



The collaboration of grouping laws in vision

Rafael Grompone von Gioi, Julie Delon, Jean-Michel Morel

► To cite this version:

Rafael Grompone von Gioi, Julie Delon, Jean-Michel Morel. The collaboration of grouping laws in vision. *Journal of Physiology - Paris*, 2012, 106 (5-6), pp.266-283. 10.1016/j.jphysparis.2011.12.002 . hal-00644851

HAL Id: hal-00644851

<https://hal.science/hal-00644851>

Submitted on 27 Nov 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The collaboration of grouping laws in vision

Rafael Grompone von Gioi, Julie Delon & Jean-Michel Morel

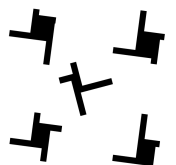
November 9, 2011

Abstract

Gestalt theory gives a list of geometric grouping laws that could in principle give a complete account of human image perception. Based on an extensive thesaurus of clever graphical images, this theory discusses how grouping laws collaborate, and conflict toward a global image understanding. Unfortunately, as shown in the bibliographical analysis herewith, the attempts to formalize the grouping laws in computer vision and psychophysics have at best succeeded to compute individual partial structures (or *partial gestalts*), such as alignments or symmetries. Nevertheless, we show here that a never formalized clever Gestalt experimental procedure, the *Nachzeichnung* suggests a numerical set up to implement and test the collaboration of partial gestalts. The new computational procedure proposed here analyzes a digital image, and performs a numerical simulation that we call *Nachtanz* or Gestaltic dance. In this dance, the analyzed digital image is gradually deformed in a random way, but maintaining the detected partial gestalts. The resulting dancing images should be perceptually indistinguishable if and only if the grouping process was complete. Like the *Nachzeichnung*, the *Nachtanz* permits a visual exploration of the degrees of freedom still available to a figure after all partial groups (or gestalts) have been detected. In the new proposed procedure, instead of drawing themselves, subjects will be shown samples of the automatic Gestalt dances and required to evaluate if the figures are similar. Several numerical preliminary results with this new Gestaltic experimental setup are thoroughly discussed.

1 Introduction

Vision and perceptual systems manage to pass from local and partial measurements to a global understanding of a scene. The laws governing this process are widely unknown, in spite of valuable efforts by philosophers, psychologists, neurologists, and computer scientists. The Gestalt school, with among others Wertheimer, Köhler, Koffka, Metzger, Kanizsa, Metelli [57, 28, 9, 35, 36, 25, 26] has concentrated in the past century many of the efforts to address this problem. The school proceeded by inventing and displaying clever geometric figures to subjects. Even if the Gestalt theories were mainly qualitative, the experimental imagination displayed and the mass of results is impressively demonstrated in the books by Kanizsa and Metzger. Their main conclusion was that the first steps in visual perception, observed on humans, work like geometric grouping machines. The phenomenological experiments show that the grouping laws are purely geometric and mostly independent of any empirical knowledge. A reduced set of grouping laws was proposed by the Gestalt school. According to Kanizsa [25], these laws include proximity, similarity, continuity of direction, good continuation, tendency to convexity, closure, common region, connectedness, constant width, symmetry, common motion, and *prägnanz* (or law of the good Gestalt). These laws were shown to work



mainly in collaborative form. However much attention was paid in the experiments to their conflicts, when two groups of laws compete to form alternative groups. Yet, the Gestalt collaboration and conflicts were only described qualitatively.

A major contribution of the Gestalt school is the creation of the largest and most comprehensive set of simple figures and images that are, at the same time, meaningful visual experiments (see figure 1). These experiments as a whole are supposed to cover all the most important mechanisms of vision. Like the pendulum or the cannon ball in physics, each one of these simple visual experiments involves, exemplifies, or disproves a hypothetical fundamental law of perception. Each of these figures, therefore, is a simple test that every theory of vision should match and explain. In particular, these figures represent a permanent challenge to computer vision, which could consider one of its main goals to process them automatically.

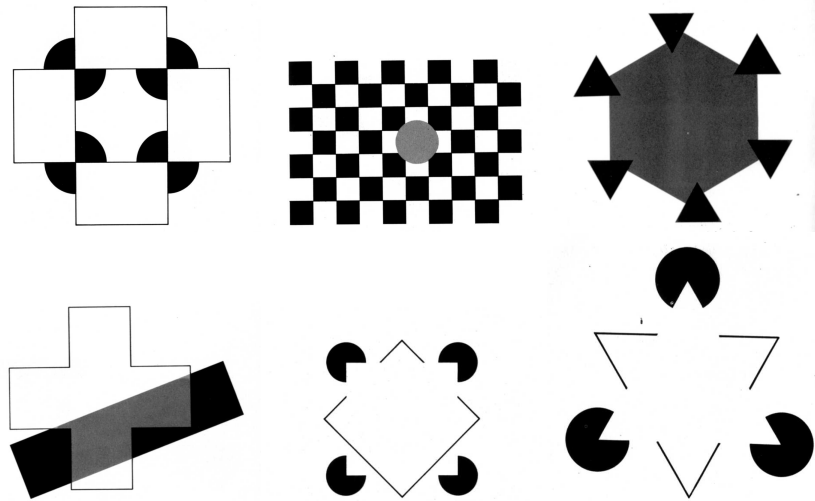


Figure 1: Some visual experiments of the Gestalt school.

