International Journal of Computer Vision 40(1), 7–23, 2000

Meaningful Alignments

AGNÈS DESOLNEUX, LIONEL MOISAN AND JEAN-MICHEL MOREL
CMLA, ENS Cachan, 61 Avenue du Président Wilson, 94235 Cachan Cedex, France

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.

Keywords: alignment, Gestalt, Helmholtz, principle, probability

1. Introduction

Most theories of image analysis tend to find in a given image geometric structures (regions, contours, lines, convex sets, junctions, etc.). These theories generally assume that the images contain such structures and then try to compute their best description. The variational framework is quite well adapted to such a viewpoint (for a complete review, see e.g. Morel and Solimini, 1995). 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 Morel and Solimini, 1995), where the energy functional to be minimized is

$$E(u, K) = \lambda^2 \int_{\Omega-K} |\nabla u|^2\, dx + \mu\lambda^2 \operatorname{length}(K) + \int_{\Omega-K} (u - u_0)^2\, dx, \tag{1}$$

  where u is the estimated image, K its discontinuity set, and the result (u, K) is called a "segmentation" of u_0, i.e. a piecewise smooth function u with a set of contours K.
- The Bayesian model (see Geman and Geman, 1984; Geman and Graffigne, 1986): 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 conditional 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

$$\Pi(\vec{x}\,|\,\vec{y}) = \frac{\Pi(\vec{y}\,|\,\vec{x})\Pi(\vec{x})}{\Pi(\vec{y})}. \tag{2}$$

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

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

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

$$V(\vec{x}, \vec{y}) + U(\vec{x}). \tag{3}$$

A main drawback of all the variational methods is that they introduce normalization constants $(\lambda, \mu, \ldots)$

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 Rudin et al., 1992). 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. Actually, we should not be fair if we claimed that what we propose immediately yields a more reliable segmentation method. In fact, we only intend to point out the possibility of checking any proposed segmentation, by any segmentation method, from the point of view of meaningfulness. So far, this check will only be analysed in detail for straight boundaries: given a segmentation performed by any other method, we can, with the method proposed here, a posteriori decide about the meaninfulness of straight parts of the proposed boundaries.

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 Maitre, 1985), the detection of globally salient structures by Sha'Ashua and Ullman (see Sha'Ashua and Ullman, 1988), the Extension Field of Guy and Medioni (see Guy and Medioni, 1992) and the Parent and Zucker curve detector (see Parent and Zucker, 1989). 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. As a main example, let us describe the Hough transform. We assume that the image under analysis is made of dots which may create aligned patterns or not. We then compute for each straight line in the image, the number of dots lying on the line. In fact, the Hough transform describes a fast algorithm to do so. The result of the Hough transform is then a map associating with each line a number of dots. Then, "peaks" of the Hough transform may be computed: they indicate the lines which have more dots. Which peaks are significant? Clearly, a threshold must be used. For the today technology, this threshold generally is given by a user or learned. The Hough transform is nothing but a particular kind of "grouping".

According to Gestalt theory, "grouping" is the law of visual perception (see Kanizsa, 1997). Its main idea is that whenever points (or previously formed visual objects) have a characteristic in common, they get grouped and form a new, larger visual object, a "Gestalt". Some of the main grouping characteristics are colour constancy, "good continuation", alignment, parallelism, common orientation, convexity and closedness (for a curve), . . . In addition, the 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.

Our purpose is *not* to propose a new segmentation method. We rather propose a computational method to decide whether 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, one of the most basic Gestalt (see Wertheimer, 1923).

In this paper, we push the study to the end for the detection of alignments, but we will first give a general definition of what we will call "a meaningful event". Many of our statements will apply to other Gestalt as well. Our 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. In that sense, we shall say that it is a "proven event". The above informal definition immediately raises an objection: if we do probabilistic estimates in an image, this means that we have an a priori model. We are therefore losing any generality in the approach, unless the probabilistic model could be proven to be "the right one" for any image. In fact, we shall do statistical estimates, but related not to a model of the images but to a general model of perception. We shall apply the so called Helmholtz principle. This principle attempts to describe when perception decides to group objects according to some quality (colour, alignment, etc.). It can be stated in the following way. Assume that objects $O_1, O_2, \ldots, O_n$ are present in an image. Assume that k of them, say $O_1, \ldots, O_k$ have a common feature, say, same colour, same orientation, etc. We are then facing the dilemna: is this common feature happening by chance or is it significant? In order to answer this question, we make the following mental experiment: we assume that the considered quality has been randomly and uniformly distributed on all objects, i.e. $O_1, \dots O_n$. Notice that this quality may be spatial (like position, orientation); then we (mentally) assume that the observed position of objects in the image is a random realization of this uniform process. Then, we may ask the question: is the observed repartition probable or not?

The Helmholtz principle states that if the expectation in the image of the observed configuration $O_1, \dots, O_k$ is very small, then the grouping of these object makes sense, is a Gestalt.

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 $\varepsilon \ll 1$, 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 ε-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. Our term ε-meaningful is related to the classical p-significance in statistics; as we shall see further on, we must use expectations in our estimates and not probabilities.

The program we state here has been proposed several times in Computer Vision. We know of at least two instances: Lowe (1985) and Witkin-Tenenbaum (1983). 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.*" Some main points of the program we shall mathematically develop are contained in the preceding quotation: particularly the idea that significant geometric objects are the ones with small probability and the idea that this probability is anyway never zero because of the necessary lack of accuracy of observations in an image. Now, the preceding program is not accurate enough to give the right principles for computing Gestalt. The above mentionned example is e.g. not complete enough to be convincing. Indeed, we simply cannot fix a priori an event such as "these two lines are parallel" without merging it into the set of all events of the same kind, that is, all parallelisms. The space of straight lines in an image depends on the accuracy of the observations, but also on the size of the image itself. The fact that the mentionned probability be "low" (1/36) does not imply that few such events will occur in the image: we have to look for the number of possible pairs of parallel lines. If this number is large, then we will in fact detect many non-significant pairs of parallel lines. Only if the expected number of such pairs is much below 1, can we decide that the observed parallelism makes sense. Before proceeding to the mathematical theory, let us give some other toy example and discuss our definition of "ε-meaningfulness".

