l) = P(k,l)$.

Proof : The continuity results from classical theorems on the regularity of parameterized integrals. Notice that the continuous extension of $\tilde{P}$ when $k = 0$ is $\tilde{P}(0,l) = 1$. Now, we prove that $\tilde{P}(k,l)$ is decreasing with respect with k. For that purpose, we introduce the map

$$A(k,l) = \frac{\displaystyle\int_0^p x^{k-1}(1-x)^{l-k}\,dx}{\displaystyle\int_p^1 x^{k-1}(1-x)^{l-k}\,dx}.$$

Since $1/\tilde{P} = 1 + 1/A$, we need to prove that A decreases with respect with k. We compute

$$\frac{1}{A}\frac{\partial A}{\partial k}(k,l) = \frac{\displaystyle\int_0^p x^{k-1}(1-x)^{l-k} \cdot \ln\frac{x}{1-x}\,dx}{\displaystyle\int_0^p x^{k-1}(1-x)^{l-k}\,dx} - \frac{\displaystyle\int_p^1 x^{k-1}(1-x)^{l-k} \cdot \ln\frac{x}{1-x}\,dx}{\displaystyle\int_p^1 x^{k-1}(1-x)^{l-k}\,dx},$$

and we apply the mean value theorem to obtain the existence of (α, β) such that

$$0 < \alpha < p < \beta < 1 \qquad \text{and} \qquad \frac{1}{A}\frac{\partial A}{\partial k}(k,l) = \ln\frac{\alpha}{1-\alpha} - \ln\frac{\beta}{1-\beta}.$$

The right hand term being negative, the proof is complete. As concerns the increasingness of P with respect with l, the proof is similar, the increasing map $x \mapsto \ln \frac{x}{1-x}$ being replaced by the decreasing map $x \mapsto \ln(1-x)$. Finally, the fact that $\tilde{P}(k,l) = P(k,l)$ for integer values of k and l is a consequence of the relation $\tilde{P}(k+1,l+1) = p\tilde{P}(k,l) + (1-p)\tilde{P}(k+1,l)$ (see [3] for example).
□

Remark : Properties (2) and (3) guarantee that $\tilde{P}$ is a "good" interpolate of P in the sense that the monotonicity of P in both variables k and l is extended to the continuous domain. Notice that a proof based on the same method (thanks to the increasingness of $x \mapsto \ln x$) will establish that

$$\frac{\partial \tilde{P}}{\partial k} + \frac{\partial \tilde{P}}{\partial l} \leqslant 0,$$

which is the natural extension of the property $P(k+1,l+1) \leqslant P(k,l)$ previously established in Proposition 1.

From now on, we shall assume that $p < 1/2$. The following property is a good example of the interest of the continuous extension of P.

Lemma 4 *If $l \geqslant 1$, then $p \leqslant \tilde{P}(p(l-1)+1,l) < \frac{1}{2}$.*

Proof : Thanks to Lemma 3, it is sufficient to prove that if $k-1 = p(l-1)$, then

$$\frac{p}{1-p}\int_p^1 x^{k-1}(1-x)^{l-k}\,dx \leqslant \int_0^p x^{k-1}(1-x)^{l-k}\,dx < \int_p^1 x^{k-1}(1-x)^{l-k}\,dx. \tag{6}$$

For that purpose, we write $f(x) = x^{k-1}(1-x)^{l-k}$ and we study the map

$$g(x) = \frac{f(p-x)}{f(p+x)}.$$

A simple computation gives that up to a positive multiplicative term, $g'(x)$ writes

$$2x^2(k-1-(1-p)(l-1)) - 2p(1-p)(k-1-p(l-1)),$$

and since $k-1 = p(l-1)$ and $p < 1/2$, we have $g' < 0$ on $]0,p]$. Hence, $g(x) < g(0) = 1$ on $]0,p]$, which implies

$$\int_0^p f(x)dx = \int_0^p f(p-x)dx < \int_0^p f(p+x)dx = \int_p^{2p} f(x)dx < \int_p^1 f(x)dx$$

and the right side of (6) is proved.

For the left side, we follow the same reasoning with the map

$$g(x) = \frac{f(p-x)}{f(p+\frac{1-p}{p}x)}.$$

After a similar computation, we obtain that $g' \geqslant 0$ on $]0,p]$, so that $f(p-x) \geqslant f(p+\frac{1-p}{p}x)$ on $]0,p]$. We integrate this inequality to obtain

$$\int_0^p f(x)dx = \int_0^p f(p-x)dx \geqslant \int_0^p f(p+\frac{1-p}{p}x)dx = \frac{p}{1-p}\int_p^1 f(x)dx,$$

which proves the left side of (6). □

5.4 Density of maximal segments

In general, it is not easy to compare $P(k,l)$ and $P(k',l')$ by performing simple computations on k, k', l and l'. However, a simple case is solved by the following

Proposition 7 *Let $A=(k,l)$ and $B=(k',l')$ be two 1-meaningful segments of a $N\times N$ image (with $N\geqslant 3$) such that*

$$\frac{k'}{l'}\geqslant\frac{k}{l}\qquad\text{and}\qquad l'>l.$$

Then, B is more meaningful than A, that is $NF(B)<NF(A)$.

Lemma 5 *Define for $p<r\leqslant 1$, $B(r,l)=\tilde{P}(rl,l)$. Then, one has*

$$g_r(r)-g_r(p)>\frac{1}{l}\qquad\Longrightarrow\qquad\frac{\partial B}{\partial l}<0,$$

where g_r is the map defined in Section 4.

Proof : We first write the Beta integral in terms of the Gamma function (see [1]),

$$\int_0^1 t^{x-1}(1-t)^{y-1}\,dt\;=\;\frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

Thanks to (5), this yields

$$B(r,l)=\frac{\Gamma(l+1)}{\Gamma(rl)\,\Gamma((1-r)l+1)}\int_0^p x^{rl-1}(1-x)^{(1-r)l}\,dx. \tag{7}$$

We now use the expansion (see [1])

$$\frac{d\ln\Gamma(x)}{dx}\;=\;-\gamma-\frac{1}{x}+\sum_{n=1}^{+\infty}(\frac{1}{n}-\frac{1}{x+n}), \tag{8}$$

where γ is Euler's constant. Using (7) and (8), we obtain

$$\begin{aligned}\frac{\partial B}{\partial l}\;=\;&-\gamma-\frac{1}{l+1}+\sum_{n=1}^{+\infty}(\frac{1}{n}-\frac{1}{l+1+n})-r\left[-\gamma-\frac{1}{rl}+\sum_{n=1}^{+\infty}(\frac{1}{n}-\frac{1}{rl+n})\right]\\&-(1-r)\left[-\gamma-\frac{1}{(1-r)l+1}+\sum_{n=1}^{+\infty}(\frac{1}{n}-\frac{1}{(1-r)l+1+n})\right]\\&+\frac{\int_0^p(r\ln x+(1-r)\ln(1-x))x^{rl-1}(1-x)^{(1-r)l}\,dx}{\int_0^p x^{rl-1}(1-x)^{(1-r)l}\,dx}.\end{aligned}$$