These figures have many different complexity levels, involving from very few grouping rules to almost all. Thus, they make the perfect test bed for computational theories of vision. Their goal was precisely to translate the problem of an explanation of vision to a reduced but complete set of test figures avoiding the luxuriance of natural scenes. Due to the simplicity of Gestalt experiments, essentially composed of geometrical figures on a flat scene, we already possess the mathematical knowledge to handle most of the concepts and relations involved in them. On the other hand, these figures which were printed in the Gestalt age can now be digitalized and fully analyzed as digital images in a computer. This technological possibility is the main new feature that has emerged in the past thirty years.

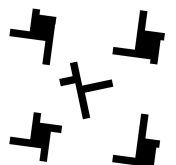
In the last 40 years, many studies have been proposed by psychologists to complete the work of the Gestalt school. These studies generally rely on psychophysics experiments to establish quantitative limits to the perception of each law. Much attention has been drawn on the good continuation law [53, 52, 12, 3, 41, 13, 60] and the different situations where it occurs. A possible neural mechanism for this law is suggested in [12]. In the same vein, several papers focus on the ability of human perception to detect symmetry in the presence of perturbations [2, 54]. In [24], psychophysical experiments

tend to prove that this ability is the result of the collaboration of several processes, one for instance detecting orientation uniformity. For a complete review on the subject of symmetry, see [55]. Numerous papers have been written on other Gestalt laws. Among them, let us mention some works on the law of proximity [29, 53], and the interesting work of Rubin on the fundamental role of junctions in our perception of occlusion [44]. A related direction of research focuses on our perception of motion and its connections with the perception of occlusions [56, 61, 34].

Several neuroscientists suggested neural interpretations of the grouping processes at work in the early stages of vision [51, 20, 19, 21]. According to Treisman [51], “*some mechanism is needed to bind the information relating to each object and to distinguish it from others. Possible candidates include cells tuned to conjunctions of features, spatial attention, and synchronized firing across separate but interconnected areas of the brain.*” From that perspective, the collaboration of Gestalt handled in the present paper is the algorithmic counterpart of the binding problem. In the same direction, the authors of [20] claim that the interaction laws between visual neurons might be used to design universal computer vision systems. Unfortunately, their description of the neural mechanisms is not complete enough to give birth to an algorithm. The grouping process is formalized as a logical inference process in [10]. Although this paper has a more computational point of view, it does not provide experiments on real or synthetic images.

Computer vision aims at a mathematical and computational formalization of the grouping machine [32], and eventually at algorithms analyzing automatically images. There have been several attempts to formalize aspects of the Gestalt program to this discipline. A first answer is provided by information theory, which permits to interpret Gestalt laws as the result of efforts of the visual system toward an optimal scene encoding. In his pioneering work, Attneave [1] suggests that most of the perceptual information in images is statistical and should be represented by sparse statistics (mean and variance on regions for instance), while the information of contours is contained in a filtered out curvature. This idea is taken up by [11], in which a formal expression of the information contained in a contour is proposed. However, this paper does not take into account the statistical aspect of curvature information. Indeed, as underlined by Attneave in [1], when observing a cat, one does not observe the peak of curvature of each hair, but rather an average contour. Information theory also underlies the work of Leclerc [30], which tackles the image partitioning problem in terms of a “descriptive language that is simplest in the sense of being shortest”. More recently, the use of the MDL (Minimum Description Length [43]) formalism in image segmentation problems [62] follows the same trend. However, some arguments against a global minimum principle are presented in [5].

Sarkar and Boyer reviewed some of the attempts to formalize perceptual organization in computer vision systems in [46]. Most early attempts and a general program were proposed in the founding book by David Marr [32]. His program is “bottom-up” and assumes that low-level visual primitives (like blobs, edges, corners) should be hierarchically grouped to achieve object perception. The formalization of these hierarchies was supposed to be a sort of logical programming based on the accumulation of empirical knowledge. Thus, it somewhat contradicted the dominance of geometry on empirical knowledge in Gestalt theory. In [31, 59, 58] was introduced the first idea of a general grouping principle: Every spatial relation which is unlikely to have arisen by accident should create a group. This generic Gestalt formation law is often called *Helmholtz principle* or *principle of non-accidentalness*. The problem of the so-called *amodal completion* is studied and formalized in [27], drawing on the concept of *relatability*, by which a



piece of edge which is discontinued can be related to another which is also discontinued in its main direction and has a similar direction. More computational attempts using similar ideas are developed in [48, 22, 49, 47]. In [14, 42, 40], Gestalt cues (convexity, parallelism or T-junctions) are used to resolve the fundamental figure/ground assignment problem. Formalizations of collinearity, cocircularity, proximity, parallelism, and symmetry were presented in [63] in a Bayesian framework. Chomskyan-Bayesian approaches have also proposed to analyze images by building up a grammar of visual elements [64, 23]. A formalization of the *Prägnanz* Gestalt was presented in [50]. The compositional model of [4] recursively composes primitives into a structure and ambiguity is handled by the minimum description length principle. Most of these approaches rely on a Bayesian approach, which notoriously assumes a learning phase to learn empirical probability distributions. Thus, they contradict the Gestalt assumption of existence of direct geometric grouping laws.

But we should not oppose formalizations which actually take things at very different levels. Nothing prevents a computational learning theory to address the early phase where geometric laws themselves were formed from experience and selection in early biological perception. The Gestalt laws seem to describe the result of this learning in advanced perceptual beings, at the stage where this biological learning seems to have converged toward geometric laws.