Example and Discussion: 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 events 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 $90 \cdot 90 \cdot (\frac{1}{2})^{100}$ 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 $(\frac{1}{2})^{100}$ 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.

5. How many Gestalt? If we make a list of "pregnant" Gestalt, following Gestalt theory, the longer the list, the higher the expectation of finding "false gestalt". Thus, perception, and also computer vision will at some time meet the following problem: to find the best trade off between number of Gestalt (which might be a priori as high as possible) and the false detection rate. For the time being, we shall not adress this problem; it will be adressed only when we are in a position to do a correct theory for many Gestalt!

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

$q = 2$). 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.

The computation of the gradient direction is based on an interpolation (we have $q = 2$). 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

$$\mathrm{Angle}(\mathrm{dir}(X), \mathrm{dir}(Y)) \leq \frac{2\pi}{n}. \tag{4}$$

In agreement with psychopysics and numerical experimentation, we consider that n should not exceed 16.

2.2. *Probabilistic Model*

According to the Helmholtz principle, our main assumption is following: we assume that the direction at all points in an image is a uniformly distributed random variable. In the following, we assume that $n > 2$ and we set $p = \frac{1}{n} < \frac{1}{2}$; p is the accuracy of the direction. We interpret p as the probability that two independent points have the "same" direction with the given accuracy p. In a structureless image, when two pixels are such that their distance is more than 2, 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. Alignments provide a more concrete way to understand Helmholtz principle. 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 highly correlated, while if they belong to two different objects, their directions should be independent (see the above quoted Lowe's program).

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 $q = 2$ having independent directions. Let A be a segment in the image made of l independent pixels (it means that the distance between two consecutive points of A is 2 and so, the real length of A is $2l$). We are interested in the number of points of A which have the property of having their direction aligned with the direction of A. Such points of A will simply be called *aligned points of A*.

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

$$\mathrm{P}[X_i = 1] = p \quad \text{and} \quad \mathrm{P}[X_i = 0] = 1 - p. \tag{5}$$

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

$$S_l = X_1 + X_2 + \cdots + X_l. \tag{6}$$

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

$$\mathrm{P}[S_l = k] = \binom{l}{k} p^k (1-p)^{l-k}. \tag{7}$$

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

$$\sum_{l=1}^{l_{\max}} m(l) = N^2(N^2-1) \simeq N^4. \tag{8}$$

The estimate N^4 is accurate enough, taking into account that what matters here will be its logarithm.

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

$$k(l) = \min\left\{k \in \mathbb{N}, \mathrm{P}\left[S_l \geq k\right] \leq \frac{\varepsilon}{N^4}\right\}. \tag{9}$$

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

$$\mathrm{P}\left[\chi_{e_i} = 1\right] = \mathrm{P}\left[S_{l_i} \geq k(l_i)\right]$$

where l_i is the length of the i-th segment. Notice that if l_i is small we may have $\mathrm{P}\left[S_{l_i} \geq k(l_i)\right] = 0$. Let R be the random variable representing the exact number of e_i occuring simultaneously in a trial. Since $R = \chi_{e_1} + \chi_{e_2} + \cdots + \chi_{e_{N^4}}$, the expectation of R is

$$\begin{aligned} E(R) &= E\left(\chi_{e_1}\right) + E\left(\chi_{e_2}\right) + \cdots + E\left(\chi_{e_{N^4}}\right) \\ &= \sum_{l=0}^{l_{\max}} m(l)\mathrm{P}\left[S_l \geq k(l)\right]. \end{aligned} \tag{10}$$

We compute here the expectation of R but not its law because it depends a lot upon the relations of dependence between the e_i. The main point is that segments may intersect and overlap, so that the e_i events are not independent, and may even be strongly dependent.

By definition we have

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

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

This notion of ε-meaningful segments has to be related to the classical "α-significance" in statistics, where α is simply ε/N^4. The difference which leads us to have a slightly different terminology is following: we are not in a position to assume that the segment detected as ε-meaningful are independent in anyway. Indeed, if (e.g.) a segment is meaningful it may be contained in many larger segments, which also are ε-meaningful. Thus, it will be convenient to compare the number of detected segments to the expectation of this number. This is not exactly the same situation as in failure detection, where the failures are somehow disjoint events. See Remark (*) below. This means that ϵ is an absolute parameter, not depending upon the size of the image, but only on the number of false detections which the user allows. Of course, if the image is larger, it may be expected that an increasing number of false detections should be allowed. However, by fixing ϵ always smaller than one, we decided not to take this opportunity. Our proposed definition of meaningfulness is also related to the statistical analysis of functional medical images (fMRI, PET) by Statistical Parameter Map (SPM), with two main differences, however. The first one is this: in the recent work of Stuart Clare (FMRIB center, Oxford, see Clare (1997)), and in the works of Friston et al. (1991) and Forman et al. (1995), an hypothesis testing method against white noise is performed in time series. As in the present work, the binomial law appears and a careful account of the effect of filtering on the number of effective degrees of freedom: this leads e.g. S. Clare to divide this number by three after a small gaussian filtering and is related to our decision of considering only nets of points at a distance larger than 2. S. Clare does as we do; he p-tests against the white noise assumption and admits a p-value of 0.005 by patient. Here is the main difference: the number of patients, and the length of the data are not taken into account in the test. In particular, the time length of the test is of course just enough to perform a significant test and the p-value is a threshold "per patient". In our case, we have two factors: the first one is that the number of "patients" is huge. Thus, with a p-test, the expectation of false detections would be much above 1, which is what we avoid by imposing ϵ much smaller than 1 and by entering into the computation the number of segments N^4. This is why we compute an expectation and not a probability: we have too many and not independent trials. The reason for introducing expectation here is the non independence (contrarily to patients) and the huge number of trials, increasing with the image size.

