Gestalt theory and Computer Vision

Agnès Desolneux, Lionel Moisan and Jean-Michel Morel
CMLA, ENS Cachan,
61 av. du président Wilson,
94235 Cachan cedex, FRANCE.

Abstract

In the Wertheimer programme there are two kinds of organizing laws. The first kind are grouping laws which, starting from atomic and local data, recursively construct larger groups in the perceived image. Each grouping law focuses on a single quality (color, shape, direction,...). The second kind are principles governing the collaboration and conflicts of gestalt laws. We explain why and how the Wertheimer programme can be translated into a Computer Vision programme. This translation is not straightforward, since Gestalt theory did not address two fundamental matters : image sampling and image information measurements. Using them, we show that a quantitative version of Kanizsa's "masking in texture" can lead to numerical simulations of several grouping laws. As illustration, we display numerical experiments on several digital images. Analyzing successes and failures of these experiments, we conclude that Computer Vision is currently facing exactly the same challenges as were discussed in Gestalt theory, namely the collaboration and conflicts between grouping laws, and the masking issue.

1 Introduction

The geometric Gestalt theory started in 1921 with Max Wertheimer's founding paper [31]. In its 1975 last edition, the Gestalt Bible *Gesetze des Sehens* by Wolfgang Metzger [22] gave a broad overview of the results of a fifty years research. At about the same date, Computer Vision was an emerging new discipline, at the meeting point of Artificial Intelligence and Robotics. The foundation of signal sampling theory by Claude Shannon [28] was already twenty years old, but computers were able to deal with images with some efficiency only at the beginning of the seventies. Two things are noticeable:

- Computer Vision did not use at first the Gestalt theory results : the founding book of David Marr [21] involves much more neurophysiology than phenomenology. Also, its programme and the robotics programme [12] founded their hopes on binocular stereo vision. This was in total contradiction to, or ignorance of the results explained at length in Metzger's chapters on *Tiefensehen*. Indeed, these chapters demonstrate that binocular stereo vision is a *parent pauvre* in human depth perception.

- Conversely, Shannon's information theory does not seem to have influenced Gestalt research, as far as we can judge from Kanizsa's and Metzger's books. The only bright exception is Attneave's attempt [2] to give a shape sampling theory adapted to shape perception.

This lack of initial interaction is surprising. Indeed, both disciplines have attempted to answer the following question : how to arrive at global percepts (let them be visual objects of gestalts[1]) from the local, atomic information contained in an image ?

In this paper, we shall propose an analysis of the Wertheimer programme adapted to computational issues. We shall distinguish two kinds of laws :
- the practical grouping laws (like vicinity or similarity), whose aim it is to build up *partial gestalts*;
- the gestalt principles like masking or *articulazione senza resti*, whose aim it is to operate a synthesis between the partial groups obtained by the elementary grouping laws. We shall review some recent methods proposed by the authors of the present paper in the computation of partial gestalts (groups obtained by a single grouping law). These results show that

- there is a simple computational principle (the so-called Helmholtz principle), inspired from Kanizsa's *masking by texture*, which allows one to compute any partial gestalt obtainable by a grouping law (section 4). Also, a particular use of the *articulazione senza resti*, which we call maximality, yields optimal partial gestalts ;

- this computational principle can be applied to a fairly wide series of examples of partial gestalts, namely alignments, clusters, boundaries, grouping by orientation, size or grey level ;

- the experiments yield evidence that in natural world images, partial gestalts often collaborate.

We push one of the experiments to prove that the partial gestalt recursive building up can be led up to the third level (gestalts built by three successive grouping principles). In contrast, we also show by numerical counterexamples that <u>all</u> partial gestalts are likely to lead to wrong scene interpretations. As we shall see, the wrong detections are always explainable by a conflict between gestalts. We eventually show some experiment suggesting that Minimal Description Length principles [26] may be adequate to resolve some of the conflicts between gestalt laws.

Our plan is as follows. We start in section 2 with an account of Gestalt theory, centered on the initial 1923 Wertheimer programme. In section 3, we focus on the problems raised by the synthesis of groups obtained by partial grouping laws : we address the conflicts between these laws and the masking phenomenon, which we discuss in the line of Kanizsa. In section 4, we point out several quantitative aspects implicit in Kanizsa's definition of masking and show that one particular kind of masking, Kanizsa's *masking by texture*, suggests computational procedures. Such computational procedures are explained in section 5. We end this paper in section 6 by the discussion of a list of numerical experiments on digital images.

2 Grouping laws and gestalt principles

2.1 Grouping laws

Gestalt theory starts with the assumption of active "grouping" laws in visual perception (see [14], [31]). These groups are identifiable with subsets of the retina. We shall talk in the following of "points" or groups of points which we identify with spatial parts of the planar rough

[1]We shall write gestalt and treat it as an English word when we talk about gestalts as groups. We maintain the German uppercase for Gestalt theory.

percept. In image analysis, we shall identify them as well with the points of the digital image. Whenever points (or previously formed groups) have one or several characteristics in common, they get grouped and form a new, larger visual object, a *gestalt*. The list of elementary grouping laws given by Gaetano Kanizsa in *Grammatica del Vedere* page 45 and following [14] is *vicinanza, somiglianza, continuita di direzione, completamento amodale, chiusura, larghezza constante, tendenza alla convessita, simmetria, movimento solidale, esperienza passata*, that is : vicinity, similarity, continuity of direction, amodal completion, closure, constant width, tendency to convexity, symmetry, common motion, past experience. This list is actually very close to the list of grouping laws considered in the founding paper of Wertheimer [31]. These laws are supposed to be at work for every new percept. The amodal completion, one of the main subjects of Kanizsa's books, is, from the geometric viewpoint, a variant of the good continuation law[2]. Figure 1 illustrates many of the grouping laws stated above. The subjects asked to describe briefly such a figure give an account of it as "three letters X" built in different ways.

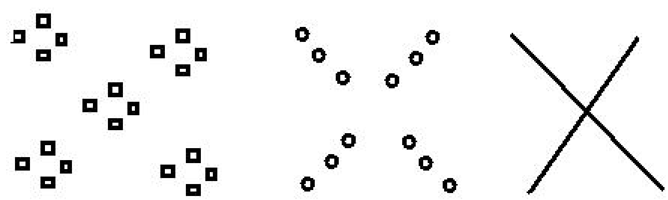

Figure 1: The building up of gestalt : *X-shapes. Each one is built up with branches which are themselves groups of similar objects; the objects, rectangles or circles are complex gestalts, since they combine color constancy, constant width, convexity, parallelism, past experience etc.*

Most grouping laws stated above work from local to global. They are of mathematical nature, but must actually be split into more specific grouping laws to receive a mathematical and computational treatment :

- *Vicinity* for instance can mean : connectedness (*i.e.* spots glued together) or clusters (spots or objects which are close enough to each other and apart enough from the rest to build a group). This vicinity gestalt is at work in all sub-figures of figure 2.
- *similarity* can mean : similarity of color, shape, texture, orientation,... Each one of these gestalt laws is very important by itself (see again figure 2).
- *continuity of direction* can be applied to an array of objects (figure 2 again). Let us add to it alignments as a grouping law by itself (constancy of direction instead of continuity of direction).
- *constant width* is also illustrated in the same figure 2 and is very relevant for drawings and all kinds of natural and artificial forms.
- Notice, in the same spirit, that *convexity,* also illustrated, is a particularization of both closure and continuity of direction.

[2]The good continuation principle has been extensively adressed in Computer Vision, first in [23], more recently in [27] and still more recently in [11]. A recent example of computer vision paper implementing "good continuation", understood a "constant curvature", is [32].

- *past experience* : In the list of partial gestalts which are looked for in any image, we can have generic shapes like circles, ellipses, rectangles, and also silhouettes of familiar objects like faces, cats, chairs, etc.

Figure 2: Illustration of gestalt principles. *From left to right and top to bottom : color constancy + proximity, similarity of shape and similarity of texture ; good continuation ; closure (of a curve) ; convexity ; parallelism ; amodal completion (a disk seen behind the square); color constancy ; good continuation again (dots building a curve) ; closure (of a curve made of dots) ; modal completion : we tend to see a square in the last figure and its sides are seen in a modal way (subjective contour). Notice also the texture similarity of the first and last figures. Most of the figures involve constant width. In this complex figures, the sub-figures are identified by their alignment in two rows and their size similarity.*

All of the above grouping laws belong, according to Kanizsa, to the so called *processo primario* (primary process), opposed to a more cognitive secondary process. Also, it may of course be asked *why and how* this list of geometric qualities has emerged in the course of biological evolution. Brunswick and Kamiya [4] were among the first to suggest that the gestalt grouping laws were directly related to the geometric statistics of the natural world. Since then, several works have addressed from different points of views these statistics and the building elements which should be conceptually considered in perception theory, and/or numerically used in Computer Vision [3], [25], [10].