The function $x\mapsto r\ln x+(1-r)\ln(1-x)$ is increasing on $]0,r[$, and we have $p<r$, so

$$\frac{\int_0^p(r\ln x+(1-r)\ln(1-x))x^{rl-1}(1-x)^{(1-r)l}\,dx}{\int_0^p x^{rl-1}(1-x)^{(1-r)l}\,dx}\leqslant r\ln p+(1-r)\ln(1-p)\ .$$

Then

$$\frac{\partial B}{\partial l}\leqslant\frac{1}{l}+\sum_{n=1}^{+\infty}(\frac{r}{rl+n}+\frac{1-r}{(1-r)l+n}-\frac{1}{l+n})+r\ln p+(1-r)\ln(1-p)\ .$$

Now, let us consider the function

$$f : x \mapsto \frac{r}{rl+x} + \frac{1-r}{(1-r)l+x} - \frac{1}{l+x},$$

defined for all $x > 0$. Since $0 < r \leqslant 1$ we have $rl + x \leqslant l + x$ and $(1-r)l + x \leqslant l + x$, so that $f(x) \geqslant 0$ and

$$f'(x) = -\frac{r}{(rl+x)^2} - \frac{1-r}{((1-r)l+x)^2} + \frac{1}{(l+x)^2} \leqslant 0.$$

We deduce that for N integer larger than 1,

$$\sum_{n=1}^{N} f(n) \leqslant \int_0^N f(x)\,dx .$$

A simple integration gives

$$\int_0^N f(x)\,dx = r\ln(1+\frac{rl}{N}) + (1-r)\ln(1+\frac{(1-r)l}{N}) - \ln(1+\frac{l}{N}) - r\ln r - (1-r)\ln(1-r).$$

Finally

$$\sum_{n=1}^{+\infty}(\frac{r}{rl+n} + \frac{1-r}{(1-r)l+n} - \frac{1}{l+n}) \leqslant -r\ln r - (1-r)\ln(1-r) ,$$

which yields

$$\frac{\partial B}{\partial l} \leqslant \frac{1}{l} - r\ln r - (1-r)\ln(1-r) + r\ln p + (1-r)\ln(1-p) = \frac{1}{l} - g_r(r) + g_r(p),$$

and the right hand term is negative thanks to the hypothesis. □

Proof of Proposition 7 : Let us define $r = k/l$, since A and B are 1-meaningful we have $r > p$ and consequently

$$P(k',l') = \tilde{P}(\frac{k'}{l'}\cdot l', l') \leqslant \tilde{P}(rl', l') = B(r,l') < B(r,l) = P(k,l)$$

thanks to Lemma 3 and Lemma 5. □

Remark : For the approximation of $\tilde{P}(k,l)$ given by the Gaussian Law

$$G(k,l) = \frac{1}{\sqrt{2\pi}}\int_{\alpha(k,l)}^{+\infty} e^{-\frac{x^2}{2}}\,dx \qquad \text{where} \qquad \alpha(k,l) = \left(\frac{k}{l} - p\right)\sqrt{\frac{l}{p(1-p)}},$$

we immediatly have the result that $G(k',l') < G(k,l)$ when $k'/l' \geqslant k/l > p$ and $l' > l$.

6 A conjecture and some partial results

Up to now, we have established some properties that permit to characterize or compare meaningful segments. We now study the structure of maximal segments, and give some evidence (a proof relying on a technical conjecture asymptotically proven) that two distinct maximal segments on a same straight line have no common point.

Definition 5 (curvature) *Consider a function $F : \mathbb{R}^2 \to \mathbb{R}$ of class C^2. The curvature of F is the function defined at each point $(k,l) \in \mathbb{R}^2$ where the gradient of F is nonzero by*

$$\mathrm{curv}(F) = \frac{F_l^2 F_{kk} - 2F_k F_l F_{kl} + F_k^2 F_{ll}}{(F_k^2 + F_l^2)^{3/2}}.$$

In this definition, the indices represent partial derivatives. Notice that if DF and $[D^2F]$ represent respectively the gradient of F and its second derivative (viewed as a bilinear form), then

$$\mathrm{curv}(F) = \frac{[D^2F](DF^\perp, DF^\perp)}{\|DF\|^3}.$$

The curvature of F at point (k,l) is the (signed) curvature of the level curve $F = cte$ passing through this point. If $k(l,\lambda)$ is defined by $F(k(l,\lambda),l) = \lambda$, then one has

$$\mathrm{curv}(F) = \left[1 + \left(\frac{\partial k}{\partial l}\right)^2\right]^{-3/2} \frac{\partial^2 k}{\partial l^2}. \tag{9}$$

The result we are going to present now partially rely on the following

Conjecture 1 *The map $(k,l) \mapsto \tilde{P}(k,l)$ defined in Lemma 3 has negative curvature on the domain $D_p = \{(k,l) \in \mathbb{R}^2,\ p(l-1)+1 \leqslant k \leqslant l\}$.*

Thanks to (9), it is equivalent to say that the level curves $l \mapsto k(l,\lambda)$ of $\tilde{P}$ defined by $\tilde{P}(k(l,\lambda),l) = \lambda$ are concave, i.e. satisfy

$$\forall (k_0,l_0) \in D_p, \qquad \frac{\partial^2 k}{\partial l^2}(l_0, \tilde{P}(k_0,l_0)) < 0.$$

Remark : All numerical computations we have realized so far for the function $\tilde{P}(k,l)$ have been in agreement with Conjecture 1. Concerning theoretical results, we shall prove that this conjecture is asymptotically true in next section. For now, the following result shows that Conjecture 1 is satisfied for the Gaussian approximation of the binomail tail (correct for small deviations, that is $k \simeq pl + C\sqrt{l}$).

Proposition 8 *The approximation of $\tilde{P}(k,l)$ given by the Gaussian law*

$$G(k,l) = \frac{1}{\sqrt{2\pi}} \int_{\alpha(k,l)}^{+\infty} e^{-\frac{x^2}{2}}\, dx \qquad \text{where} \qquad \alpha(k,l) = \frac{k - pl}{\sqrt{lp(1-p)}}$$

has negative curvature on the domain D_p.