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 $m(l) \cdot [S_l \geq k'(l)] \leq \varepsilon$. 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 A be a segment of length l_0 with at least k_0 points having their direction aligned with the direction of A. We define the number of false alarms of A as

$$\begin{aligned} NF(k_0, l_0) &= N^4 \cdot \mathrm{P}\left[S_{l_0} \geq k_0\right] \\ &= N^4 \cdot \sum_{k=k_0}^{l_0} \binom{l_0}{k} p^k (1-p)^{l_0-k}. \quad (11) \end{aligned}$$

Interpretation of this definition: the number $NF(k_0, l_0)$ of false alarms of the segment A 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 A be a segment and $NF(k_0, l_0)$ its number of false alarms. Then A is ε-meaningful if and only if $NF(k_0, l_0) \leq \varepsilon$, 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 A, 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 $10 = E(R) \geq 100 \cdot \mathrm{P}[R = 100]$. 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.*

This proposition is a consequence of the definition and properties of the binomial distribution (see Feller, 1968).

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

$$E(S_l) = \sum_{i=1}^{l} E(X_i) = \sum_{i=1}^{l} \mathrm{P}[X_i = 1] = p \cdot l. \quad (12)$$

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

$$k_0 \geq pl_0 + (1 - p). \tag{13}$$

This is a "sanity check" for the model.

4. Thresholds

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

$$P(k, l) = \mathrm{P}\,[S_l \geq k] = \sum_{i=k}^{l} \binom{l}{i} p^i (1 - p)^{l-i}. \tag{14}$$

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

$$k(l) = \min\left\{k \in \mathbb{N},\ \mathrm{P}\,[S_l \geq k] \leq \frac{\varepsilon}{N^4}\right\}. \tag{15}$$

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

$$p^l \leq\ \mathrm{P}\,[S_l \geq k(l)] \leq \frac{\varepsilon}{N^4}, \tag{16}$$

so that

$$l \geq \frac{-4 \ln N + \ln \varepsilon}{\ln p}. \tag{17}$$

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 $l_{\min} = 9$.

We can also give estimates of the thresholds $k(l)$. The mathematical theorems are given in the Appendix. They roughly say that

$$k(l) \simeq pl + \sqrt{C \cdot l \cdot \ln \frac{N^4}{\varepsilon}}, \tag{18}$$

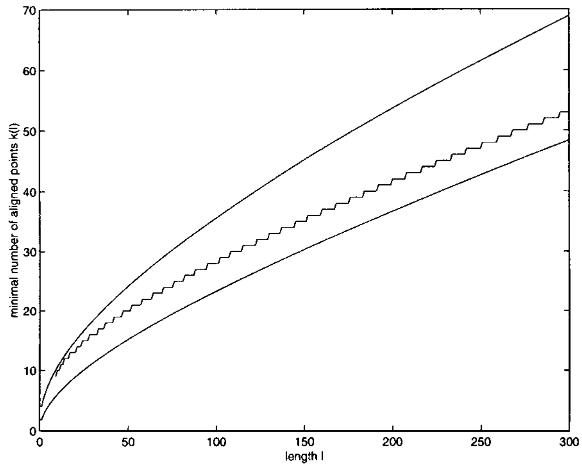


Figure 1. Estimates for the threshold of meaningfulness $k(l)$. The middle (stepcase) curve represents the exact value of the minimal number of aligned points $k(l)$ to be observed on a 1-meaningful segment of length l in an image of size 512, for a direction precision of $1/16$. The upper and lower curves represent estimates of this threshold obtained by Proposition 5 and Proposition 7 (see Appendix).

where $2p(1 - p) \leq C \leq 1/2$. Some of these results are illustrated by Fig. 1. These estimates are not necessary for the algorithm (because $\mathrm{P}\,[S_l \geq k]$ is easy to compute) but they provide an interesting order of magnitude for $k(l)$.

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. $NF(S) \ll 1$). 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: $\forall B \subset A, NF(B) \geq NF(A)$,
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.

Proposition 3 (*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.*

These elementary properties are simple consequences of Proposition 1.

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

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

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

An interesting application of this proposition is the concatenation of meaningful segments. Let $A = (k, l)$ and $B = (k', l')$ be two meaningful segments lying on the same line. Moreover we assume that A and B are consecutive, so that $A \cup B$ is simply the segment $(k + k', l + l')$. Then, since

$$\frac{k + k'}{l + l'} \geq \min\left(\frac{k}{l}, \frac{k'}{l'}\right),$$

we deduce, thanks to the above proposition, that

$$NF(A \cup B) < \max(NF(A), NF(B)). \qquad (19)$$

This shows that the concatenation of two meaningful segment is a meaningful segment.

6. A Conjecture about Maximality

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 that two distinct maximal segments on a same straight line have no common point.

Conjecture 1. If for $i = 1, 2, 3$, k_i and l_i are integers such that $l_i \neq 0$ and $k_i \leq l_i$, then

$$\begin{aligned} &\min(p, P(k_1, l_1), P(k_1 + k_2 + k_3, l_1 + l_2 + l_3)) \\ &\quad < \max_{i \in \{2,3\}} P(k_1 + k_i, l_1 + l_i) \end{aligned} \qquad (20)$$

This conjecture can be deduced from a stronger (but simpler) conjecture: the concavity in a particular domain of the level lines of a natural continuous extension of P involving the incomplete Beta function.