The grouping laws usually collaborate to the building up of larger and larger objects. A simple object like a square, whose boundary has been drawn in black with a pencil on a white sheet, will be perceived by connectedness (the boundary is a black line), by constant width (of the stroke), convexity and closure (of the black pencil stroke), parallelism (between opposite sides), orthogonality (between adjacent sides), again constant width (of both pairs of opposite sides).

We must therefore distinguish between *global* gestalt and the *partial* gestalts. The square is a global gestalt, but it is the synthesis of a long list of concurring local groupings, leading to parts of the square endowed with some gestalt quality. Such parts we shall call *partial gestalts.*

Notice also that all grouping gestalt laws are *recursive* : they can be applied first to atomic inputs and then in the same way to partial gestalts already constituted. Let us illustrate this by an example. In figure 3, the same partial gestalt laws, namely alignment, parallelism, constant width and proximity, are recursively applied not less than six times : the single elongated dots first aligned in rows, these rows in groups of two parallel rows, these groups again in groups of five parallel horizontal bars, these groups again in groups of six parallel vertical bars. The final groups appear to be again made of two macroscopic horizontal bars. The whole organization of such figures is seeable at once.

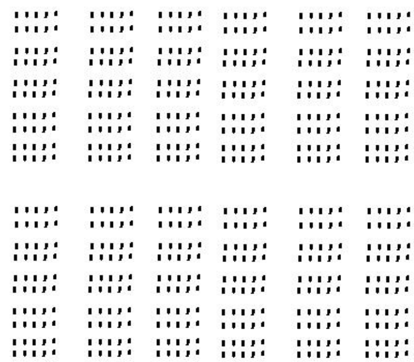

Figure 3: Recursiveness of gestalt laws : *here, constant width and parallelism are applied at different levels in the building up of the final group not less than six times, from the smallest bricks which are actually complex gestalts, being roughly rectangles, up to the final rectangle. Many objects can present deeper and more complex constructions.*

2.2 Global Gestalt principles

While the partial, recursive, grouping gestalt laws do not bring so much doubt about their definition as a computational task from atomic data, the global gestalt principles are by far more challenging. For many of them, we do not even know whether they are properly constitutive laws, or rather an elegant way of summarizing various perception processes. They constitute, however, the only cues we have about the way the partial gestalt laws could be derived from a more general principle. On the other hand, these principles are absolutely necessary in the description of the perception process, since they should fix the way grouping laws interact or compete to create the final global percepts, the final gestalts. Let us go on with the gestalt principles list which can be extracted from [14]. We have :

ragruppamento secondo la direzionalita della struttura (Kanizsa ,*Grammatica del Vedere*, op. cit., page 54) : inheritance by the parts of the overall group direction. This is a statement which might find its place in Platon's Parmenides: "the parts inherit the whole's qualities".

pregnancy, structural coherence, unity (pregnanza, coerenza strutturale, carattere unitario, ibidem, page 59), tendency to maximal regularity (ibidem, p. 60), articulation whole-parts, (in German, Gliederung), articulation without remainder (ibidem p. 65) : These seven Gestalt laws are not partial gestalts ; in order to deal with them from the computer vision viewpoint, one has to assume that all partial grouping laws have been applied and that a synthesis of the groups into the final global gestalts must be thereafter performed. Each principle describes some aspect of the synthesis made from partial grouping laws into the most wholesome, coherent, complete and well-articulated percept.

3 Conflicts of partial gestalts and the masking phenomenon

With the computational discussion to come in mind, we wish to examine the relationship between two important technical terms of Gestalt theory, namely *conflicts* and *masking*.

3.1 Conflicts

The gestalt laws are stated as independent grouping laws. Now, they start from the same building elements. Thus, conflicts between grouping laws can occur and therefore conflicts between different interpretations, that is, different possible groups in a given figure. Three cases :

a) two grouping laws act simultaneously on the same elements and give raise to two overlapping groups. It is not difficult to build figures where this occurs, as in figure 4. In that example, we can group the black dots and the white dots by similarity of color. All the same, we see a rectangular grid made of all the black dots and part of the white ones. We also see a good continuing curve, with a loop, made of white dots. These groups do not compete.

b) two grouping laws compete and one of them wins, the other one being inhibited. This case is called *masking* and will be discussed thoroughly in the next section.

c) conflict : in that case, both grouping laws are potentially active, but the groups cannot exist simultaneously. In addition, none of the grouping laws wins clearly. Thus, the figure is ambiguous and presents two or more possible interpretations.

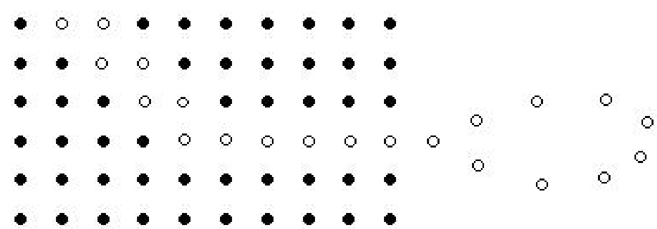

Figure 4: *Gestalt laws in simultaneous action without conflict : the white dots are elements of the grid (alignment, constant width) and simultaneously participate of a good continuing curve.*

A big section of Kanizsa's second chapter [14], is dedicated to conflicts of gestalts. Their study leads to the invention of clever figures where an equilibrium is maintained between two conflicting gestalt laws struggling to give the final figure organization. The viewer can direct his attention in both ways, see both organizations and perceive their conflict. A seminal experiment due to Wertheimer [3] gives an easy way to construct such conflicts. In figure 5, we see on the left a figure made of rectangles and ellipses. The prominent grouping laws are : a) shape similarity (L_1), which leads us to group the ellipses together and the rectangles as two conspicuous groups ; b) the vicinity law L_2, which makes all of these elements build anyway a unified cluster. Thus, on the left figure, both laws coexist without real conflict. On the right figure instead, two clusters are present. Each one is made of heterogeneous shapes, but they fall apart enough to enforce the splitting of the ellipses group and of the rectangles group. Thus, on the right, the vicinity law L_2 tends to win. Such figures can be varied, by changing (e.g.) progressively the distance between clusters until the final figure presents a good equilibrium between conflicting laws.

Some laws, like good continuation, are so strong that they almost systematically win, as is illustrated in figure 6. In this figure, two figures with a striking axial symmetry are concatenated in such a way that their boundaries are put in "good continuation". The result is a different

[3]Op. cit..

Figure 5: *Conflict of similarity of shapes with vicinity. We can easily view the left hand figure as two groups by shape similarity, one made of rectangles and the other one of ellipses. On the right, two different groups emerge by vicinity. Vicinity "wins" against similarity of shapes.*

interpretation, where the symmetric figures literally disappear. This is a conflict, but with a total winner. It therefore enters into the category of masking.

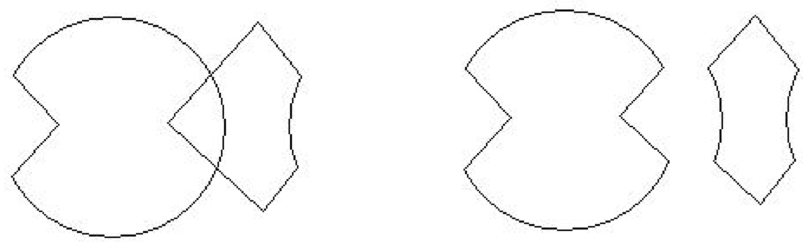

Figure 6: *A "conflict of gestalts" : two overlapping closed curves or, as suggested on the right, two symmetric curves which touch at two points ? We can interpret this experiment as a masking of the symmetry gestalt law by the good continuation law. (From Kanizsa, Grammatica del Vedere p 195, op. cit.)*

3.2 Masking

Masking is illustrated by a lot of puzzling figures, where partial gestalts are literally hidden by other partial gestalts giving a better global explanation of the final figure. The masking phenomenon is the outcome of a conflict between two partial gestalts L_1 and L_2 struggling to organize a figure. When one of them, L_1, wins, a striking phenomenon occurs : the other possible organization, which would result from L_2, is hidden. Only an explicit comment to the viewer can remind her of the existence of the possible organization under L_2 : the parts of the figure which might be perceived by L_2 have become invisible, *masked* in the final figure, which is perceived under L_1 only.

Kanizsa considers four kinds of masking : *masking by embedment in a texture ; masking by addition (the Gottschaldt technique) ; masking by substraction (the Street technique)* ; masking by manipulation of the figure-background articulation (Rubin, many famous examples by Escher). The first technique we shall consider is *masking in texture.* Its principle is : a geometrically organized figure is embedded into a texture, that is, a whole region made of similar building elements. This masking may well be called *embeddedness* as suggested by Kanizsa [4]. Figure

[4]Vedere e pensare, op. cit., p 184

7 gives a good instance of the power of this masking, which has been thoroughly studied by the schools of Beck and Juslesz [13]. In this clever figure, the basis of a triangle is literally hidden in set of parallel lines. We can interpret the texture masking as a conflict between an arbitrary organizing law L_2 and the similarity law, L_1. The masking technique works by multiple additions embedding a figure F organized under some law L_2 into many elements which have a shape similar to the building blocks of F.