Proof : The level lines of $G(k,l)$ can be written under the form

$$k(l,\lambda) = pl + f(\lambda)\sqrt{l},$$

with $f > 0$ on the domain $\{k > pl\}$. Hence, we have

$$\frac{\partial^2 k}{\partial l^2}(l,\lambda) = -\frac{f(\lambda)}{4l^{3/2}}$$

and consequently $\mathrm{curv}(G) < 0$ on D_p thanks to (9). $\square$

6.1 Consequences of Conjecture 1

Lemma 6 (under Conjecture 1) *If $k-1 \geqslant p(l-1)$ and $\mu > 0$, then the map*

$$f : x \mapsto \tilde{P}(k+\mu x, l+x)$$

has no local minimum in $x = 0$.

Proof : It is sufficient to prove that either $f'(0) \neq 0$ or ($f'(0) = 0$ and $f''(0) < 0$). If $f'(0) = 0$, then

$$\mu = -\frac{\tilde{P}_l}{\tilde{P}_k}(k,l),$$

so that

$$f''(0) = \mathrm{curv}(\tilde{P})(k,l) \cdot (1+\mu^2)^{3/2} < 0$$

thanks to Conjecture 1.

□

Proposition 9 *If $(l,l',l'') \in [1,+\infty)^3$ and $(k,k',k'') \in [0,l] \times [0,l'] \times [0,l'']$, then*

$$\min\Big(p, \tilde{P}(k,l), \tilde{P}(k+k'+k'', l+l'+l'')\Big) < \max\Big(\tilde{P}(k+k', l+l'), \tilde{P}(k+k'', l+l'')\Big). \qquad (10)$$

Proof : Because the inequality is symmetric in k' and k'', we can suppose that $k''/l'' \geqslant k'/l'$. If $k+k'-1 \leqslant p(l+l'-1)$, then $P(k+k', l+l') > p$ and we have finished. If not, let us define the map

$$f(x) = P(k+x(k'+k''), l+x(l'+l'')) \qquad \text{for} \quad x \in [0,1].$$

We remark that for $x_0 = l'/(l'+l'') \in]0,1[$, we have

$$k + x_0(k'+k'') = k + \frac{l'}{l'+l''}(k'+k'') \geqslant k + \frac{l'}{l'+l''}(k' + \frac{k'l''}{l'}) = k+k',$$

which implies that $P(k+k', l+l') \geqslant f(x_0)$. Hence, it is sufficient to prove that

$$\min\Big(p, f(0), f(1)\Big) < f(x_0).$$

The set

$$S = \Big\{x \in [0,1], \quad k + x(k'+k'') - 1 \geqslant p(l + x(l'+l'') - 1)\Big\}$$

is a connected segment that contains 0 or 1, and whose inside part contains x_0. Since f has no local minimum on S thanks to Lemma 6, we conclude as announced that

$$f(x_0) > \min_{x \in S} f(x) = \min_{x \in \partial S} f(x) \geqslant \min\left(p, f(0), f(1)\right),$$

since if $x \in \partial S \cap]0,1[$, then $f(x) \geqslant p$ thanks to Lemma 4.

□

Corollary 3 (Union and Intersection) *If A and B are two segments on the same straight line, then*

$$\min\Big(pN^4, NF(A \cap B), NF(A \cup B)\Big) < \max\Big(NF(A), NF(B)\Big).$$

This is a direct consequence of Proposition 9 for integer values of k, k', k'', l, l' and l''. Numerically, we checked this property for all segments A and B such that $|A \cup B| \leqslant 256$. For $p = 1/16$, we obtained

$$\min_{|A\cup B|\leqslant 256} \frac{\max\Big((NF(A), NF(B)\Big) - \min\Big(pN^4, NF(A\cap B), NF(A\cup B)\Big)}{\max\Big((NF(A), NF(B)\Big) + \min\Big(pN^4, NF(A\cap B), NF(A\cup B)\Big)} \simeq 0.000754697... > 0,$$

this minimum (independant of N) being obtained for $A = (23, 243)$, $B = (23, 243)$ and $A \cap B = (22, 230)$ (as before, the couple (k, l) we attach to each segment represents the number of aligned points (k) and the segment length (l)).

Theorem 5 (maximal segments are disjoint) *Suppose that Conjecture 1 is true. Then, any two maximal segments lying on the same straight line have no intersection.*

Notice that this property applies to maximal segments and not only to maximal meaningful segments.

Proof : Suppose that one can find two maximal segments $(k + k', l + l')$ and $(k + k'', l + l'')$ that have a non-empty intersection (k, l) Then, according to Proposition 9 we have

$$\min\Big(p, P(k,l), P(k+k'+k'', l+l'+l'')\Big) < \max\Big(P(k+k', l+l'), P(k+k'', l+l'')\Big).$$

If the left hand term is equal to p, then we have a contradiction since one of $(k + k', l + l')$ or $(k+k'', l+l'')$ is strictly less meaningful than the segment $(1, 1)$ it contains. If not, we have another contradiction because one of $(k + k', l + l')$ or $(k + k'', l + l'')$ is strictly less meaningful than one of (k, l) or $(k + k' + k'', l + l' + l'')$. □

Remark : The numerical checking of Proposition 9 ensures that for $p = 1/16$ (but we could have checked for another value of p), two maximal meaningful segments with total length smaller than 256 are disjoint, which is enough for most practical applications.

6.2 Partial results about Conjecture 1

In all the following, we assume that p and r are fixed and satisfy $0 < p < r < 1$ and $p < 1/2$.

Proposition 10 (large deviations estimate)

$$\tilde{P}(rl+1, l+1) \quad \underset{l\to+\infty}{\sim} \quad \frac{p(1-p)}{(r-p)\sqrt{2\pi l r(1-r)}}\; e^{-l\left(r\ln\frac{r}{p} + (1-r)\ln\frac{1-r}{1-p}\right)}. \tag{11}$$

Compared to Hoeffding's inequality, where only the exponential term appears (see Theorem 1), we have a more precise estimate of $\tilde{P}(k, l)$.

Proposition 11 *For any $\lambda \in [0, 1]$ and $l > 0$, there exists a unique $k(l, \lambda)$ such that*

$$\tilde{P}(k(l,\lambda) + 1, l + 1) = \lambda. \tag{12}$$

Moreover, when $r \in]p, 1[$ is fixed one has

$$\frac{\partial^2 k}{\partial l^2}(l, \tilde{P}(rl+1, l+1)) \quad \underset{l\to+\infty}{\sim} \quad -\frac{\left(r\ln\frac{r}{p} + (1-r)\ln\frac{1-r}{1-p}\right)^2}{l\;\cdot\; r(1-r)\;\cdot\;\left(\ln\frac{r(1-p)}{(1-r)p}\right)^3}. \tag{13}$$

We shall prove these results at the end of this section. It is interesting to notice that (13) remains true when $k(l,\lambda)$ is defined not from $\tilde{P}$ but from its estimate obtained in (11). In the same way, one can prove that

$$\frac{\partial k}{\partial l}(l, \tilde{P}(rl+1, l+1)) \underset{l\to+\infty}{\longrightarrow} \frac{\ln \dfrac{1-p}{1-r}}{\ln \dfrac{r(1-p)}{(1-r)p}}$$

is satisfied with both definitions of $k(l,\lambda)$. This prove that (11) is in fact a very good estimate of $\tilde{P}$, since it not only approximates the values of $\tilde{P}$ but also its level lines up to second order.