Corollary 1 (*Union and Intersection***).** *Suppose that Conjecture* 1 *is true. Then, if A and B are two segments on the same straight line such that $A \not\subseteq B$ and $B \not\subseteq A$, one has*

$$\begin{aligned} &\min(pN^4, NF(A \cap B), NF(A \cup B)) \\ &\quad < \max(NF(A), NF(B)). \end{aligned} \qquad (21)$$

This is a direct consequence of Conjecture 1. Numerically, we checked this property for all segments A and B such that $|A \cup B| \leq 256$. For $p = 1/16$, we obtained

$$\begin{aligned} &\min_{|A \cup B| \leq 256} \\ &\times \frac{\max((NF(A), NF(B)) - \min(pN^4, NF(A \cap B), NF(A \cup B))}{\max((NF(A), NF(B)) + \min(pN^4, NF(A \cap B), NF(A \cup B))} \\ &\simeq 0.000754697\ldots > 0, \end{aligned}$$

this minimum (independent 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)).

Notice also that Conjecture 1 can be proven when $P(k,l)$ is replaced by its approximation by the Gaussian law (asymptotic estimate when $k \simeq pl$)

$$G(k,l) = \frac{1}{\sqrt{2\pi}} \int_{\alpha(k,l)}^{+\infty} e^{-\frac{x^2}{2}}\, dx \quad \text{where} \qquad \alpha(k,l) = \frac{k - pl}{\sqrt{lp(1-p)}} \tag{22}$$

or by its Large Deviation estimate (asymptotic estimate when $l \to +\infty$ and $\frac{k}{l} > r > p$),

$$H(k,l) = \exp\left(k \ln p + (l-k)\ln(1-p) - k \ln \frac{k}{l} - (l-k)\ln \frac{l-k}{l} \right). \tag{23}$$

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

Remark. The numerical checking of Corollary 1 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.

7. Experiments

In all the following experiments, the direction at a pixel in an image is computed on a 2×2 neighborhood with the method described in section 2.1 ($q = 2$) and the precision is $p = 1/16$.

The direction is computed at all pixels, unless the gradient is strictly equal to zero (up to machine precision). Let N denote the size of the considered image. The algorithm used to find the meaningful segments is the following. For each one of the four sides of the image, we consider for each pixel of the side the lines starting at this pixel, and having an orientation multiple of $\pi/48$. And then on each line, we compute the meaningful segments. For each segment, let l be its length counted in independant pixels (which means that the real length of the segment is $2l$), then among the l points we count the number k of points having their direction aligned with the direction of the segment (with the precision p), and finally we compute $P(k,l)$: if it is less than $\frac{1}{48N^3} \times \frac{1}{10}$, we say that the segment is meaningful. The value $48N^3$ is an estimate of the number of considered segments and we took $\varepsilon = 1/10$. Because of the angle precision $2\pi/16$ (to be compared with $\pi/48$), the sampling of directions is enough to cover all possible alignments in a 512×512 image. Motice that $P(k,l)$ can be simply tabulated at the begining of the algorithm using Newton's law $P(k+1, l+1) = pP(k,l) + (1-p)P(k+1,l)$.

It must be made clear that we applied *exactly* the same algorithm to all presented images, which have very different origins. The only parameter of the algorithm is precision. We fixed it equal to $1/16$ in all experiments; this value corresponds to the very rough accuracy of 22.5 degrees; this means that (e.g.) two points can be considered as aligned with, say the 0 direction if their angles with this direction are up to ± 11.25 degrees! It is clear that these bounds are very rough, but in agreement with the more pessimistic estimates for the vision accuracy in psychophysics and the numerical experience as well. Moreover, in all experiments, we only keep the meaningful segments having in addition the property that their endpoints have their direction aligned with the one of the segment: black points represent points on a meaningful segment which have the same direction as the one of the segment (with the precision p), and gray points represent points on a meaningful segment which do not have the same direction as the segment.

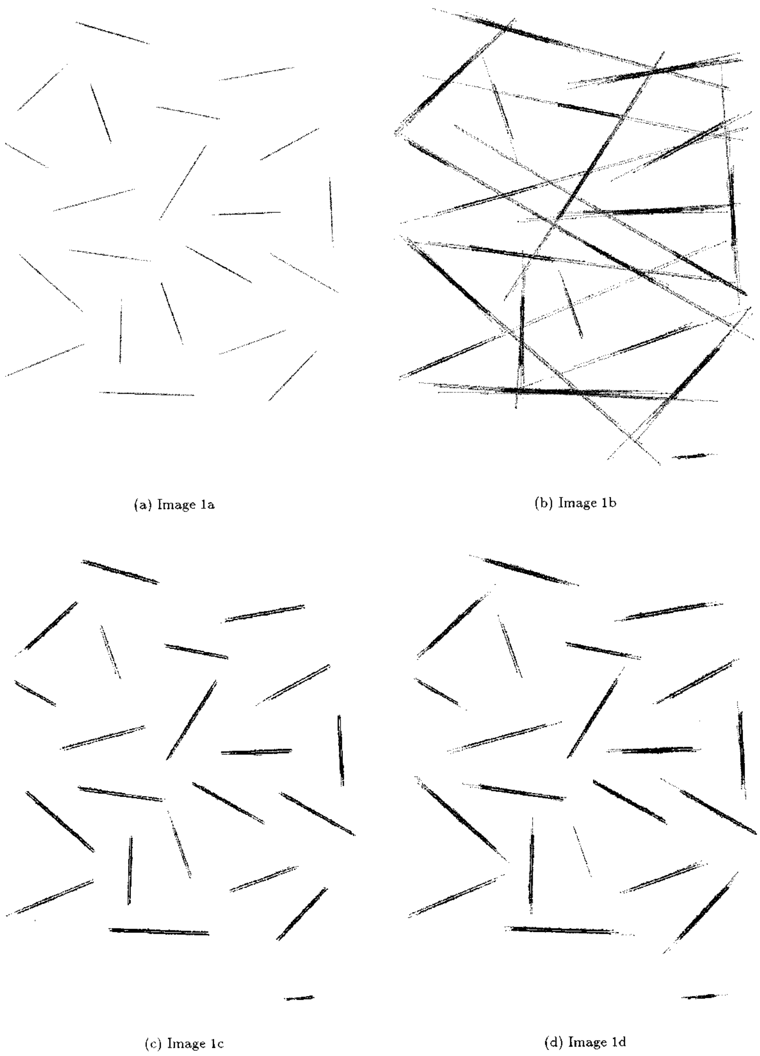
For each one of the following images, we comput

1. all the meaningful segments.
2. the maximal meaningful segments.
3. for some of them: meaningful segments with length less than 30 or 20. These segments have a small length (close to the minimal length $l_{\min} = -4 \ln N / \ln p$), and consequently a density of aligned points close to 1.