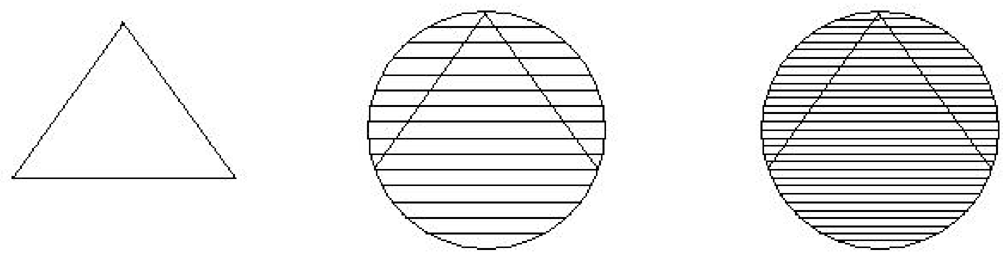

Figure 7: *Masking by embedding in a texture. The basis of the triangle becomes invisible as it is embedded in a group of parallel lines. (Galli and Zama, quoted in Vedere e pensare, op. cit.).*

The same proceeding is at work in figure 8. In that figure, one sees that a curve made of roughly aligned pencil strokes is embedded in a set of many more parallel strokes.

Figure 8: *Masking by embedding in a texture again. On the right, a curve created from strokes by "good continuation". This curve is present, but masked on the left. We can consider it as a conflict between L_2, "good continuation" and L_1 : similarity of direction. The similarity of direction is more powerful, as it organizes the full figure* (articulazione senza resti).

In the *masking by addition* technique, due to Gottschaldt, a figure is concealed by addition of new elements which create another and more powerful organization. Here, L_1 and L_2 can be any organizing law. In figure 9, an hexagon is thoroughly concealed by the addition to the figure of two parallelograms which include in their sides the initial sides of the hexagon. Noticeably, the "winning laws" are the same which made the hexagon so conspicuous before masking, namely closure, symmetry, convexity and good continuation.

As figure 10 shows, L_1 and L_2 can revert their roles. On the right, the curve obtained by good continuation is made of perfect half circles concatenated. This circular shape is masked in the good continuation. Surprisingly enough, the curve on the left is present in the figure on the right, but masked by the circles. Thus, on the left, good continuation wins against the past experience of circles. On the right, the converse occurs ; convexity, closure and circularity win against good continuation and mask it. The third masking technique considered by Kanizsa is substraction (Street technique), that is, removal of parts of the figure. As is apparent in figure 11, where a square is amputated in three different ways, the technique results effective only

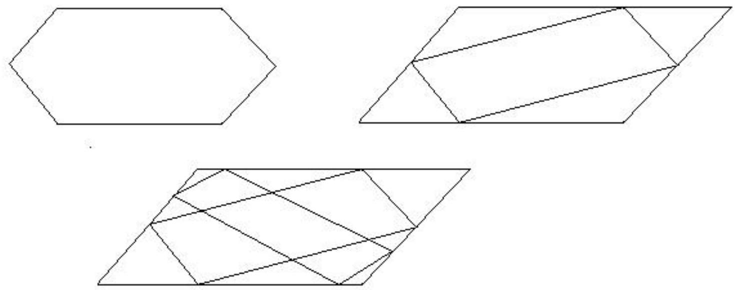

Figure 9: *Masking by concealment (Gottschaldt 1926). The hexagon on the left is concealed in the figure on the right, and more concealed below. The hexagon was built by the closure, symmetry, convexity gestalt laws. The same laws are active to form the winner figures, the parallelograms.*

Figure 10: *Masking of circles in good continuation, or, conversely, masking of good continuation by closure and convexity. We do not really see arcs of circles on the left, although significant and accurate parts of circles are present : we see a smooth curve. Conversely, we do not see the left "good" curve as a part of the right figure. It is nonetheless present in it.*

when the removal creates a new gestalt. The square remains in view in the third figure from the left, where the removal has been made at random and is assimilable to a random perturbation. In the second and fourth figure, instead, the square disappears, although some parts of its sides have been preserved.

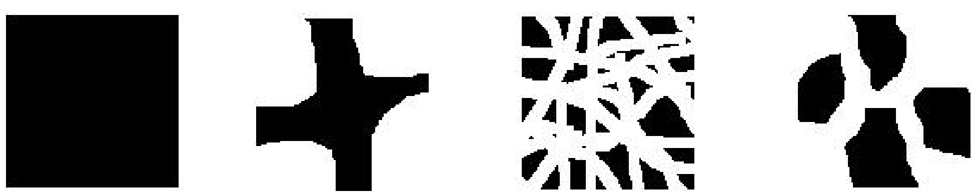

Figure 11: *Masking by the Street substraction technique (1931), inspired from Kanizsa (Vedere e pensare p 176, op. cit.). Parts are removed from the black square. When this is done in a coherent way, it lets appear a new shape (a rough cross in the second subfigure, four black spots in the last one) and the square is masked. It is not masked at all in the third, though, where the removal has been done in a random way and does not yields a competing interpretation.*

We should not end this section without considering briefly the last category of masking mentioned by Kanizsa, the masking by inversion of the figure-background relationship. This kind of masking is well known thanks to the famous Escher drawings. Its principle is "the background is not a shape" (*il fondo non é forma*). Whenever strong gestalts are present in an image, the space between those conspicuous shapes is not considered as a shape, even when it has itself a familiar shape like a bird, a fish, a human profile. Again here, we can interpret masking as the result of a conflict of two partial gestalt laws, one building the form and the other one, the loser, not allowed to build the background as a gestalt.

4 Quantitative aspects of Gestalt theory

In this section, we open the discussion on quantitative laws for computing partial gestalts. We shall first consider some numerical aspects of Kanizsa's *masking by texture*. In continuation, we shall make some comments on Kanizsa's paradox and its answer pointing out the involvement of a quantitative image resolution. These comments lead to Shannon's sampling theory.

4.1 Quantitative aspects of the masking phenomenon

In his fifth chapter of *Vedere e pensare*, Kanizsa points out that "it is licit to sustain that a black homogeneous region contains all theoretically possible plane figures, in the same way as, for Michelangelo, a marble block virtually contains all possible statues". Thus, these virtual statues could be considered as masked. This is the so called Kanizsa paradox. Figure 12 shows that one can obtain any simple enough shape by pruning a regular grid of black dots. In order

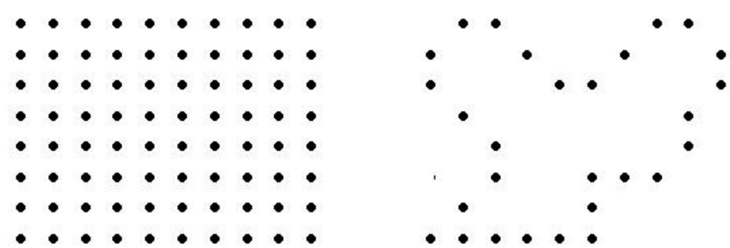

Figure 12: *According to Kanizsa's paradox, the figure on the right is potentially present in the figure on the right, and would indeed appear if we colored the corresponding dots. This illustrates the fact that the figure on the left contains a huge number of possible different shapes.*

to go further, it seems advisable to the mathematician to make a count : how many squares could we see, for example, in this figure ? Characterizing the square by its upper left corner and its side length, it is easily computed that the number of squares whose corners lie on the grid exceeds 400. The number of curves with "good continuation" made of about 20 points like the one drawn on the right of figure 12 is equally huge. We can estimate it in the following way : we have 80 choices for the first point, and about five points among the neighbors for the second point, etc. Thus, the number of possible good curves in our figure is grossly $80 \cdot 5^{20}$ if we accept the curve to turn strongly, and about $80 \cdot 3^{20}$ is we ask the curve to turn at a slow rate. In both cases, the number of possible "good" curves in the grid is very large.

This multiplicity argument suggests that a grouping law can be active in an image, only if its application would not create a huge number of partial gestalts. Or, to say it in another way, we can sustain that the multiplicity implies a masking by texture. Masking of all possible good curves in the grid of figure 12 occurs, just because too many such curves are possible.

In the above figure 8 (subsection 3.2), we can repeat the preceding quantitative argument. In this figure, the left hand set of strokes actually contains, as an invisible part, the array of strokes on the right. This array of strokes is obviously organized as a curve (good continuation gestalt). This curve becomes invisible on the left hand figure, just because it gets endowed in a more powerful gestalt, namely parallelism (similarity of direction). As we shall see in the computational discussion, the fact that the curve has been masked is related to another fact which is easy to check on the left hand part of the figure : one could select on the left many curves of the same kind as the one given on the right.