Corollary 4 (asymptotic proof of Conjecture 1) *For any $r \in]p,1[$, there exists $L > 0$ such that the map $(k,l) \mapsto P(k,l)$ has negative curvature on the domain*

$$D_p^{L,r} = \{(rl+1, l+1), \quad l \in [L, +\infty[\}.$$

Proof : Define $k(l,\lambda)$ by (12). If r is fixed, for L large enough we have

$$\forall l > L, \qquad \frac{\partial^2 k}{\partial l^2}(l, \tilde{P}(rl+1, l+1)) < 0$$

thanks to Proposition 11. Hence, from (9) we deduce that $\mathrm{curv}(\tilde{P})(rl+1, l+1) < 0$ for $l > L$. $\square$

6.3 Proof of Proposition 10 and 11

We are going to apply a technique called "stationary phase" or "Laplace method", which has been used a lot in microlocal analysis (see [7], [10]). First, we study more precisely the map g_r defined in Lemma 1.

Lemma 7 *Consider the map*

$$g_r(x) = r \ln x + (1-r)\ln(1-x) \tag{14}$$

and define the increasing maps $a : [0, +\infty[\to [-p, 0[$ and $b :]-\infty, +\infty[\to]0,1[$ by

$$g_r(-a(X)) = g_r(p) - X \qquad \text{and} \qquad g_r(b(U)) = g_r(r) - U^2. \tag{15}$$

Then, a'' and b'' are bounded and

$$a(0) = -p, \qquad a'(0) = \frac{p(1-p)}{r-p}, \qquad b(0) = r, \qquad b'(0) = \sqrt{2r(1-r)}.$$

Proof : We first derive the left side of (15) to obtain, thanks to Lemma 1,

$$a'(X) = \frac{1}{g_r'(-a(X))} \quad \Longrightarrow \quad a'(0) = \frac{1}{g_r'(p)} = \frac{p(1-p)}{r-p}.$$

Then, we derive the right side of (15) twice to get

$$-2 = b''(U) g_r'(b(U) + b'(U)^2 g_r''(b(U)), \tag{16}$$

and since $g_r'(b(0)) = g_r'(r) = 0$, we get as announced

$$b'(0) = \sqrt{\frac{-2}{g_r''(r)}} = \sqrt{2r(1-r)}.$$

Using higher order derivatives, one can check without difficulties that a'' and b'' are bounded. $\square$

We now recall some classical computations.

Lemma 8 *For $l > 0$ and n integer,*

$$\begin{aligned}
\int_0^\infty X^n e^{-lX}\, dX &= \frac{n!}{l^{n+1}}, \\
\int_0^\infty U^{2n} e^{-lU^2}\, dU &= \frac{(2n)!}{4^n n!\, l^n}\sqrt{\frac{\pi}{l}}, \\
\int_0^\infty U^{2n+1} e^{-lU^2}\, dU &= \frac{p!}{2l^{n+1}}.
\end{aligned}$$

Proof of Proposition 10 : Le us write $Q(k,l) = P(k+1, l+1)$. We first recall that

$$Q(k,l) = \frac{\displaystyle\int_0^p x^k (1-x)^{l-k}\, dx}{\displaystyle\int_0^1 u^k (1-u)^{l-u}\, du} = \frac{\displaystyle\int_0^p e^{lg_r(x)}\, dx}{\displaystyle\int_0^1 e^{lg_r(u)}\, du}, \tag{17}$$

where g_r is the map previously defined and $r = k/l$. Now we perform on (17) the variable changes $x = -a(X)$ and $u = b(U)$ to obtain

$$Q(k,l) = e^{-l(g_r(r)-g_r(p))} \frac{\displaystyle\int_0^\infty e^{-lX} a'(X)\, dX}{\displaystyle\int_{-\infty}^\infty e^{-lU^2} b'(U)\, dU}.$$

But since a'' and b'' are bounded, we have, using Lemma 8,

$$Q(k,l) = e^{-l(g_r(r)-g_r(p))} \frac{a'(0)\frac{1}{l} + O\left(\frac{1}{l^2}\right)}{b'(0)\sqrt{\frac{\pi}{l}} + O\left(\frac{1}{l}\right)} \sim e^{-l(g_r(r)-g_r(p))} \frac{a'(0)}{b'(0)\sqrt{\pi l}}$$

and we conclude with Lemma 7. □

Proof of Proposition 11 : First, the monotonicity of $Q(k,l)$ with respect with k ensures the existence and uniqueness of the map $k(l,\lambda)$ such that $Q(k(l,\lambda),l) = \lambda)$. Moreover, from (17) we get

$$\frac{\partial k}{\partial l} = -\frac{\dfrac{\partial Q}{Q\partial l}}{\dfrac{\partial Q}{Q\partial k}} = R \qquad \text{with} \qquad R(k,l) = \frac{\displaystyle\int_0^p \int_0^1 e^{l(g_r(x) + g_r(u))} \ln\frac{1-x}{1-u}\, du\, dx}{\displaystyle\int_0^p \int_0^1 e^{l(g_r(y) + g_r(v))} \ln\frac{(1-y)v}{y(1-v)}\, dv\, dy},$$

so that

$$\begin{aligned}
\frac{\partial^2 k}{\partial l^2} &= \frac{\partial k}{\partial l}\frac{\partial R}{\partial k} + \frac{\partial R}{\partial l} \\
&= \frac{\displaystyle\int_{d^3} e^{l(g_r(x) + g_r(y) + g_r(z) + g_r(u) + g_r(v) + g_r(w))}\, f(x,y,z,u,v,w)\, dxdudydvdzdw}{\left(\displaystyle\int_d e^{l(g_r(x) + g_r(u))}\, g(x,u)\, dxdu\right)^3},
\end{aligned}$$

where $f(x,y,z,u,v,w) =$

$$\ln\frac{1-x}{1-u}\ln\frac{(1-y)v}{y(1-v)}\cdot\left(\ln\frac{1-z}{1-w}\ln\frac{xu(1-y)(1-v)}{(1-x)(1-u)yv}+\ln\frac{(1-z)w}{z(1-w)}\ln\frac{(1-x)(1-u)}{(1-y)(1-v)}\right),$$