The Helmholtz (non-accidentalness) principle is an attempt to formalize the geometric laws, not the learning theory by which basic Gestalt emerged. Desolneux, Moisan and Morel [7, 8] formulated the Gestalt grouping laws as independent detectors for geometrical events (based on the non-accidentalness argument) and called them “partial gestalts”. In their original formulation the result of each detector, or partial gestalt, was not directly comparable to the others. Each one produced a particular kind of analysis, and the problem of how to compare, relate, collaborate and solve conflicts between different partial gestalts was stated but left open in [8].

The present paper proposes a new direction to attack the question of collaboration of partial gestalts. We argue that the Gestalt theory itself provides an experimental setting to this question, which can be translated into computation terms. This answer is the so called *Nachzeichnung*, meaning “drawing after a model” set up, extensively described in [35]. The formalization proposed here is general, but will be tested first on a reduced set of Gestalt images, roughly formed by flat zones separated by straight edges. This, in accordance with the Gestalt experimental credo, avoids the prolificacy of natural images, but still tries to retain a rich enough set of samples to cover the general perception problem. In our current digital experiments the basic primitive will be the (automatically detected) line segments from real digital images. We will therefore study all grouping laws that start from segments. This keeps a rich enough set of laws including most if not all Gestalt primitives: parallelism, constant width, several kinds of good continuation, and convexity.

The image in figure 2 illustrates the method. This figure can easily be described geometrically by a set of line segments bounding the rectangles. But what are the relevant properties of this set of edges from the Gestalt point of view? This image is perceived as three black bars of different sizes. So the relevant properties are that the set of edges is organized into three sets of connected edges, each one delimiting one of the bars. Then, each of these three sets is composed of four parts, each one somewhat aligned (the sides), and forming right angles between them. The lower side of each of the three

¹We drop here and in the sequel the German initial capital on purpose, to mark the difference between gestalt in this computational theory and the references to the original *Gestalt* theory.

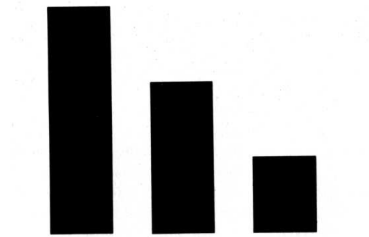


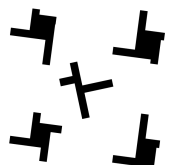
Figure 2: A simple figure from [25].

groups are aligned. Also, the width of the three bars is the same. However, the particular width of the bars, or the particular separation between them, or even their exact heights are not perceptually relevant in this figure. (The relative heights of the bars are nevertheless relevant). If the Gestalt analysis is correct, a reduced set of geometric relations on the segments should specify how this and any similar figure is organized, at least for the low level perception. From it, a similar if not identical figure can be reproduced if all relevant geometric relationships have been perceived, which is precisely what gestaltists called a *Nachzeichnung*.

The problem of finding all relevant geometric relations in a figure is related to the *vectorization* problem in computer graphics. A *vector graphic* is a figure determined by a geometrical description, while a *raster graphic* is a figure described by its pixels. In the former case, the geometrical description lets us produce an image, or *raster graphic*, at an arbitrary resolution. For example, a vector description of figure 2 would be a list of the coordinates of the three bars. With that description one can draw the figure to any desired resolution. The process of vectorization is the opposite. It takes an image like figure 2 and produces a geometrical description of it, which is nothing but a list of geometric relations, or *gestalts*.

Thus Gestalt analysis is, in a sense, a symbolic vectorization problem, because we would like to retain merely the geometric relationships, and discard the other quantitative values that would be also discarded by a person performing a drawing after the model (*a Nachzeichnung*). Thus the goal is to perform a symbolic vectorization revealing the geometric relations, but also the hidden degrees of freedom left on the figure. On the same example of figure 2, in addition to changing the image resolution, we would like to be allowed to change the width of the bars, their separating space, or the height of each one. Indeed, from the Gestalt analysis point of view there are only a few partial *gestalts*: the fact that there are three rectangular bars, that they share the same width, that they are aligned in the baseline, and that they have decreasing heights values with equal decreasing steps. The other features, like the bar widths, are arbitrary and therefore should be allowed to vary, as shown on figure 3. For performing this symbolic vectorization in the general case, we must be able to select automatically the perceptually relevant properties, and to discard those which are irrelevant. A Gestalt analysis is a geometrical description of the perceptually relevant properties only.

According to Kanizsa [25], the Gestalt grouping laws include proximity, similarity, continuity of direction, good continuation, tendency to convexity, closure, common region, connectedness, constant width, symmetry, common motion, and *prägnanz* (or law of the good Gestalt). Nevertheless, the computational formalization requires some changes even in the terminology. For example, the complex and imprecise law of



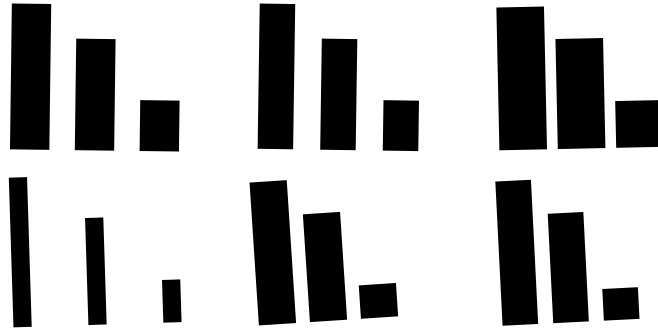


Figure 3: An example of *Nachttanz* automatically performed by our algorithms on the image of figure 2.

Prägnanz will be replaced by a set of related structures, including alignments of points and a stroke detector (related to width constancy and parallelism).