In short, we do not consider Kanizsa's paradox as a difficulty to solve, but rather as an arrow pointing towards the computational formulation of gestalt : In section 5, we shall define a partial gestalt as *a structure which is not masked in texture.*

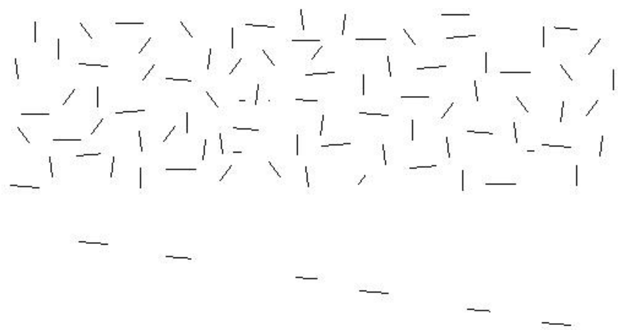

Figure 13: *Below, an array of roughly aligned segments. Above, the same figure is embedded into a texture in such a way that it still is visible as an alignment. We are in the limit situation associated with Vicario's proposition : "is masked only what can be unmasked".*

We shall therefore not rule out the extreme masking cases, in contradiction with Vicario's principle "é mascherato solo ciò che può essere smascherato" (is masked only what can be unmasked). Clearly, all psychophysical masking experiments must be close enough to the "conflict of gestalts" situation, where the masked gestalt still is attainable when the subject's attention is directed. Thus, psychological masking experiments must remain close to the non masking situation and therefore satisfy Vicario's principle. From the computational viewpoint instead, figures 12 and 8 are nothing but very good masking examples.

In this masking issue, one feels the necessity to pass from qualitative to quantitative arguments : a gestalt can be more or less masked. How to compute the right information to quantize this "more or less" ? It is actually related to a *precision parameter.* In figure 13, we constructed a texture by addition from the alignment drawn below. Clearly, some masking is at work and we would not notice immediately the alignment in the texture if our attention was not directed. All the same, the alignment remains somewhat conspicuous and a quick scan may convince us that *there is no other alignment of such an accuracy in the texture.* Thus, in that case, alignment is not masked by parallelism. Now, one can now suspect that this situation can be explained in yet quantitative terms : the precision of the alignment matters here and should be evaluated. Precision will be one of the three parameters we shall use in computational gestalt.

4.2 Shannon theory and the discrete nature of images

The preceding subsection introduced two of the parameters we shall have to deal with in the computations, namely the *number of possible partial gestalts* and a *precision* parameter. Before proceeding to computations, we must discuss the rough datum itself, namely the computational nature of images, let them be digital or biological. Kanizsa addresses briefly this problem in the fifth chapter of *Vedere e pensare*, in his discussion of the masking phenomenon : "We should not consider as masked elements which are too small to attain the visibility threshold". Kanizsa was aware that the amount of visible points in a figure is finite.[5] He explains in the same chapter why this leads to work with figures made of dots ; we can consider this decision as a way to quantize the geometric information.

In order to define mathematically an image, be it digital or biological, in the simplest possible way, we just need to fix a point of focus. Assume all photons converging towards this focus are intercepted by a surface which has been divided into regular cells, usually squares or hexagons. Each cell counts its number of photons hits during a fixed exposure time. This count gives a grey level image, that is, a rectangular, (roughly circular in biological vision) array of grey level values on a grid. In the case of digital images, C.C.D. matrices give regular grids made of squares. In the biological case, the retina is divided into hexagonal cells with growing sizes from the fovea. Thus, in all cases, a digital or biological image contains a finite number of values on a grid. Shannon [28] made explicit the mathematical conditions under which, from this matrix of values, a continuous image can be reconstructed. By Shannon's theory, we can compute the grey level at all points, and not only the points of the grid. Of course, when we zoom in the interpolated image it looks more and more blurry : the amount of information in a digital image is indeed finite and the resolution of the image is bounded. The points of the grid together with their grey level values are called *pixels*, an abbreviation for *pict*ure *el*ements.

The pixels are the computational atoms from which gestalt grouping procedures can start. Now, if the image is finite, and therefore blurry, **how can we infer sure events as lines, circles, squares and whatsoever gestalts from discrete data ?** If the image is blurry, all of these structures cannot be inferred as completely sure ; their exact location must remain uncertain. This is crucial : all basic geometric information in the image has a *precision*[6]. Figure 13 shows it plainly. It is easy to imagine that if the aligned segments, still visible in the figure, are slightly less aligned, then the alignment will tend to disappear. This is easy checked with figure 14, where we moved slightly up and down the aligned segments.

Let us now say briefly which local, atomic, information can be the starting point of computations. Since every local information about a function u at a point (x, y) boils down to its Taylor expansion, we can assume that these atomic informations are :

- the value $u(x, y)$ of the grey level at each point (x, y) of the image plane. Since the function u is blurry, this value is valid at points close to (x, y).

[5] *"non sono da considerare mascherati gli elementi troppo piccoli per raggiungere la soglia della visibilita pur potendo essere rivelati con l'ausilio di una lente di ingrandimento, il che dimostra che esistono come stimoli potenziali. E altrettanto vale per il caso inverso, nel quale soltanto con la diminuzione dell'angolo visivo e la conseguente riduzione della grandezza degli elementi e dei loro interspazi (mediante una lente o la visione a grande distanza) é possibile vedere determinate strutture."*

[6] It is well known by gestaltists that a right angle "looks right" with some ± 3 degrees precision, and otherwise does not look right at all.

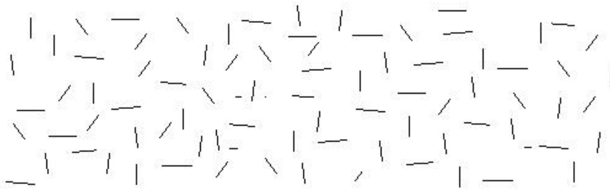

Figure 14: *When the alignment present in figure 13 is made less accurate, the masking by texture becomes more efficient. The precision plays a crucial role in the computational gestalt theory outlined in the next section.*

- the gradient of u at (x, y), the vector

$$Du(x,y) = \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right)(x,y) \ ;$$

- the *orientation* at (x, y),

$$Orient(x) = \frac{1}{\|Du(x,y)\|}\left(-\frac{\partial u}{\partial y}, \frac{\partial u}{\partial x}\right)(x,y).$$

  This vector is visually intuitive, since it is tangent to the boundaries one can see in an image.

These local informations are known at each point of the grid and can be computed at any point of the image by Shannon interpolation. They are quantized, having a finite number of digits, and therefore noisy. Thus, each one of the preceding measurements has an intrinsic precision. The orientation is invariant when the image contrast changes (which means robustness to illumination conditions). Attneave and Julesz [13] refer to it for shape recognition and texture discrimination theory. Grey level, gradient and orientation are the only local information we shall retain for the numerical experiments of the next section, together with their precisions.

5 Computing partial gestalts in digital images

In this section, we shall summarize a computational theory which permits to find automatically partial gestalts in digital images. This theory essentially predicts *perception thresholds* which can be computed on every image and give a usually clear cut decision between what is seeable as a geometric structure (gestalt) in the image and what is not. Those thresholds are computable thanks to the discrete nature of images. Many more details can be found in [9]. All computations below will involve three fundamental numbers, whose implicit presence in Gestalt theory has just been pointed out, namely

- a relative precision parameter p which we will treat as a probability ;
- a number N_{conf} of possible configurations for the looked for partial gestalt. This number is finite because the image resolution is bounded ;
- the number N of pixels of the image.

Of course, p and N_{conf} will depend upon the kind of gestalt grouping law under consideration. We can relate p and N to two fundamental qualities of any image : its noise and its blur.

5.1 A general detection principle

In [6], [7], [8], we outlined a computational method to decide whether a given partial gestalt (computed by any segmentation or grouping method) is reliable or not. As we shall recall, our method gives *absolute thresholds*, depending only on the image size, permitting to decide when a given gestalt is perceptually relevant or not.