$$g(x,u)=\ln\frac{(1-x)u}{x(1-u)}\qquad\text{and}\qquad d=[0,p]\times[0,1].$$

Now, performing the variable changes $x=-a(X)$, $y=-a(Y)$, $z=-a(Z)$, $u=b(U)$, $v=b(V)$ and $w=b(W)$ yields

$$\frac{\partial^2 k}{\partial l^2}=\frac{\displaystyle\int_{D^3} e^{-l(X+Y+Z+U^2+V^2+W^2)}\ F(X,Y,Z,U,V,W)\ dX\,dU\,dY\,dV\,dZ\,dW}{\left(\displaystyle\int_D e^{-l(X+U^2)}G(X,U)\ dX\,dU\right)^3},\tag{18}$$

where $F(X,Y,Z,U,V,W) =$

$$f\Big(-a(X),-a(Y),-a(Z),b(U),b(V),b(W)\Big)\ a'(X)a'(Y)a'(Z)b'(U)b'(V)b'(W),$$

$$G(X,U)=g(-a(X),b(U))\ a'(X)b'(U)\qquad\text{and}\qquad D=]0,\infty[\times]-\infty,\infty[.$$

After a tedious computation (a formal computation software is welcome), one can check that

$$F(\mathbf{0})=0,\qquad\left(\frac{\partial F}{\partial X}+\frac{\partial F}{\partial Y}+\frac{\partial F}{\partial Z}\right)(\mathbf{0})=0,$$

so that the principal term of the numerator of (18) is due to

$$C_1=\left(\frac{\partial^2 F}{\partial U^2}+\frac{\partial^2 F}{\partial V^2}+\frac{\partial^2 F}{\partial W^2}\right)(\mathbf{0})=\frac{8p^3(1-p)^3\sqrt{2r(1-r)}}{(r-p)^3}\left(r\ln\frac{r}{p}+(1-r)\ln\frac{1-r}{1-p}\right)^2.$$

Concerning the denominator of (18), the main contribution is due to

$$C_2=G(\mathbf{0})=\frac{-p(1-p)\sqrt{2r(1-r)}}{r-p}\ln\frac{(1-p)r}{p(1-r)}.$$

Hence, we have, as $l\to+\infty$,

$$\frac{\partial^2 k}{\partial l^2}\quad\sim\quad\frac{\dfrac{\pi^{3/2}}{2l^{11/2}}\cdot\dfrac{C_1}{2}}{\left(\dfrac{\pi^{1/2}}{l^{3/2}}\cdot C_2\right)^3}\quad\sim\quad-\frac{\left(r\ln\dfrac{r}{p}+(1-r)\ln\dfrac{1-r}{1-p}\right)^2}{l\ \cdot\ r(1-r)\ \cdot\ \left(\ln\dfrac{r(1-p)}{(1-r)p}\right)^3}$$

thanks to Lemma 8. □

7 Experiments

7.1 Image 1 : the DNA image (size 512×512)

In all the experiments on this image the direction at a point is computed as the morphological one ($q=3$ and $p=1/16$). We first computed the detection of all the meaningful segments having in addition the property that their endpoints have their direction aligned with the one of the segment (image 1.a): black points represent points on a meaningful segment which have the same direction as the one of the segment, and gray points represent points on a meaningful segment which don't have the same direction as the segment. The result shows that all the expected alignments are found, and even the horizontal ones. On this image, one can notice that some segments are grouped together according to the alignment property. We next computed the detection of the maximal meaningful alignments (image 1.b). The resulting image is of course simpler than the preceding one, but one can see that all the main alignments are still found. The third experiment on the ADN image is the image 1.c. On that image, we only displayed the meaningful segments of length less than 20 independent pixels (it means that their real length is less than 60 pixels). The result is that all the expected segments are found, but we loose the grouping of distant meaningful aligned segments.

7.2 Image 2 : the Megawave image (size 256×256)

On image 2.a, we compute the detection of the meaningful segments of length less than 15, but now the direction is the "gradient" one (still $p = 1/16$ and $q = 3$). To avoid aliasing effects, we only keep meaningful segments which are such that the mean of the directions of all well-aligned points of the segment is very close to the direction of the segment itself : numerically we keep a meaningful segment when the angular distance between the mean and the direction is less than $\frac{\pi}{24}$.

References

[1] J. Anastassiadis, *Définition des fonctions eulériennes par des équations fonctionnelles"*, ed. Gauthier-Villars 1965.

[2] H. Bateman, *Higher Transcendental Functions*, vol 1.

[3] W. Feller, *An introduction to probability theory and its applications*, vol. I, 3rd edition, Wiley 1968.

[4] M. Fréchet, *Les probabilités associées à un système d'événements compatibles et dépendants*, 1ère et 2ème parties, Actualités Scientifiques et Industrielles, Hermann 1940.

[5] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images", *IEEE Trans. Pattern Anal. Machine Intell.* 6 , pp. 721-741, 1984

[6] S. Geman and C. Graffigne, "Markov Random Field Image Models and Their Applications to Computer Vision", *Proceedings of the International Congress of Mathematicians*, Berkeley USA 1986.

[7] A. Grigis and J. Sjostrand, "Microlocal analysis for Differential Operators : An introduction", *London Mathematical Society, Lecture Note Series 196*, Cambridge University Press, 1994.

[8] G. Guy and G. Medioni, "Inferring global perceptual contours from local features", *IEEE Trans. Pattern Anal. Machine Intell.*, 1992.

[9] W. Hoeffding, "Probability inequalities for sum of bounded random variables", *Journal of the American Statistical Association*, No.58, p 13-30.

[10] L. Hörmander, *The Analysis of Linear Partial Differential Operators I*, Grundlehren, Springer, 256 (1983).

[11] G. Kaniza, *La Grammaire du Voir*, Editions Diderot, arts et sciences.

[12] D. Lowe, *Perceptual Organization and Visual Recognition*, Kluwer Academic Publishers.

[13] H. Maitre, "Un panorama de la transformation de Hough", *Traitement du signal*, vol.2 No.4, 1985.

[14] J.-M. Morel and S. Solimini, *Variational Methods In Image Segmentation*, Birkhähauser.

[15] P. Parent and S.W. Zucker, "Trace inference, curvature consistency and curve detection", *IEEE Trans. Pattern Anal. Machine Intell.* vol.2, No.8, 1989.

[16] L. Rudin, S. Osher and E. Fatemi, "Nonlinear total variation based noise removal algorithms", *Physica D 60*, No.1-4, pp. 259-268, 1992.

[17] A. Sha'Ashua and S. Ullman, "Structural saliency : the detection of globally salient structures using a locally connected network", *Proceedings of the 2d Int. Conf. on Computer Vision*, pp. 321-327, 1988.