Typical CPU time for a 512×512 image is ten seconds, and one second for a 256×256 image. As a general comment to all experiments, we shall see that the (non maximal) meaningful events are too long: indeed, if we find a very meaningful segment (and this happens very systematically in the experiments), then much larger segments containing this very meaningful one will still be meaningful. We display, for a sake of completeness, several images with all meaningful alignments. In continuation, we display the maximal meaningful alignments, as a way to check by comparison that these maintain the whole alignment information, and are by far more accurate. We think the experiments clearly demonstrate the necessity of maximality. We also diplay in several images the only alignments whose length is smaller than a given threshold (20 or 30). This is a way to check that, in "natural" images, most alignments can be locally detected. Indeed, we see that most maximal detected alignements are a concatenation of small, still meaningful, alignments.

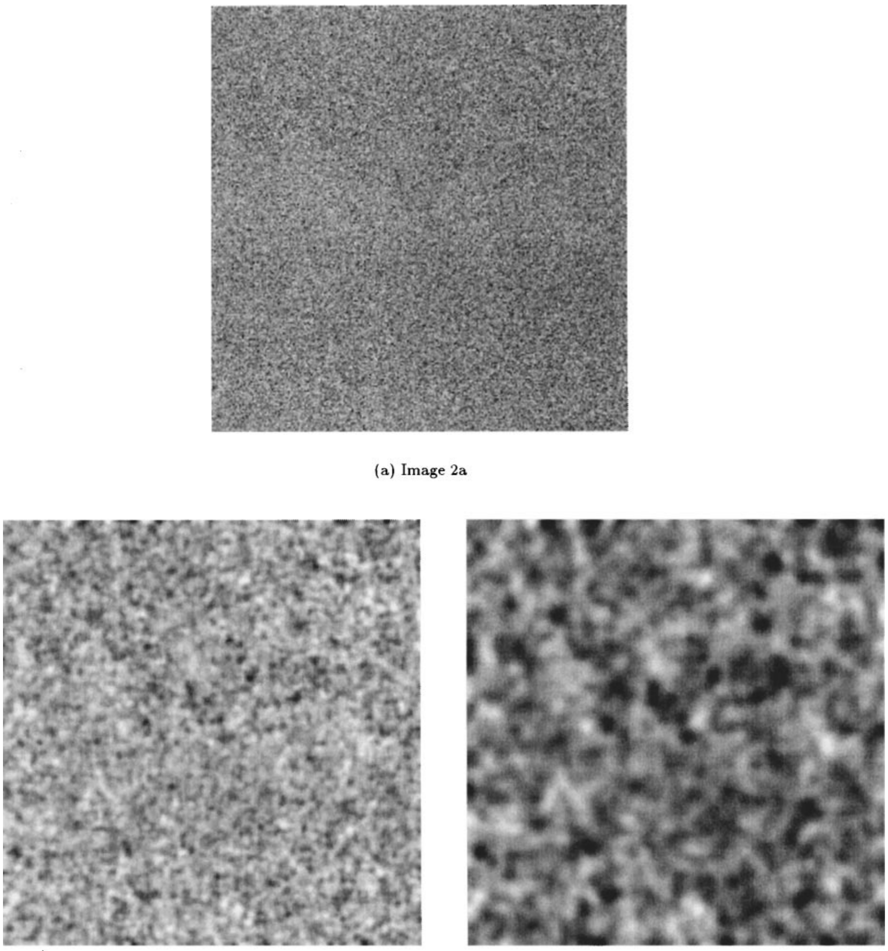
Image 1: Pencil strokes (Fig. 2). This digital image was first drawn with a ruler and a pencil on a standard A4 white sheet of paper, and then scanned into a 478 $\times$ 598 digital image (image 1a); the scanner's apparent blurring kernel is about two pixels wide and some aliasing is perceptible, making the lines somewhat blurry and dashed. Two pairs of pencil strokes are aligned on purpose. We display in the first experiment all meaningful segments (image 1b). Four phenomena occur, which are very apparent in this simple example, but will be perceptible in all further experiments.

1. Too long meaningful alignments: we commented this above; clearly, the pencil strokes boundaries are very meaningful, thus generating larger meaningful segments which contain them.
2. Multiplicity of detected segments. On both sides of the strokes, we find several parallel lines (reminder: the orientation of lines is modulo 2π). These parallel



(a) Image 1a (b) Image 1b (c) Image 1c (d) Image 1d

Figure 2. Pencil strokes.



(a) Image 2a

(b) Image 2b

(c) Image 2c

Figure 3. White noise blurred images.

lines are due to the blurring effect of the scanner's optical convolution. Classical edge detection theory would typically select the best, in terms of contrast, of these parallel lines.

3. Lack of accuracy of the detected directions: We do not check that the directions along a meaningful segment be distributed on both sides of the lines direction. Thus, it is to be expected that we detect lines which are actually slanted with respect to the edge's "true" direction. Typically, a blurry edge will generate several parallel and more or less slanted alignements. It is not the aim of the actual algorithm to filter out this redundant information; indeed, we do not know at this point whether the detected parallel or slanted alignments are due to an edge or not: this must be the object of a more complex algorithm. Everything indicates that an edge is no way an elementary phenomenon in Gestalt.