We applied a general perception principle which we called Helmholtz principle (figure 15). This principle yields computational grouping thresholds associated with each gestalt quality. It can be stated in the following generic way. Assume that atomic objects O_1, O_2,...,O_n are present in an image. Assume that k of them, say O_1,...,O_k, have a common feature, say, same color, same orientation, position etc.. We are then facing the dilemma : is this common feature happening by chance or is it significant and enough to group O_1,..., O_k ? In order to answer this question, we make the following mental experiment : we assume *a priori* that the considered quality has been randomly and uniformly distributed on all objects O_1, ...,O_n. Then we (mentally) assume that the observed position of objects in the image is a random realization of this uniform process. We finally ask the question : is the observed repartition probable or not ? If not, this proves *a contrario* that a grouping process (a gestalt) is at stake. Helmholtz principle states roughly that in such mental experiments, the numerical qualities of the objects are assumed to be equally distributed and independent. Mathematically, this can be formalized by

Definition 1 (ε-meaningful event [6]) *We say that an event of type "such configuration of geometric objects has such property" is ε-meaningful if the expectation of the number of occurrences of this event is less than ε under the uniform random assumption.*

As an example of generic computation we can do with this definition, let us assume that the probability that a given object O_i has the considered quality is equal to p. Then, under the independence assumption, the probability that at least k objects out of the observed n have this quality is

$$B(p,n,k) = \sum_{i=k}^{n} \binom{n}{k} p^i (1-p)^{n-i},$$

i.e. the tail of the binomial distribution. In order to get an upper bound for the number of false alarms, *i.e.* the expectation of the number of geometric events happening by pure chance, we can simply multiply the above probability by the number of tests we perform on the image. This number of tests N_{conf} corresponds to the number of different possible positions we could have for the searched gestalt. Then, in most cases we shall consider in the next subsections, a considered event will be defined as ε-meaningful if

$$NFA = N_{conf} \cdot B(p,n,k) \leqslant \varepsilon.$$

We call in the following the left hand member of this inequality the "number of false alarms" (NFA). The number of false alarms of an event measures the "meaningfulness" of this event : the smaller it is, the more meaningful the event is. We refer to [6] for a complete discussion of this definition. To the best of our knowledge, the use of the binomial tail, for alignment detection, was introduced by Stewart [30].

5.2 Computation of six partial gestalts

Alignments of points.

Points will be called aligned if they all fall into a strip thin enough, and in sufficient number. This qualitative definition is easily made quantitative. The precision of the alignment is measured by the width of the strip. Let S be a strip of width a. Let $p(S)$ denote the prior probability for a point to fall in S, and let $k(S)$ denote the number of points (among the M) which are in S. The following definition permits to compute all strips where a meaningful alignment is observed (see figures 15 and 18).

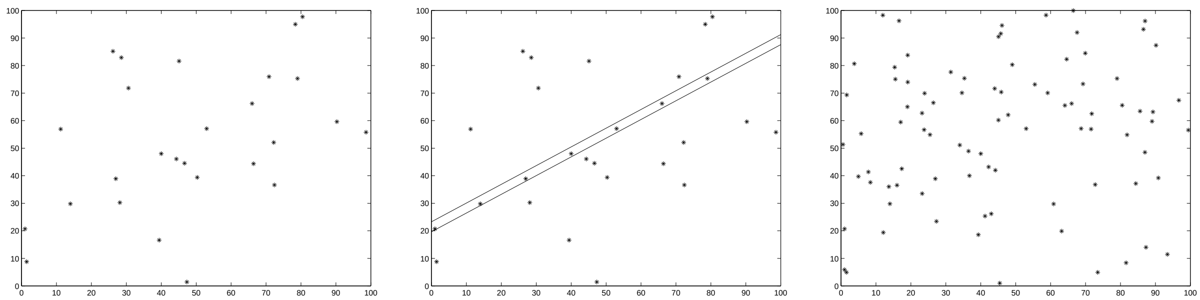

Figure 15: An illustration of Helmholtz principle : non casual alignments are automatically detected by Helmholtz principle as a large deviation from randomness. *Left, 20 uniformly randomly distributed dots, and 7 aligned added. Middle : this meaningful (and seeable) alignment is detected as a large deviation from randomness. Right : same alignment added to 80 random dots. The alignment is no more meaningful (and no more seeable). In order to be meaningful, it would need to contain at least 11 points.*

Definition 2 ([9]) *A strip S is ε-meaningful if*

$$NF(S) = N_s \cdot B(p(S), M, k(S)) \leqslant \varepsilon,$$

where N_s is the number of considered strips (one has $N_s \simeq 2\pi(R/a)^2$, where R is the half-diameter of Ω and a the minimal width of a strip).

Let us give immediately a summary of several algorithms based on the same principles, most of which use, as above, the tail of the binomial law. This is done in table 1, where we summarize the (very similar) formulae permitting to compute the following partial gestalts : *alignments (of orientations in a digital image), contrasted boundaries, all kinds of similarities for some quality measured by a real number (grey level, orientation,. . .)*, and of course the most basic one, treated in the last row, namely the *vicinity* gestalt.

Alignments in a digital image

The first row of table 1 treats the alignments in a digital image. As we explained in subsection 4.2, an orientation can be computed at each point of the image. Whenever a long enough segment occurs in the image, whose orientations are aligned with the segment, this segment is perceived as an alignement. We consider the following event : "on a discrete segment of the image, joining two pixel centers, and with length l, at least k independent points have the same direction as the segment with precision p." The definition of the number of false alarms is given in the first row of the table and an example of the alignments, in a digital aerial image, whose number of false alarms is less than 1 is given in figure 16.

GROUP LOOKED FOR	MEASUREMENTS	NUMBER OF FALSE ALARMS
Alignment of directions on a segment [6]		
a discrete segment with points at Nyquist distance (i.e. 2)	k: number of aligned points l: number of points on the segment	$N_{segments} \cdot B(p,l,k)$ $N_{segments} = N^3$ N: number of pixels in the image $p = 1/16$ (angular precision)
Contrasted edges and boundaries [7]		
a level line (or a piece of) with points at Nyquist distance (i.e. 2)	μ : minimum contrast (gradient norm) along the curve l: length of the curve	$N_{level\ lines} \cdot H(\mu)^l$ H is the empirical cumulative distribution of the gradient norm on the image
Similarity of a uniform scalar quality (grey level, orientation, etc.) [9]		
a group of objects having a scalar quality q such that $a \leqslant q \leqslant b$	k: number of points in the group M: total number of objects	$\frac{L(L+1)}{2} \cdot B\left(\frac{b-a+1}{L}, M, k\right)$ L: number of values ($q \in \{1..L\}$)
Similarity of a scalar quality with decreasing distribution (area, length, etc.) [9]		
a group of objects having a scalar quality q such that $a \leqslant q \leqslant b$	k: number of points in the group M: total number of objects	$\frac{L(L+1)}{2} \cdot \max_{p\in\mathcal{D}} B\left(\sum_{i=a}^{b} p(i), M, k\right)$ L: number of values ($q \in \{1..L\}$) $\mathcal{D}$: set of decreasing distributions on $\{1..L\}$
Alignment of points (or objects) [9]		
a goup of points falling in a strip (region enclosed by two parallel lines)	p: relative area of the strip k: number of points falling in the strip	$N_{strips} \cdot B(p,M,k)$ M: total number of points The strips are quantized in position, width and orientation
Vicinity : clusters of points (or objects) [9]		
a group of points falling in a region enclosed by a low-resolution curve	σ: relative area of the region σ': relative area of the thick low-resolution curve k: number of points falling in the region	$N_{regions} \cdot \sum_{i=k}^{M} \binom{M}{i} \sigma^i (1-\sigma-\sigma')^{M-i}$ M: total number of points $N_{regions} = N^2 qr2^L$: the low resolution curves are quantized in resolution (q), thickness (r), location (N), and bounded in length (L).

Table 1: List of gestalts computed so far

Maximal meaningful gestalts and *articulazione senza resti*
On this example of alignments, we can address a problem encountered by the mathematical formalization of gestalt. Assume that on a straight line we have found a very meaningful segment S. Then, by enlarging slightly or reducing slightly S, we still find a meaningful segment. This means that meaningfulness cannot be a univoque criterion for detection, unless we can point out the "best meaningful" explanation of what is observed as meaningful. This is done by the following definition, which can be adapted as well to meaningful boundaries [7], meaningful edges [7], meaningful modes in a histogram and clusters [9].

Definition 3 ([8]) *We say that an ε-meaningful geometric structure A is maximal meaningful if*

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

It is proved in [8] that maximal structures cannot overlap, which is one of the main theoretical outcomes validating the above definitions. This definition formalizes the *articulazione senza resti* principle in the case of a single grouping law.