A problem that arises immediately is the validation of the geometrical description performed by a complete Gestalt analysis. How do we know that we are producing reasonably complete results? The problem is rendered more complex by the conflicts and collaboration between laws that preclude an independent evaluation for each one. The main contribution of this paper is a numerical method to perform a joint evaluation of any given set of perceptual grouping algorithms. The proposed test is the generalization and computational analog of the *Nachzeichnung*.

In the *Nachzeichnung* Gestalt experimental setting, a figure was shown to a subject for a varying period of time. Then, the subject was asked to make a drawing of it. Relevant elements and perceived relations were kept in the copy, while irrelevant relations were not.

The proposed computational test literally replaces the subjects performing the *Nachzeichnung* by a computer program implementing automatic detection of partial grouping laws and a reconstruction algorithm of an image from its segments. In addition, with a computer it is actually possible to implement a dynamic time varying reconstruction. Thus, we called the result *Nachttanz*, “after dance²” because it produces a dancing figure emulating the original. Given a figure and the geometric relations found by the partial gestalt algorithms, each figure is automatically modified in a random way, but keeping all the gestalt relations that have been detected. The figure is thus gradually modified by this Gestalt dance, but hopefully in a way that respects its fundamental relations. When it does, the experimenter knows that the algorithms have captured the relevant relations; otherwise the algorithms need revision or completion. A preliminary version of the *Nachttanz* was described in [18]. In this paper we will show some static frames of this dance, but the animated version are accessible at the web page <http://iie.fing.edu.uy/~jirafa/nachttanz/>.

The book *Pattern theory* by Mumford and Desolneux [38] develops many algorithms deriving from a single and strong principle due to Grenander: to recognize a pattern, one must be able to synthesize it. The *Nachttanz* proposed in this paper is a

²*Nachttanz* seems to be an old German term, the following is a definition from <http://www.encyclopedia.com/doc/1076-Nachttanz.html>. “*Nachttanz* (Ger.). After-dance. Term applied to the 2nd of the two dance tunes which were commonly paired from the 15th to the 17th cents., i.e. pavan and galliard, passamezzo and saltarello, sarabande and gigue, etc. (The Saltarello, especially, is known by this name.)”

(much milder) application of this reconstruction principle, since our goal here is to be able to demonstrate an implicit global analysis of the pattern from many partial detectors. Thus the Gestalt dance does not give back a model of the observed pattern, like in [38], but it gives an indirect proof that the pattern structure is globally apprehended by the partial gestalt detectors.

This paper is organized as follows. Section 2 develops the *Nachtanz*. A preliminary implementation is described in section A, also examining the partial gestalts involved. Section 4 presents and discusses several experiments. Section 5 concludes.

2 Nachtanz

A natural question when dealing with perceptually inspired algorithms is how to validate them. In our Gestalt laws approach we face two problems: the validation of each particular detector, and the selection of the set of Gestalt laws to use. It is hard and perhaps impossible to evaluate one grouping law independently of all others. Gestalt laws collaborate to form structures. For example, a weak alignment may become strongly perceived if reinforced by a shape similarity of its elements. Visual information is highly redundant and an occasional missed detection can be implicitly restored by other grouping laws. Inversely, two or more laws can be in conflict and yield alternative image interpretations. Gestalt conflicts fascinated gestaltists, but computational experience shows that the collaboration question, the “binding problem” is far more challenging and promising.



Figure 4: An example of *Nachzeichnung* by children of growing ages. Taken from *Gesetze des Sehens* by Wolfgang Metzger, page 124.

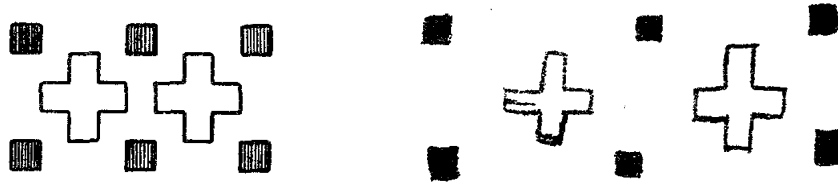
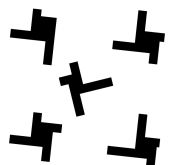


Figure 5: A second example of *Nachzeichnung* from [45].

In the *Nachzeichnung* Gestalt experimental set up, a figure was shown to the subject for a short period of time, or sometimes at some large distance, or in other unfavorable conditions. The experiments were also performed on children of growing age to assess the progress and emergence of each grouping law. Subjects were asked to make a drawing (*Zeichnung*) of the figure they had been shown. The result was invaluable to gestaltists: the relevant elements and perceived relations were kept in the copy made



by subjects, while irrelevant relations were not! Figure 4 shows some experiments with children of ages growing from four to six years [35]. One can see how the perception of various partial gestalt relations, such as the general shape, the number of dots, alignments, the presence of a square etc., improves with age. The perception of that figure differs for adults and small children, as revealed by these reproductions, reasonably satisfactory for the child who made them, but definitely not for an adult. In the second example shown on figure 5 the copy is quite similar to the original (up to the drawing skills of the subject). Only the spacing between the white crosses to the black squares is larger in the copy. The interspace was perceptually indistinguishable to the subject. The last example on figure 6 shows two copies of the same picture. The first one is quite accurate. On the second one, as in the previous example, we observe an uneven separation of the Celtic maze patterns. Again, the even separation was not perceptually relevant to the subject. The *Nachzeichnung* is a complex experimental set up involving two different subjects: one who makes the drawing, and one who compares how the first subject perceives a figure to how other humans would perceive it in good conditions.