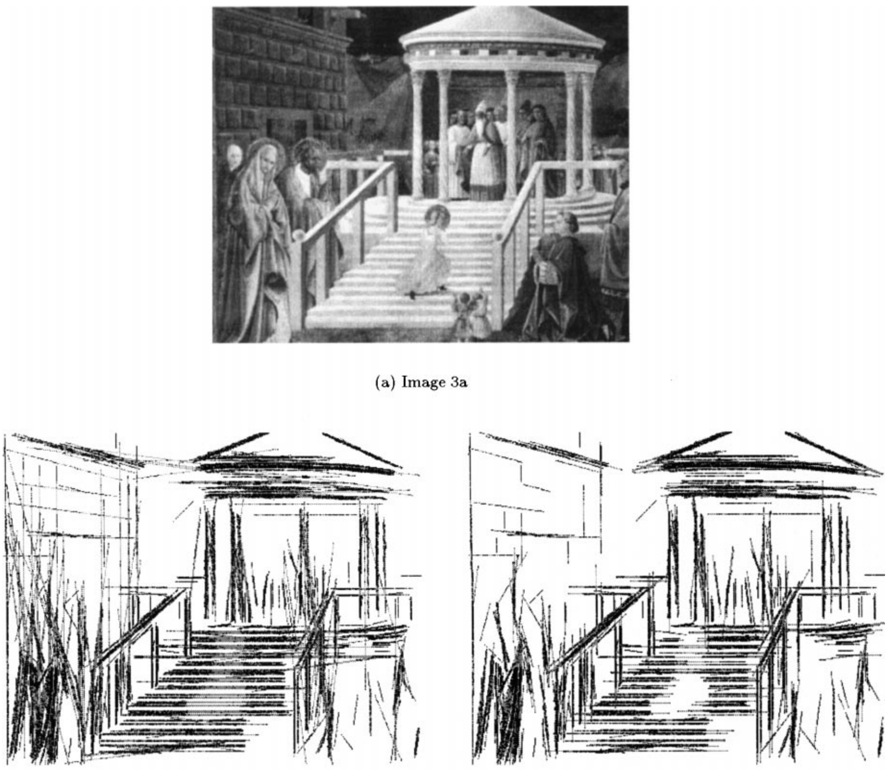
We display in the second experiment for this image all maximal meaningful segments (image 1c), which show for each stroke two bundles of parallel lines on each side of the stroke. In the third one, we display all meaningful segments whose length is less than 60 pixels (image 1d). This achieves a kind of localization of the segments. Now, a visual comparison between this experiment and the former one (1c) shows that maximality achieves a better, more accurate localization. Thus, we will not show the "small segments" in all experiments to follow.

Image 2: White noise blurred images (Fig. 3). Image 2a is a white noise, all pixels values being independent and identically distributed with a gaussian law. Image 2b is Image 2a convolved with a gaussian kernel with standard deviation 4 pixels and Image 2c is Image 2a convolved with a gaussian kernel with standard deviation 16 pixels. We apply the same algorithm as before to all of these images. The outcome was for all of three: no alignement detected! This experiment was devised to show that the local independence of pixels can be widely violated without affecting the final outcome. Indeed, a blurring creates local alignments but not global ones.

Image 3: Uccello's painting (Fig. 4). This image (3a) is a result of the scan of an Uccello's painting: "Presentazione della Vergine al tempio" (from the book "L'opera completa di Paolo Uccello", Classici dell'arte, Rizzoli). In image 3b we display all maximal meaningful segments and in image 3c all meaningful segments with length less than 60. Notice how maximal segments are detected on the staircase in spite of the occlusion by the going up child. Compare with the small meaningful segments. All remarks made in Image 1 apply here (parallelisms due to the blur, etc.)

Image 4: Airport image (Fig. 5). This digital image also has a noticeable aliasing which creates horizontal and vertical dashes along the edges. We display in image 4b all maximal detectable segments, always for $\varepsilon = 1/10$. We compare in image 4c and 4d with the same image with $\varepsilon = 1/100$ and $\varepsilon = 1/1000$.

Image 5: A road (courtesy of INRETS) (Fig. 6). We display all maximal meaningful segments (image 5b) and all meaningful segments with length less than 60 (image 1c). Notice the detected horizontal lines in 5b: they correspond to "horizon lines", that is, lines parallel to the horizon. They tend to accumulate towards the horizon of the image. Such lines correspond to nonlocal alignments (they are not present in Image 5c). They are due to a perspective effect: all visual objects on the road (shadows, spots, etc.) are seen in very slanted view. Thus, their contours are mostly parallel

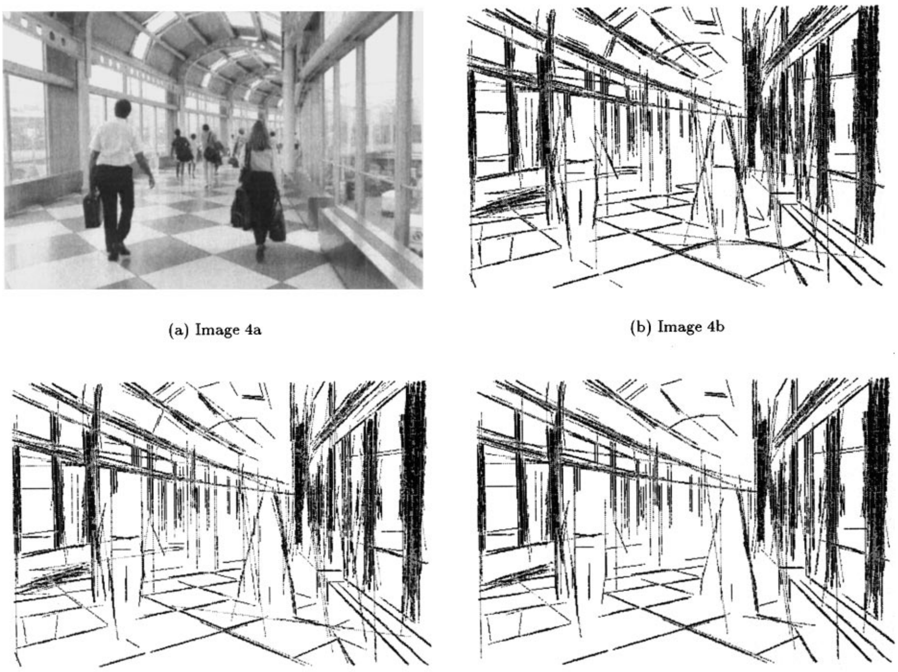


(a) Image 3a

(b) Image 3b

(c) Image 3c

Figure 4. Uccello's painting.