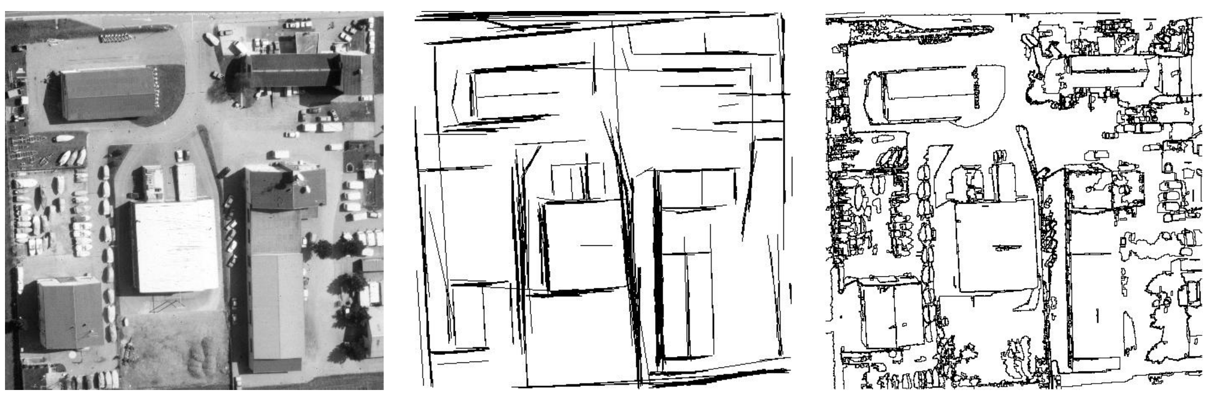

Figure 16: Two partial gestalts, alignments and boundaries. *Left : original aerial view (source : INRIA), middle : maximal meaningful alignments, right : maximal meaningful boundaries.*

Boundaries
One can define in a very similar way the "boundary" grouping law. This grouping law is never stated explicitly in gestaltism, because it is probably too obvious for phenomenologists. From the computation viewpoint it is not, at all.

The definition of the number of false alarms for boundaries involves again two variables : the length l of the level line, and its minimal contrast μ, which is interpreted as a precision. An example of boundary detection is given on figure 16.

Similarity
The third row of table 1 addresses the main gestaltic grouping principle : points or objects having a feature in common are being grouped, just because they have this feature in common. Assume k objects $O_1, \ldots, O_k$, among a longer list $O_1, \ldots, O_M$, have some quality q in common. Assume that this quality is actually measured as a real number. Then our decision of whether the grouping of $O_1, \ldots, O_k$ is relevant must be based on the fact that the values $q(O_1), \ldots, q(O_k)$

make a *meaningfully dense interval* of the histogram of $q(O_1), \ldots, q(O_M)$. Thus, the automatic quality grouping is led back to the question of an automatic, parameterless, histogram mode detector. Of course, this mode detector depends upon the kind of feature under consideration. We shall consider two paradigmatic cases, namely the case of orientations, where the histogram can be assumed by Helmholtz principle to be flat, and the case of the objects sizes (areas) where the null assumption is that the size histogram is decreasing (see figure 17).

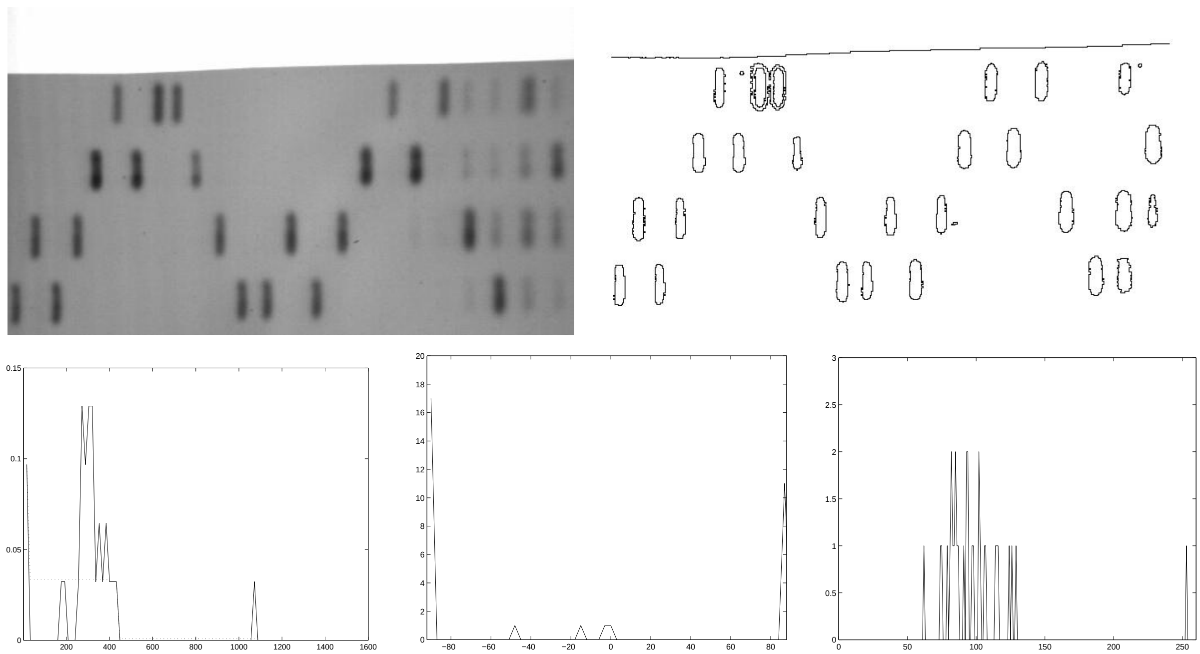

Figure 17: Collaboration of gestalts. The objects tend to be grouped similarly by several different partial gestalts. *First row : original DNA image (left) and its maximal meaningful boundaries (right). Second row, left : histogram of areas of the meaningful blobs. There is a unique maximal mode (256-416). The outliers are the double blob, the white background region and the three tiny blobs. Second row, middle : histogram of orientations of the meaningful blobs (computed as the principal axis of each blob). There is a single maximal meaningful mode (interval). This mode is the interval 85-95. It contains 28 objects out of 32. The outliers are the white background region and three tiny spots. Second row, right : histogram of the mean grey levels inside each block. There is a single maximal mode containing 30 objects out of 32, in the grey level interval 74-130. The outliers are the background white region and the darkest spot.*

Thus, the third and fourth row of our table permit to detect all kinds of similarity gestalt : objects grouped by orientation, or grey level, or any perceptually relevant scalar quality.

5.3 Study of an example

We are going to perform a complete study of a digital image, figure 17, involving all computational gestalts defined in table 1. The analyzed image is a common digital image. It is a scan of photograph and has blur and noise. The seeable objects are electrophoresis spots which have all similar but varying shape and color and present some striking alignments. Actually, all of these perceptual remarks can be recovered in a fully automatic way by combining several partial gestalt grouping laws.

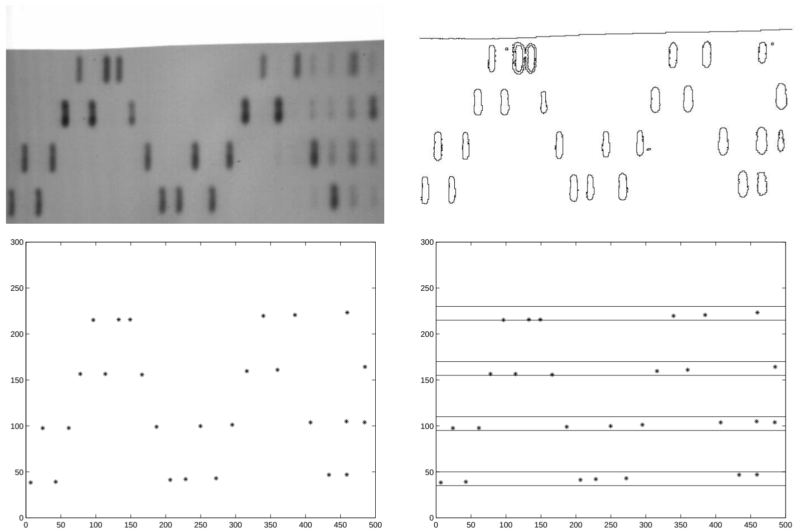

Figure 18: Gestalt grouping principles at work for building an "order 3" gestalt (alignment of blobs of the same size). *First row : original DNA image (left) and its maximal meaningful boundaries (right). Second row : left, barycenters of all meaningful regions whose area is inside the only maximal meaningful mode of the histogram of areas ; right, meaningful alignments of these points.*

First, the *contrasted boundaries* of this electrophoresis image are computed (above, right). Notice that all closed curves found are indeed perceptually relevant as they surround the conspicuous spots. Other many possible boundaries in the noisy background have been ruled out and remain "masked in texture". Let us apply a second layer of grouping laws. This second layer will use as atomic objects the blobs found at the first step. For each of the detected boundaries, we compute three qualities, namely

a) *the area* enclosed by the boundary, whose histogram is displayed on the top left of figure 17. There is a unique maximal mode in this figure, which actually groups all and exactly the blobs with similar areas and rules out two tiny blobs and a larger one enclosing two different blobs. Thus, almost all blobs get grouped by this quality, with the exception of two tiny spots and a double spot.

b) *the orientation* of each blob, an angle between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$. This histogram (figure 17, bottom, middle) again shows a single maximal mode, again computed by the formula of the third row of table 1. This mode appears at both end points of the interval, since the dominant direction is $\pm\frac{\pi}{2}$ and these values are identified modulo π. Thus, about the same blobs as in b) get grouped by their common orientation.

c) *the average grey level* inside each blob : its histogram is shown on the bottom right of figure 17. Again, most blobs, but not all get grouped with respect to this quality.