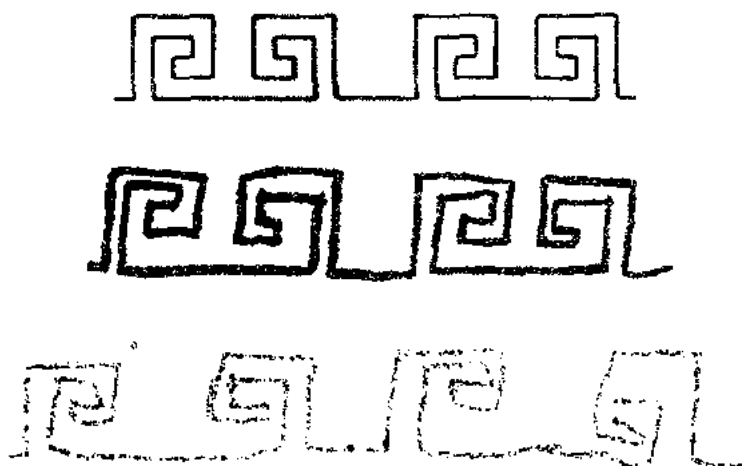


Figure 6: *Nachzeichnung* from [45].

In our proposed set up, the first subject is replaced by the machine, but the experimenters who compare the original to the computer *Nachzeichnung* are still there, and use their own perception. Thus, they actually must become the real subjects of the experiment. In the future they should be reduced to the subjects of an experimental psychophysical procedure where they are asked yes-no questions about the similarity or dissimilarity of two figures produced by the computerized after-dance.

In the Gestalt experiments, subjects drew figures that were, in a sense, modifications of the observed figures, but those modifications were compatible with the Gestalt organization they perceived. Similarly in the *Nachtanz* the algorithms modify images in a random way, but keeping the detected structures. For example, if three points were detected as aligned, then the figure can be modified in any way that maintains these points aligned. A trivial example of such a modification is to scale the original image: all geometrical relations would be maintained. But other modifications are possible.

For example, the dot pattern of figure 4 is reproduced with slight variations of the size of the dots while keeping the same interspace. Nobody will notice the difference.

The *Nachzeichnung* experiments were manual and static. Computers allow for gradual modifications of the figures into an animated “Gestaltic dance”. The more elaborate the figures for which the *Nachtanz* succeeds in indistinguishable figures, the more complete the theory will be. Our end goal will be to perform it on any digital image, and not just on scans of the Gestalt experimental thesaurus.

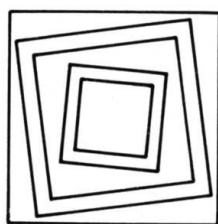


Figure 7: Scan of figure 1.44 from [25].

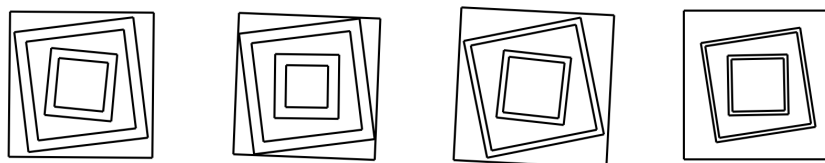


Figure 8: An example of *Nachtanz* automatically performed by our algorithm with a manually selected set of restrictions that maintain the essence of the figure.

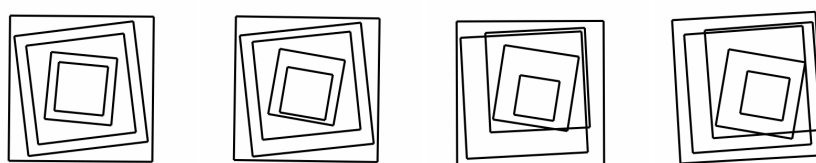
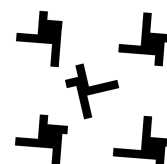


Figure 9: An example of *Nachtanz* with an incomplete set of restrictions.

Figure 7 contains five non intersecting squares, all centered on the same point, which we can number from 1 to 5, starting from the exterior square. Squares 2 and 3 are parallel, and so are squares 4 and 5. Furthermore, the distance of 2 to 3 and the distance of 4 to 5 are equal. This short description is a near complete perceptual description of the figure. Any variant of the figure that satisfies this list of relations will be termed perceptually similar or even indistinguishable at first sight. All of the mentioned relations can be expressed as geometric restrictions to the elements of the picture. This permits a computer program to gradually and randomly modify the picture, while maintaining a list of geometric restrictions. Figure 8 shows such a result. An animated version is



accessible at the web page <http://iie.fing.edu.uy/~jirafa/nachtanz/>, as well as for the other *Nachtänze* on this paper.

If more restrictions than what is perceptually needed were imposed, the figure would not move at all. On the other hand, if the set of restrictions is not perceptually complete, the “dance” will gradually destroy the essence of the figure. A manually selected set of incomplete restrictions was used to generate the samples in figure 9.

When all perceptually relevant geometric structures (partial gestalts) are formalized, it will be possible to automatically produce a valid *Nachtanz* with them. Detectors for each partial gestalt will detect the relevant structures and transform them into a set of restrictions to impose to the figure. Then the full set of restrictions will be formed as the sum of all restrictions provided by partial gestalts and used to generate the *Nachtanz*. In a sense, the full set of restrictions will therefore encode the global Gestalt of the figure.

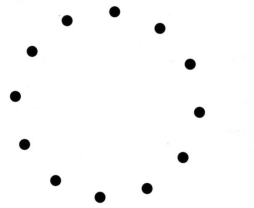


Figure 10: Scan of figure 10.03 from [25].