(a) Image 4a

(b) Image 4b

(c) Image 4c

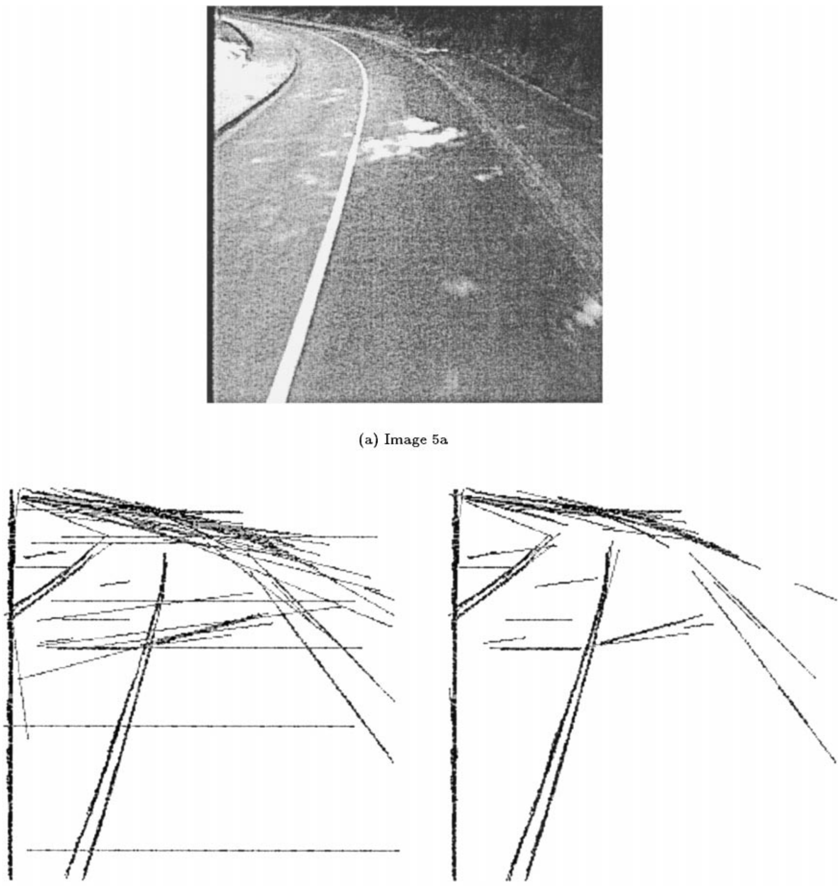
(d) Image 4d

Figure 5. Airport image.

to the horizon, thus generating what we should call "perspective alignements".

8. Conclusion

This preliminary study about Gestalt has tried to build the correct mathematical framework for the widespread idea that significant geometric structures in an image correspond to very low probability events. They are two ways to interpret this statement: the wellspread one is to define a probabilistic functional which is minimized, thus yielding the most likely geometric structures. Now, we emphasized the fact that the detection of structure has an intermediate stage, clearly missed in the variational framework: before we look for the most likely structures, we have to make a list of all proven structures. Experiments show well the difference between both approaches: where edge detection algorithm (which always look for the best position for an edge) directly yield a single edge, we find multiple alignments. In many cases, it is plain from the experiments that edge detection could be interpreted as a selection procedure among the alignments. To summarize, we have two different qualities which are mixed in the variational framework: the feasibility and the optimality. By looking for optimality only, we forget to prove that the found, optimal structures indeed exist. Next, we proposed an alternative to global variational principles: the notion of maximal event. In some extent, maximal alignments are local minimizers of a probability functional. The main difference is first that we do a minimization among feasible structures only; second, that we get aditional structure properties from maximality, as the fact that maximal alignments do not intersect. It may well be asked at that point what we can further do. We have considered one Gestalt quality only: the alignment. A first question is: to which other qualities the notions developped here apply? We do not intend to give here a detailed answer. We will develop this general viewpoint in a further work. A second question which was raised by Lowe is the combination of several Gestalt qualities to generate more elaborate geometric structures. Edge detection is such an elaborate geometric structure: it is a combination of alignment (or curviness), of contrast along the edge curve, of homogeneity on both sides, of maximality of the slope and finally of stability accross scales! (see the scale space theory). All of these criteria contribute to more and more sophisticated edge detectors. In this paper, we have shown that one of the



(a) Image 5a

(b) Image 5b

(c) Image 5c

Figure 6. A road.

qualities involved, the alignment, can be proved separately. The other qualities can receive an analogous, if not sometimes identical theory of meaningfulness. Now, the question of how we should let such qualities collaborate seems open.

A. Appendix

In the following, ε and p are fixed numbers smaller than 1. We recall the notation

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

We also 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

$$k(l) = \min\left\{k \in \mathbb{N},\, P(k,l) \leq \frac{\varepsilon}{N^4}\right\}. \qquad (24)$$

A.1. Sufficient Condition of Meaningfulness

In this appendix, 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 Hoeffding, 1963).

Theorem 2 (*Hoeffding's inequality*). *If k, l are positive integers with $k \leq l$, and if p is a real number such that $0 < p < 1$.*

Then if $r = k/l \geq p$, we have the inequalities

$$\begin{aligned} P(k,l) &\leq \exp\left(lr \ln \frac{p}{r} + l(1-r)\ln\frac{1-p}{1-r}\right) \\ &\leq \exp(-l(r-p)^2 h(p)) \\ &\leq \exp(-2l(r-p)^2), \end{aligned} \qquad (25)$$

where h is the function defined on]0, 1[*by*

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

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

$$k \ge pl + \sqrt{\frac{4 \ln N - \ln \varepsilon}{h(p)}} \sqrt{l}, \tag{26}$$

then A is ε-meaningful.