[18] E. Slud, "Distribution inequalities for the binomial law", *Annals of Probability*, vol. 5, pp. 404-412.

[19] J. Stojanov, I. Mirazchiiski, Z. Ignatov and M. Tanushev, *Exercise Manual in Probability Theory*, Kluwer Academic 1989.

[20] M. Wertheimer, "Untersuchungen zur Lehre der Gestalt", II, *Psychologische Forschung*, vol. 4, pp. 301-350, 1923.

[21] A.P.Witkin and J.P.Tenenbaum, "On the role of structure in vision", Beck, Hope and Rosenfeld, Eds. (New York: Academic Press, 1983), pp. 481-543.

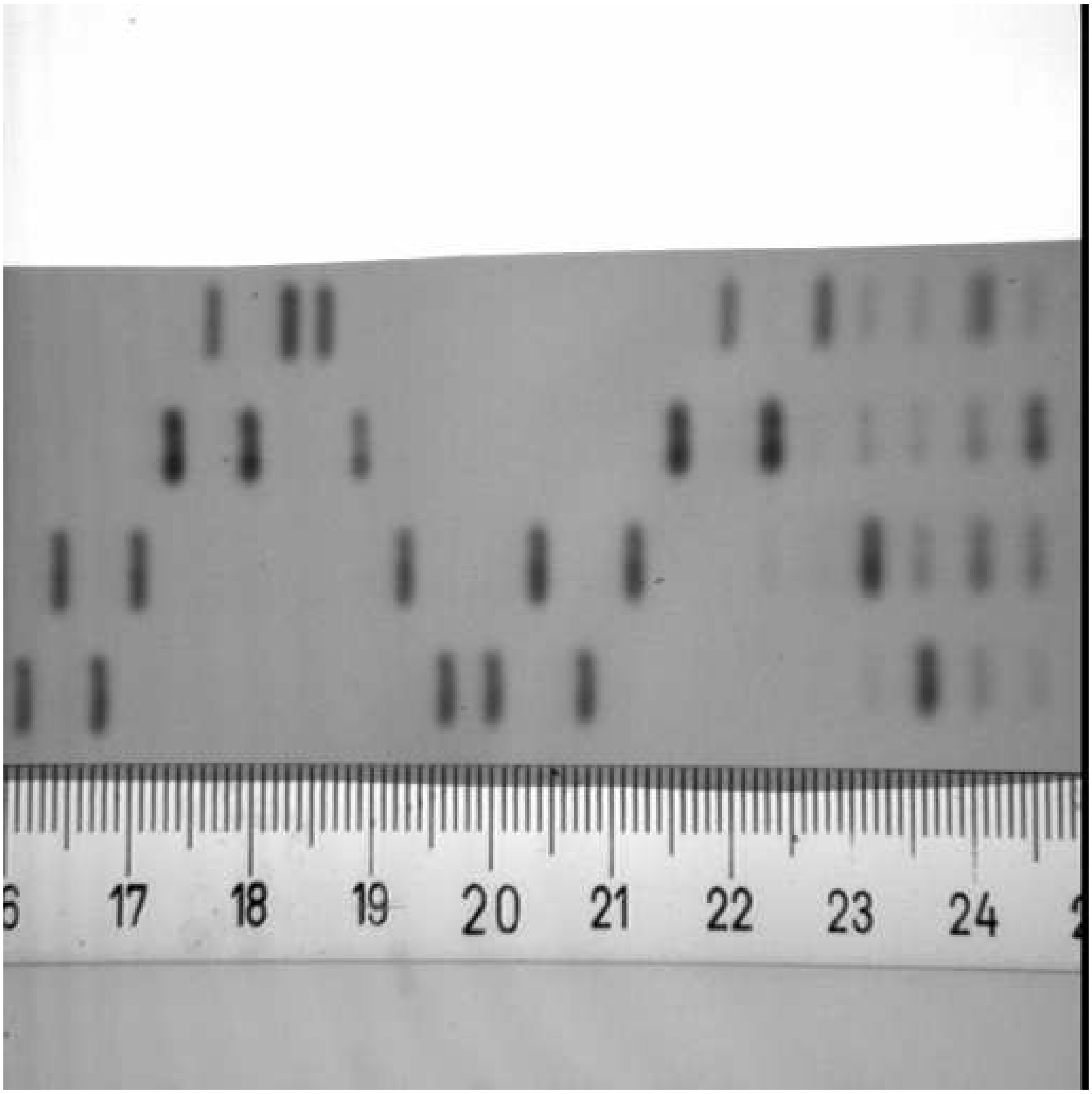