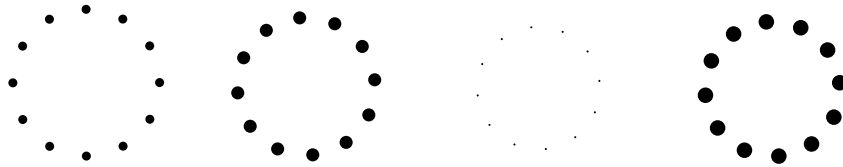


Figure 11: *Nachtanz* of the image on figure 10 with a manually selected set of restrictions.

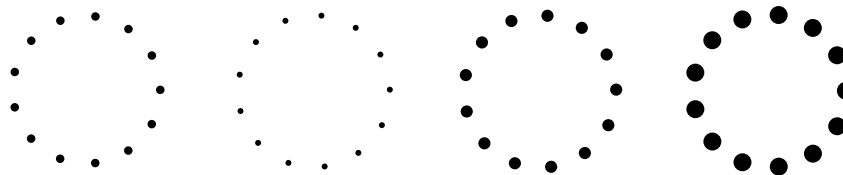


Figure 12: *Nachtanz* of the image on figure 10 with a manually selected set of restrictions. Note that in this *Nachtanz* there are 13 dots instead of 12 in the original version.

The goal is indeed ambitious. Consider the image on figure 10. Even if it is extremely simple, it involves identifying structures at various levels. A recursive analysis is probably necessary: a first level would identify the black disks, which are already complex gestalts, and a second level would be necessary to find out that these disks are uniformly distributed along a circle, a particular case of good continuation. A *Nachtanz*

generated with a manually selected set of restrictions is shown on figure 11. Actually, the *Nachtanz* of this figure should be able to change (at least to some extent) the number of dots, which would be perceptually indistinguishable as shown on figure 12. This is a further challenge.

In the rest of this paper we will show some preliminary results of this research program toward automatic gestalt analysis and its *Nachtanz*. The next section details a first implementation of the *Nachtanz*, working on a reduced set of partial gestalts. Then, section 4 shows the current results.

3 Preliminary Implementation

In this section we briefly describe our preliminary formalization of some partial gestalts and the *Nachtanz* generator. The algorithms used to detect those gestalts and implement the respective restrictions in the *Nachtanz* are more thoroughly detailed in the appendix A. The structures currently analyzed are line segments, line segment chains, alignments of points (vertices of chained line segments), a restricted form of parallelism (also related to the constancy of width) that we called *strokes*, modes of angles between chained line segments, and modes of stroke widths. Our current implementation of the *Nachtanz* is limited to figures composed by straight contours defined by control points. A limited set of geometrical restrictions are currently implemented. Nevertheless, the algorithmic chain starts from a raw digital image of the figure and proceeds at all stages with automatic detections. Several simple figures actually get the correct and autonomous *Nachtanz*.

3.1 Line Segment Chaining

Line segments give the main basic geometric information on almost any image: each one corresponds to a locally straight part of an edge. It is statistically observed in digital images that detected line segments are connected to other line segments forming a vertex, or being part of a smooth edge locally approximated by line segments. Both cases are illustrated on figure 13. The line segments detected by the parameterless and fully automatic algorithm LSD [17] provide a good description of the image provided that the relative positions of the line segments are considered. Thus, detecting and linking adjacent line segments belonging to the same edge is a relevant task and should constitute the basis of the *Nachtanz*. Figure 14 shows the result for the image in figure 13. The linking algorithm is described in section A.1.

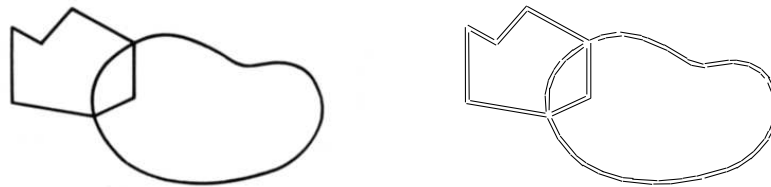
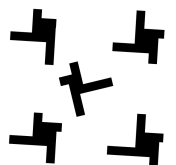


Figure 13: An image (left) and the line segments detected by LSD [17] (right).



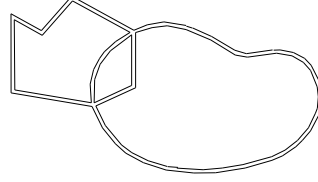


Figure 14: Result of the line segment chaining algorithm described in section A.1 for the example in figure 13. Linked line segments share endpoints.

3.2 Alignment of Points

The alignment of points is one of the basic geometrical structures in images. Figure 15 shows an example. To exemplify why alignment detection is a subtle task, the conspicuous alignment on the left figure was also placed in the two examples on the right in figure 15. Observe how it simply vanishes perceptually on the right, an effect that Gestaltists called *masking by texture* [26]. The algorithm for detecting point alignments is described in section A.2, and the respective restriction is described in section A.4.

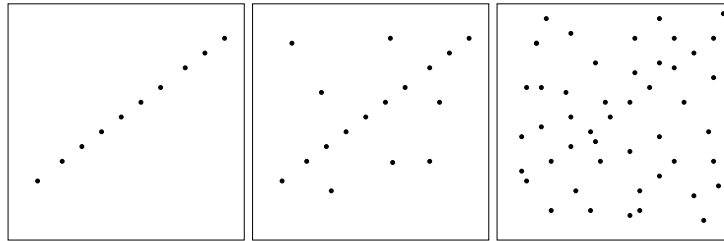


Figure 15: An alignment of dots and an increasing quantity of random dots. The perceptually visible alignment of the left vanishes on the right image.