A further structural grouping law can be applied to build subgroups of blobs formed by alignment. This is illustrated in figure 18 (bottom, left), where the meaningful alignments are found.

This experiment illustrates the usual strong collaboration of partial gestalts : most salient objects or groups come to sight by several grouping laws.

6 The limits of every partial gestalt detector

The preceding section argued in favor of a very simple principle, Helmholtz principle, applicable to the automatic and parameterless detection of any partial gestalt, in full agreement with our perception. In this section, we shall show by commenting briefly several experiments that *tout n'est pas rose* : there is a good deal of visual illusion in any apparently satisfactory result provided by a partial gestalt detector on a digital image. We explained in the former section that partial gestalts often collaborate. Thus, in many cases, what has been detected by a partial gestalt will be corroborated by another one. For instance, boundaries and alignments in the experiment 16 are in good agreement. But what can be said about the experiment of figure 19 ? In this cheetah image, we have applied the alignment detector explained above. It works wonderfully on the grass leaves when they are straight. Now, we also see some unexpected alignments in the fur. These alignments do exist : these detected lines are tangent to several of the convex dark spots on the fur. This generates a meaningful excess of aligned points on these lines, the convex sets being smooth enough and having therefore on their boundary a long enough segment tangent to the detected line.

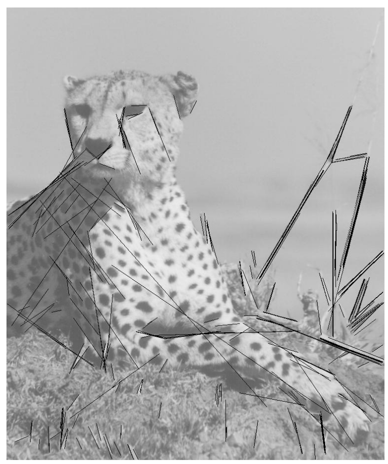

Figure 19: *Smooth convex sets or alignments ?*

We give an illustration of this fact in figure 20. The presence of only two circles can actually create a (to be masked) alignment, since the circles have bitangent straight lines. In order to discard such spurious alignments, the convexity (or good continuation) gestalt should be systematically searched when we look for alignments. Then, the alignments which only are tangent lines to several smooth curves, could be inhibited.

We should detect as alignment what indeed is aligned, but only under the condition that the alignment does not derive from the presence of several smooth curves... This statement can be generalized : no gestalt is just a positive quality. The outcome of a partial gestalt detector is valid only when all other partial gestalts have been tested and the eventual conflicts dealt with.

The same argument applies to our next experimental example, in figure 21. In that case, a dense cluster of points in present. Thus, it creates a meaningful amount of dots in many strips

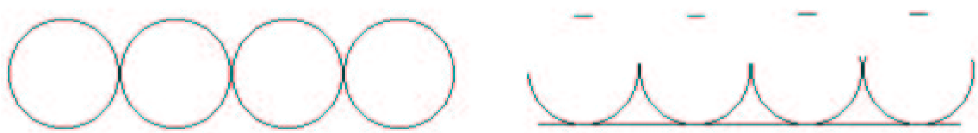

Figure 20: *Alignment is masked by good continuation and convexity : the small segments on the right are perfectly aligned. Any alignment detector should find them. All the same, this alignment disappears on the left figure, as we include the segments into circles. In the same way, the casual alignments in the Cheetah fur (figure 19) are caused by the presence of many oval shapes. Such alignments are perceptually masked and should be computationally masked !*

and the result is the detection of obviously wrong alignments. Again, the detection of a cluster should inhibit such alignment detections. We defined an alignment as "many points in a thin strip", but must add to this definition : "provided these points do not build one or two dense clusters".

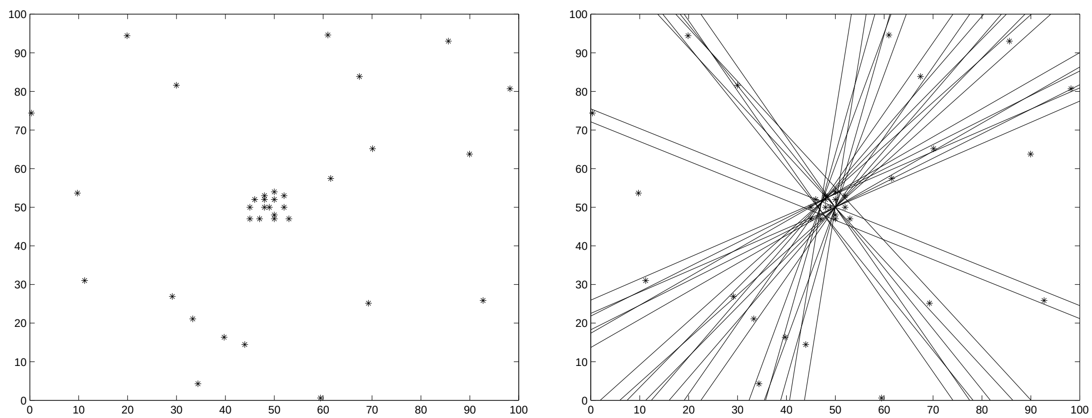

Figure 21: *One cluster, or several alignments ?*

One can reiterate the same problematic with another gestalt conflict (figure 22). In this figure, a detector of arcs of circles has been applied. The arc of circle detection grouping law is easily adapted from the definition of alignments in table 1. The main outcome of the experiment is this : since the MegaWave figure contains many smooth boundaries and several straight lines, lots of meaningful circular arcs are found. It may be discussed whether those circular arcs are present or not in the figure : clearly, any smooth curve is locally tangent to some circle. In the same way, two segments with an obtuse angle are tangent to several circular arcs (see figure 23). Thus, here again, a partial gestalt should mask another one. Hence the following statement, which is of wide consequence in Computer Vision : *We cannot hope any reliable explanation of any figure by summing up the results of one or several partial gestalts. Only a global synthesis, treating all conflicts of partial gestalts, can give the correct result.*

In view of these experimental counterexamples, it may well be asked why partial gestalt detectors often work "so well". This is due to the redundancy of gestalt qualities in most natural images, as we explained in the first section with the example of a square. Indeed, most natural or synthetic objects are simultaneously conspicuous, smooth and have straight or convex parts,

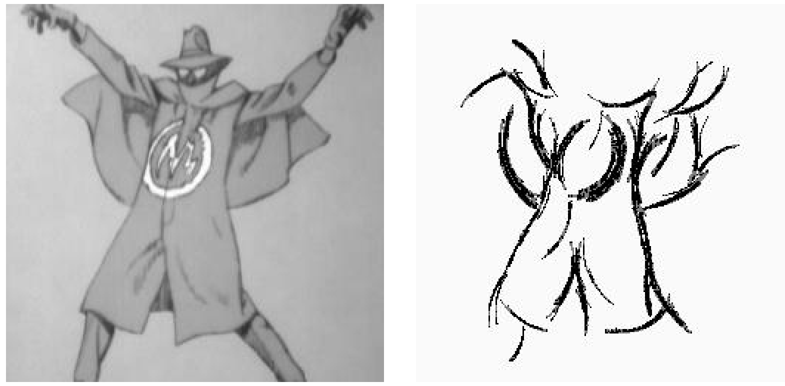

Figure 22: *Left : original "MegaWave" image. Right : an circular arc detector is applied to the image. Now, this image contains many smooth curves and obtuse angles to which meaningful circular arcs are tangent. This illustrates the necessity of the interaction of partial gestalts : the best explanation for the observed structures is "good continuation" in the gestaltic sense, i.e. the presence of a smooth curve, or of straight lines (alignments) forming obtuse angles. Their presence entails the detection of arcs of circles which are not the final explanation.*

etc. Thus, in many cases, each partial gestalt detector will lead to the same group definition. Our experiments on the electrophoresis image (figure 17) have illustrated the *collaboration of gestalt* phenomenon[7]. In that experiment, partial gestalts collaborate and seem to be redundant. This is an illusion which can be broken when partial gestalts do not collaborate.