Figure 2: Image 1

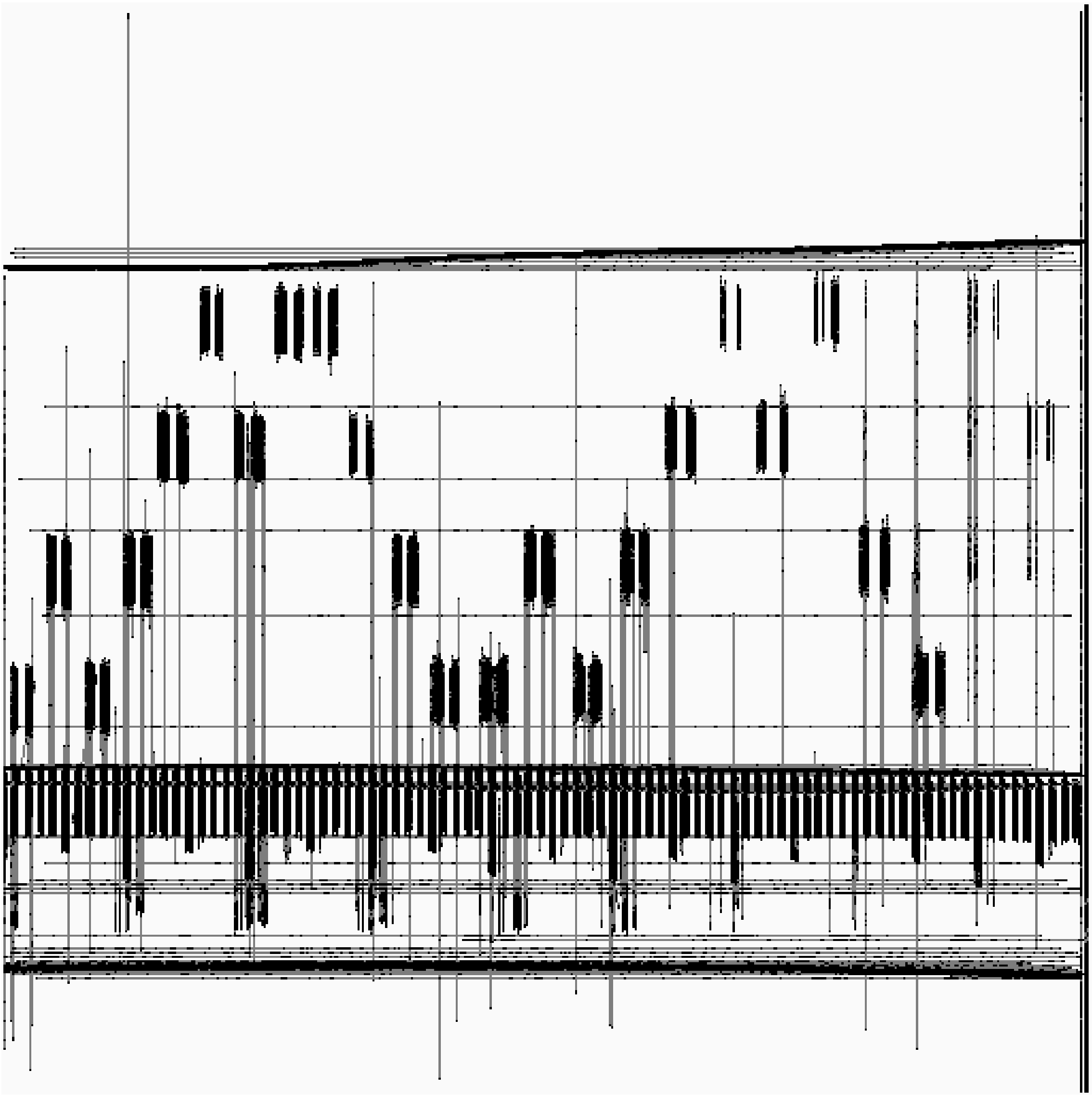

Figure 3: Image 1.a

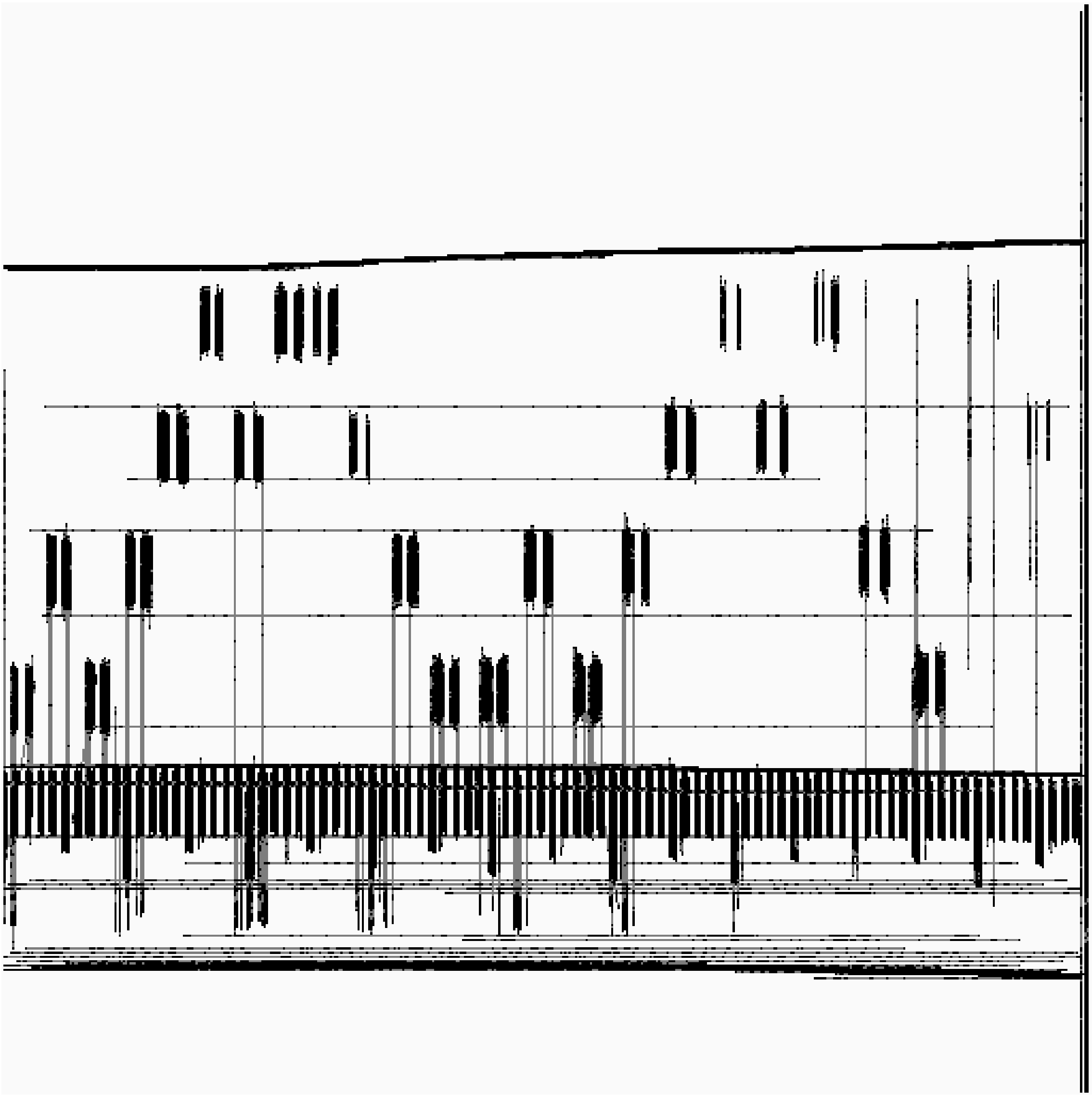

Figure 4: Image 1.b

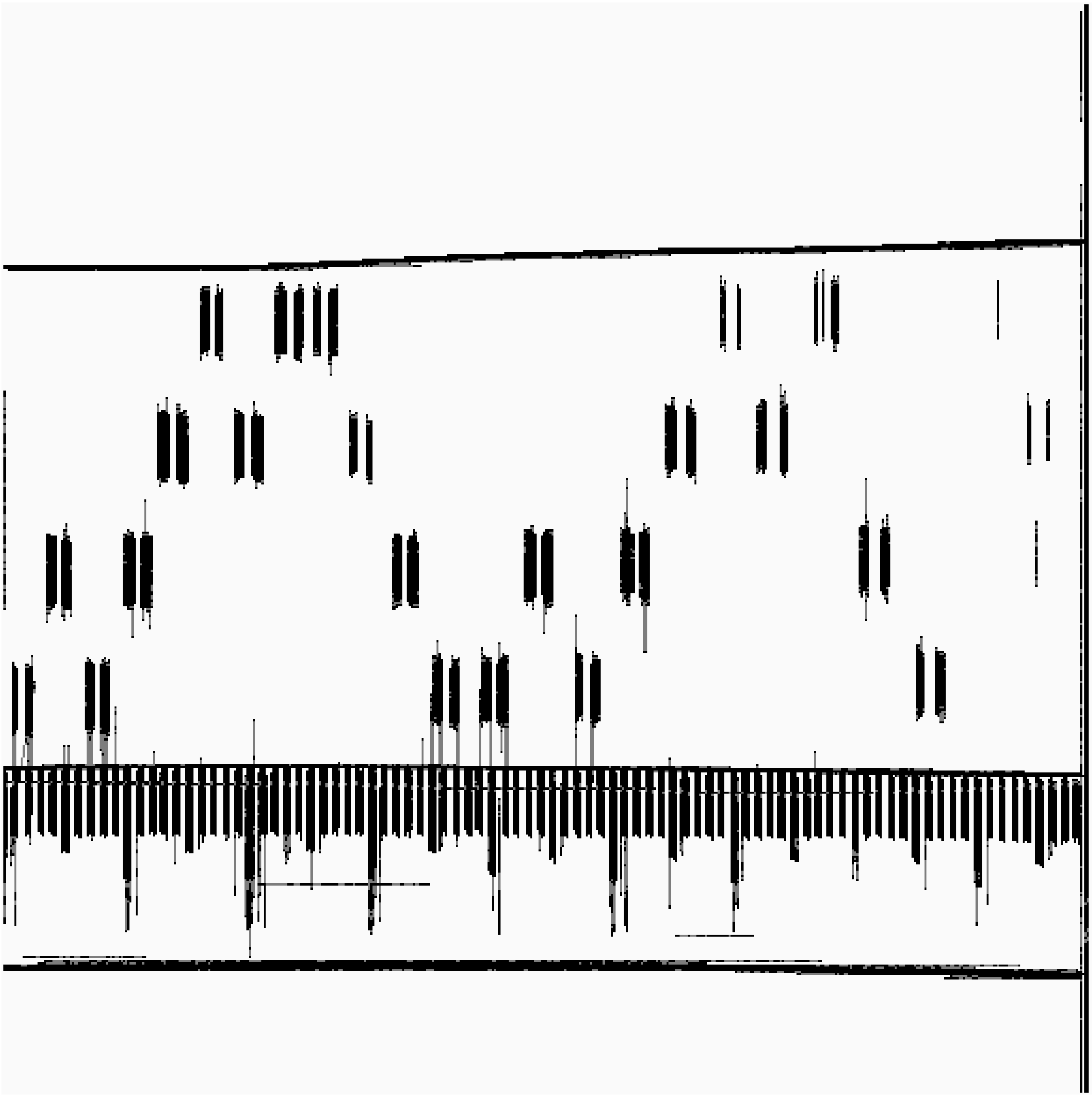

Figure 5: Image 1.c

Figure 6: Image 2

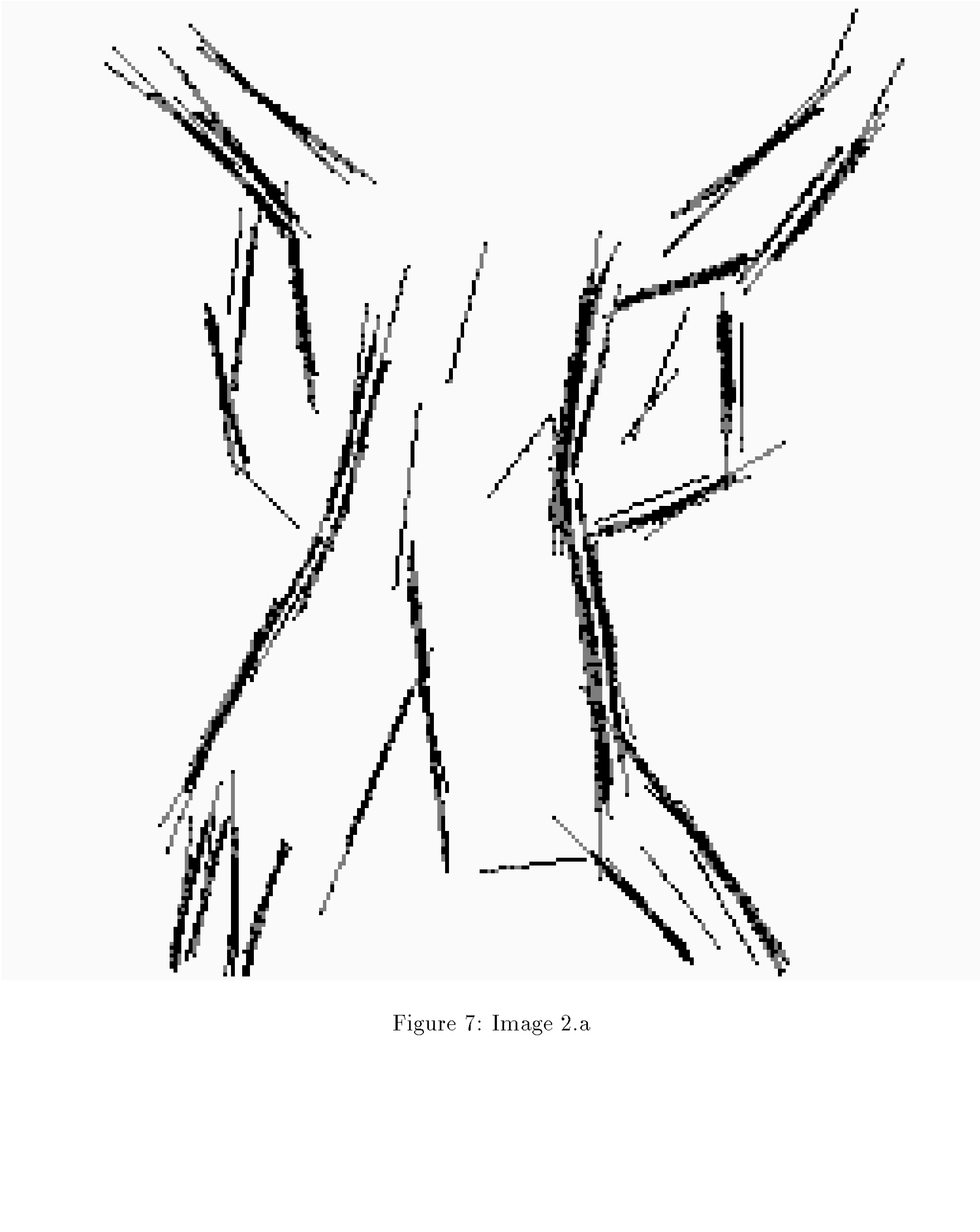

Figure 7: Image 2.a