3.3 Strokes

We now consider the analysis of parallelism, in a restricted form that we have called “strokes”. What we are looking for is a zone of the image generated by a straight stroke of uniform color. Such a stroke produces parallel line segments of opposite orientation, with no other detectable edge between them, see figure 16. There are two kinds of strokes: dark and light. In *dark* strokes the uniform color between the two line segments is darker than the surrounding, while in *light* strokes it is the opposite. This geometric structure is an implementation of the *constant width* Gestalt law studied by Kanizsa [25]. The strokes detection algorithm is described in section A.3.

3.4 Angle and stroke width modes

A particular type of the similarity Gestalt law is the similarity of angles and width of elements. We propose to handle these structures by detecting relevant groups of similar angles and width and imposing them as perceptual restrictions to the figures. The



Figure 16: A straight stroke with uniform color (the shaded rectangle) would produce two line segments, one on each side (the arrows).

key point is to decide when similarity for some numerical value (angle, distance, width, orientation, etc.) is “relevant”. As shown in Delon et al. [6, 8], this amounts to building a histogram mode detector. Based on the *a contrario* approach and the Helmholtz principle, this mode detector is parameterless and can be used indistinctly for all proposed angle, orientation, length, or width histogram.

Since our analysis relies on chains of line segments, the natural measure of angles of elements is the angle between two chained line segments. Figure 17 shows how we measure it. This process gives one angle value for each pair of chained line segments or vertex. With them, a histogram is built and meaningful modes are detected by the variant of Delon et al. method [6] for circular histograms. The process is similar for the widths of dark strokes, using the non-circular variant of the Delon et al. method.

Now, note that the same process permits to handle a special case of the proximity law in a simple way: by looking for the similarity of the light strokes widths. Indeed, in the set of Gestalt images with straight elements, the separating space is often delimited by parallel contours, thus producing light strokes according to our definition.

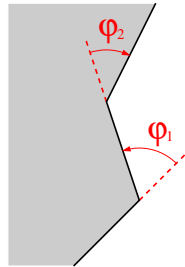
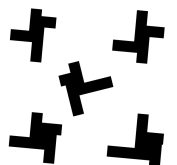


Figure 17: Measurement of the angle between chained line segments. The black line represents three chained line segments, separating a dark area on the left from a light area on the right. The angle is measured from the extension of the first line segment to the next one. The order of the line segments is important, since the angles are signed (or equivalently, measured in the range from 0 to 360 degrees). The convention we use is that a positive angle (as φ_1 in the figure) corresponds to a convex area on the dark side, while a negative angle (as φ_2 in the figure) corresponds to a convex area on the light side.

A natural extension of the method will also look for the modes of the lengths of the line segments. Some examples of angles and width histograms are shown on section 4, with the resulting modes and their meaning for the analysis of images.

3.5 Nachtanz Implementation

Our implementation of the *Nachtanz* is based on restricted perturbations of points. The first step is to parameterize the figure by control points. Figure 18 shows how a figure composed of line segments is parameterized by their endpoints. The figure



information is decomposed into two parts: a list of points with their coordinate values, and a list of couples of points between which a line segment must be drawn. Given that decomposition, the *Nachtanz* is obtained by a random perturbation of the control points subject to the partial gestalt restrictions that have been automatically detected on the original figure. Images of this Gestaltic dance are regenerated using the new points.

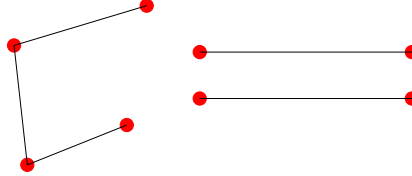


Figure 18: Line segments parameterized by its endpoints.

To impose the restrictions to the control points, we opted for a variational formulation. An energy is defined on a set of points that takes a zero value when the restrictions are perfectly satisfied, and positive values when not. The energy is defined as a sum of terms, one for each restriction imposed. Ideally, a *Nachtanz* would be a trajectory of points through null energy configurations. In practice, however, it is generated by an iterative algorithm that approximates it by alternating perturbation and correction steps (see the following pseudo-code). In this Monte Carlo exploration, a random point is selected and a random perturbation (a small displacement) is added. Then, correction terms are applied to each point and for all the restrictions imposed to this point. This step is repeated until the restrictions are satisfied up to a precision δ . Details about the definition of the energy and correction terms are given in section A.4.

Algorithm: Nachtanz for a set of points P and a set of restrictions R .

do forever:

 Select a random point $p_i \in P$.

 Select a random perturbation n .

$p_i \leftarrow p_i + n$

while RestrictionsError(R, P) $> \delta$:

for all $p_j \in P$ **with** $j \neq i$:

for all $r \in R$:

$p_j \leftarrow p_j + \lambda \cdot \text{CorrectionTerm}(r, p_j)$

A set of restrictions R can be inconsistent, meaning that no configuration of points can satisfy them. Otherwise, starting from a set of points P satisfying R , one can apply an arbitrary perturbation to one of its points and there exists at least one configuration P' also satisfying R : the translation of the full set of points always satisfies R . However, there is no guarantee that the simple algorithm that we use will find such a configuration. To improve the chances of success, the algorithm advances by small perturbations in an attempt to simulate a “differential dance”. We also limited the number of correction steps: if the precision required is not obtained, the perturbation is rejected and a new one is attempted.