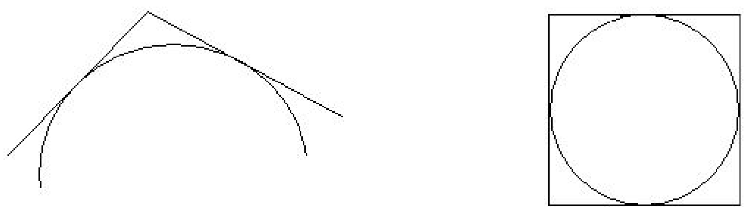

Figure 23: *Left : Every obtuse angle can be made to have many points in common with some long arc of circle. Thus, an arc of circle detector will make wrong detections when obtuse angles are present (see Figure 22). In the same way, a circle detector will detect the circle inscribed in any square and, conversely, a square detector will detect squares circumscribing any circle.*

6.1 Conclusion

We shall close the discussion by expressing some hope, and giving some arguments in favor of this hope. First of all, gestaltists pointed out the relatively small number of relevant gestalt qualities for biological vision. We have briefly shown in this paper that many of them (and probably all) can be computed by the Helmholtz principle followed by a maximality argument. Second, the discussions of gestaltists about "conflicts of gestalts", so vividly explained in the books of Kanizsa, might well be solved by a few information theoretical principles. As a good

[7] Krüger and Wörgötter [19] gave strong statistical computational evidence in favor of a collaboration between partial gestalt laws, namely collinearity, parallelism, color, contrast and similar motion.

example of it, let us mention how the dilemma alignment-versus-parallelism can be solved by an easy minimal description length principle (MDL) [26], [8]. Figure 24 shows the problem and its simple solution. On the middle, we see all detected alignments in the Brown image on the left. Clearly, those alignments make sense but many of them are slanted. The main reason is this : all straight edges are in fact blurry and therefore constitute a rectangular region where all points have roughly the same direction. Thus, since alignment detection is made up to some precision, the straight alignments are mixed up with slanted alignments which still respect the precision bound. We can interpret the situation as a conflict between alignment and parallelism, as already illustrated in figure 8.

The spurious, slanted alignments are easily removed by the application of a MDL principle : it is enough to retain for each point only the most meaningful alignment to which it belongs. We then compute again the remaining maximal meaningful alignments and the result (right) shows that the conflict between parallelism and alignment has been solved. Clearly, information theoretical rules of this kind may be applied in a general framework and put order in the proliferation of "partial gestalts". Let us mention an attempt of this kind in [20], where the author proposed a MDL reformulation of segmentation variational methods ([24]).

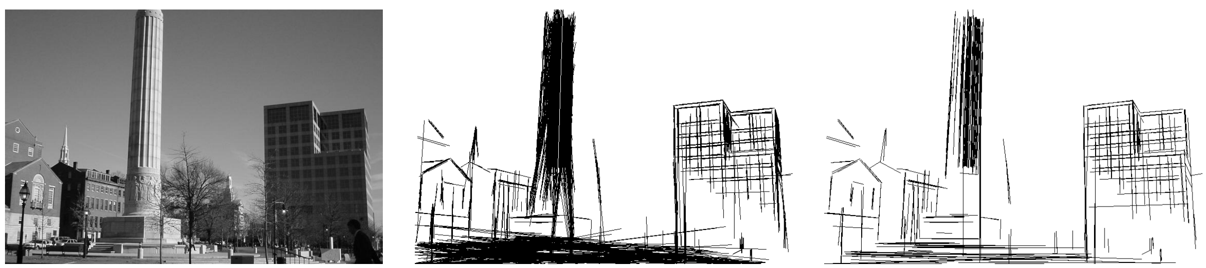

Figure 24: *Parallelism against alignment. Left, original Brown image. Middle : maximal meaningful alignments. Here, since many parallel alignments are present, secondary, parasite slanted alignments are also found. Right : Minimal description length of alignments, which eliminates the spurious alignments. This last method outlines a solution to conflicts between partial gestalts.*

Acknowledgments Work partially supported by Office of Naval Research under grant N00014-97-1-0839, Centre National d'Etudes Spatiales, Centre National de la Recherche Scientifique et Ministère de la Recherche et de la Technologie. We thank the Fondation des Treilles for having hosted the authors during the redaction of this paper.

References

[1] J.-P. d'Alès, J. Froment and J.-M. Morel, "Reconstruction visuelle et généricité", *Intellectica* 1:28, p. 11-35, 1999.

[2] F. Attneave, 'Some informational aspects of visual perception", *Psych. Rev.* 61, pp. 183–193, 1954.

[3] A.J. Bell A.J. and T.J. Sejnowski, "Edges are the 'independent components' of natural scenes", *Advances in Neural Information Processing Systems*, 9, 1996.

[4] E. Brunswik and J. Kamiya, "Ecological cue-validity of "proximity" and other Gestalt factors", *American Journal of Psychology* 66, pp. 20-32, 1953.

[5] V. Caselles, B. Coll et J.-M. Morel, "A Kanizsa programme", *Progress in Nonlinear Differential Equations and their Applications*, 25, pp. 35-55, 1996.

[6] A. Desolneux, L. Moisan and J.-M. Morel, "Meaningful Alignments", *International Journal of Computer Vision*, vol 40:1, pp.7-23, 2000.

[7] A. Desolneux, L. Moisan and J.-M. Morel, "Edge Detection by Helmholtz Principle", *Journal of Mathematical Imaging and Vision*, vol 14:3, pp 271-284, 2001.

[8] A. Desolneux, L. Moisan and J.-M. Morel, "Maximal Meaningful Events and Applications to Image Analysis", preprint CMLA no 2000-22, available on `http://www.cmla.ens-cachan.fr/Cmla/Publications/`, submitted.

[9] A. Desolneux, L. Moisan and J.-M. Morel, "Partial Gestalts", preprint CMLA no 2001-22, submitted to *IEEE Trans. on Pattern Analysis and Machine Intelligence*, available on `http://www.cmla.ens-cachan.fr/Cmla/Publications/`.

[10] W.S. Geisler, J.S. Perry, B.J. Super and D.P. Gallogly, "Edge co-occurrence in natural images predicts contour grouping performance", *Vision Research* 41, pp. 711-724, 2001.

[11] G. Guy and G. Medioni, "Inferring global perceptual contours from local features", *Int. Journal of Comp. Vision* 20:1, pp. 113-133, 1996.

[12] B.K. Horn, *Robot Vision*, MIT Press, 1987.

[13] J.R. Bergen and B. Julesz. "Textons, the fundamental elements of preattentive vision and perception of textures", *Bell System Techical Journal* 62:6, pp. 1619-1645, 1983.

[14] G. Kanizsa, *Grammatica del Vedere*, Il Mulino, Bologna, 1980.

[15] G. Kanisza, *Vedere e pensare*, Il Mulino, Bologna, 1991.

[16] G. Kanizsa, *Organization in Vision*, Holt, Rinehart & Winston, 1979.

[17] M. Kass, A. Witkin and D. Terzopoulos, "Snakes: active contour models", *1st Int. Comp. Vision Conf.* IEEE 777, 1987.

[18] N. Kiryati, Y. Eldar and A.M. Bruckstein, "A Probabilistic Hough Transform", *Pattern Recognition* 24:4, pp 303-316, 1991.

[19] N. Krüger and F. Wörgötter, "Multi-modal estimation of collinearity and parallelism in natural image sequences", submitted to *Network Computation in Neural Systems*, 2001.

[20] Leclerc, Y. Constructing Simple Stable Descriptions for Image Partitioning. *Int. J. of Comp. Vision* , 3, pp. 73-102, 1989.

[21] D. Marr, *Vision*, Freeman and co., 1982.

[22] W. Metzger, *Gesetze des Sehens.* Waldemar Kramer, 1975.

[23] U. Montanari, "On the Optimal Detection of Curves in Noisy Pictures", CACM 14:5, pp. 335-345, 1971.

[24] Mumford, D. and J. Shah. Boundary detection by minimizing functionals *IEEE Conf. on Comp. Vision and Pattern Recognition*, San Francisco, 1985.

[25] B.A. Olshausen and D.J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images", *Nature* 381:6583, pp. 607-609, 1996.

[26] Rissanen, J. A universal prior for integers and estimation by Minimum Description Length. *Annals of Statistics* 11 (2), 1983.

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

[28] C. E. Shannon, "A mathematical theory of communication", *Bell System Technical Journal*, 27, pp. 379-423 and 623-656, 1948.

[29] M. Sigman M, G.A. Cecchi, C.D. Gilbert, and M.O. Magnasco. "On a common circle: Natural scenes and Gestalt rules", *Proc. Nat. Acad. Sci. USA* 98:4, pp. 1935-1940, 2001.

[30] C. V. Stewart, "MINPRAN : a new robust estimator for Computer Vision", *I.E.E.E. Trans. on Pattern Analysis and Machine Intelligence* 17, pp. 925-938, 1995.

[31] M. Wertheimer, "Untersuchungen zur Lehre der Gestalt, II". *Psychologische Forschung* 4, pp. 301-350, 1923. Translation published as Laws of Organization in Perceptual Forms, in Ellis, W. (1938). A source book of Gestalt psychology (pp. 71-88). London: Routledge & Kegan Paul.

[32] D.M. Wuescher and K.L. Boyer, "Robust contour decomposition using constant curvature criterion", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13:1, pp. 41-51, 1991.