Notice that Proposition 5 is interesting only when

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

that is when

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

Numerical example: for $\varepsilon = 1$, $N = 512$ and $p = 1/16$, we obtain $l \ge 10$.

A.2. Necessary Conditions for Meaningfulness

We use a comparison between the Binomial and the Gaussian laws given by the following

Theorem 3 (Slud, 1977)**.** *If* $0 < p \le 1/4$ *and* $pl \le k \le l$, *then*

$$\mathrm{P}\,[S_l \ge k] \ge \frac{1}{\sqrt{2\pi}} \int_{\alpha(k,l)}^{+\infty} e^{-x^2/2}\, dx \quad \text{where}$$
$$\alpha(k,l) = \frac{k - pl}{\sqrt{lp(1-p)}}, \tag{27}$$

Proposition 6 (*necessary condition of meaningfulness*)**.** *We assume that* $0 < p \le 1/4$ *and N are fixed. If a segment* $S = (k, l)$ *is ε-meaningful then*

$$k \ge pl + \alpha(N) \sqrt{lp(1-p)}, \tag{28}$$

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

$$\frac{1}{\sqrt{2\pi}} \int_{\alpha(N)}^{+\infty} e^{-x^2/2}\, dx = \frac{\varepsilon}{N^4}. \tag{29}$$

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

A.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 Feller, 1968).

Theorem 4 (*De Moivre-Laplace limit theorem*)**.** *If α is a fixed positive number, then as l tends to* $+\infty$,

$$\mathrm{P}\,[S_l \ge pl + \alpha\sqrt{l \cdot p(1-p)}] \longrightarrow \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{+\infty} e^{-x^2/2}\, dx. \tag{30}$$

Our aim is to get the asymptotic behaviour when l is large of the threshold $k(l)$ defined by (15). 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 $l \le \sqrt{2}N$). 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 Feller, 1968).

Theorem 5 (Feller)**.** *If* $\alpha(l) \to +\infty$ *and* $\alpha(l)^6/l \to 0$ *as* $l \to +\infty$, *then*

$$\mathrm{P}\left[S_l \ge pl + \alpha(l)\sqrt{l \cdot p(1-p)}\right] \sim \frac{1}{\sqrt{2\pi}} \int_{\alpha(l)}^{+\infty} e^{-x^2/2}\, dx. \tag{31}$$

Proposition 7 (*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*

$$k(l) = pl + \sqrt{2p(1-p) \cdot l \cdot \left(\ln \frac{N^4}{\varepsilon} + O(\ln \ln N) \right)}. \tag{32}$$

This proposition shows that the lower estimate given in Proposition 6 in fact gives the right asymptotic estimate. The condition $l/(\ln N)^3$ does not make much sense for the considered values of N (about 1000). Nonetheless, Proposition 7 conforts the previous ones.

Acknowledgments

Work supported by Office of Naval Research under grant N00014-97-1-0839. We thank Jean Bretagnolle, Nicolas Vayatis, Frédéric Guichard and Isabelle Gaudron-Trouvé for valuable suggestions.

References

Clare, S. 1997. Ph.D. Thesis. http://www.fmrib.ox.ac.uk/∼stuart.

Feller, W. 1968. *An Introduction to Probability Theory and Its Applications, Vol. I*, 3rd edn. Wiley: New York.

Friston, K.J., Frith, C.D., Liddle, P.F., and Frackowiak, R.S.J. 1991. Comparing functional (PET) images: The assessment of significant change. *J. Cerb. Blood Flow Metab.*, 11:690–699.

Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintum, M.A., and Noll, D.C. 1995. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magn. Reson. Med.*, 33:636–647.

Geman, S. and Geman, D. 1984. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Machine Intell.*, 6:721–741.

Geman, S. and Graffigne, C. 1986. Markov random field image models and their applications to computer vision. In *Proceedings of the International Congress of Mathematicians*, Berkeley, USA.

Guy, G. and Medioni, G. 1996. Inferring global perceptual contours from local features. *International Journal on Computer Vision*, 20(1/2):113–133.

Hoeffding, W. 1963. Probability inequalities for sum of bounded random variables. *Journal of the American Statistical Association*, 58(58):13–30.

Kanizsa, G. 1997. *La Grammaire du Voir*. Editions Diderot, arts et sciences.

Lowe, D. 1985. *Perceptual Organization and Visual Recognition*. Kluwer Academic Publishers: Dordrecht, The Netherland.

Maitre, H. 1985. Un panorama de la transformation de Hough. *Traitement du Signal*, 2(1):305–317.

Morel, J.-M. and Solimini, S. 1995. *Variational Methods In Image Segmentation*. Birkhähauser: Cambridge, MA.

Parent, P. and Zucker, S.W. 1989. Trace inference, curvature consistency and curve detection. *IEEE Trans. Pattern Anal. Machine Intell.*, 11(8):823–839.

Rudin, L., Osher, S., and Fatemi, E. 1992. Nonlinear total variation based noise removal algorithms. *Physica D*, 60(1–4):259–268.

Sha'Ashua, A. and Ullman, S. 1998. Structural saliency: The detection of globally salient structures using a locally connected network. *Proceedings of the 2nd Int. Conf. on Computer Vision*, pp. 321–327.

Wertheimer, M. 1923. Untersuchungen zur Lehre der Gestalt, II. *Psychologische Forschung*, 4:301–350.

Witkin, A.P. and Tenenbaum, J.P. 1983. *On the Role of Structure in Vision*, Beck Hope and Rosenfeld (Eds.). Academic Press: New York, pp. 481–543.