The *Nachtanz* as described selects a random point and a random perturbation at each step. The result is a controlled joint random walk of all points. Therefore a large number of steps is needed to observe a shy motion. Indeed, without restrictions each point would follow a classic random walk which is slow by nature. The restrictions

reduce even more the observed motion by relating and averaging the different perturbations. In an attempt to produce a more dynamic *Nachtanz*, each random perturbation is applied repetitively to the same point for a fixed number of iterations, or until the correction loop is unable to satisfy the restrictions. Only then, a new point and a new perturbation are selected. The resulting dance shows bursts of coherent motion for short periods of time. The computation time needed to obtain interesting motions is thereby much reduced, and the aggressive motion does a better job illustrating the constraints imposed to the figure and the degrees of freedom still available.

4 Experiments

We first illustrate the results of the *Nachtanz* on the simple image shown on figure 2 for which a complete analysis is reached. More complex images are dealt with in continuation. Our sample images are obviously very simple, being described exclusively by straight edges. Nevertheless, we show that they contain enough complexity to produce interesting developments, and even to overwhelm the procedure.

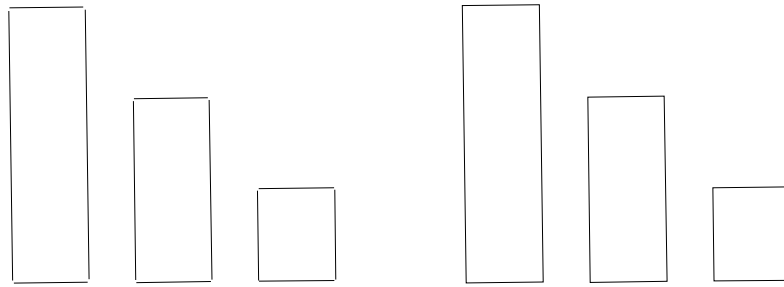
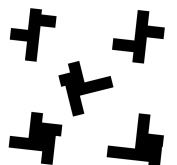


Figure 19: Left: line segments detected on the image of figure 2 by the LSD algorithm [17]. Right: result of the line segment chaining algorithm for the same image.

Figure 19 shows the line segments detected in the image of figure 2 and its chained version. In these simple images, the first step of the line segment chaining does a good job, providing us with a full set of edges in geometrical form, composed of linked line segments. This can be termed a first-order *vectorization* of the figure, meaning that we are able to reproduce it at an arbitrary resolution. With our current tools we can perform this kind of first-order vectorization only for the reduced set of images formed by flat zones divided by straight edges. Yet, this is not a full-fledged vectorization that should include all the other relations, such as the relevant angles, widths, lengths, etc., that would allow to reproduce the figure, not only at an arbitrary precision, but also allowing to change parameters like the width of bars, the height of each one, or the spacing between them when they do not form meaningful modes.

The point alignment step is applied to the vertices obtained after the line segment chaining; in this way we avoid artificial alignments produced by counting twice common endpoints. Figure 20 shows the alignment of control points for the same image. The detected alignment corresponds to the baseline to which the three bars are aligned.

The stroke detector provides us the information about parallelism, see figure 21. In this case, six dark strokes are found, two for each bar: one vertical, and one horizontal, almost coincident in the drawing. All put together, these partial gestalts imply that



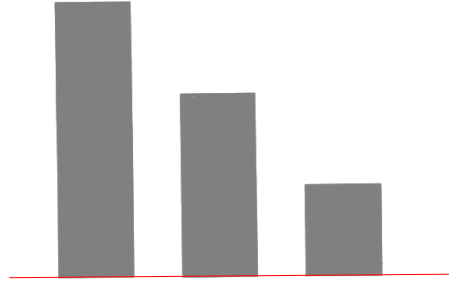


Figure 20: Alignment automatically found of the vertices of the image in figure 2. It corresponds to the aligned base of the bars.

the bars are parallelograms. Two light strokes are found, one for each bar-interspace, implying the parallelism between bars.

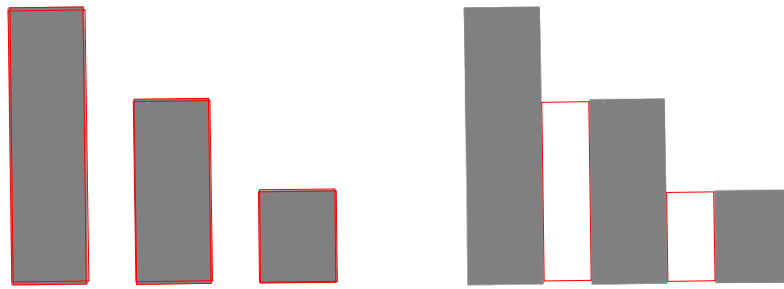


Figure 21: Dark (left) and light (right) strokes found on the image of figure 2. Each stroke is drawn in red as a rectangle, including the two parallel sides forming it. Note that there are two dark strokes detected for each bar: one vertical, corresponding to the vertical parallel sides of the bars; the other, horizontal, corresponding to the horizontal and parallel sides of the bars.



Figure 22: Top: histogram of angles between chained line segments corresponding to the image of figure 2. Bottom: meaningful histogram mode detected.

Figure 22 shows the histogram of angles between chained line segments. All of them are close to 90 degrees, so it is no surprise that the mode detector algorithm finds

only one mode, see figure 22 bottom. This partial gestalt detection imposes that all the angles grouped by this mode should share the same angle; this implies, in fact, that the bars are rectangles, since only *right* angles can make all four angles equal. This is an interesting example of *implicit* gestalt detection. Kanizsa certainly considered right angles as a lawful gestalt. But here, it is enforced implicitly by the general similarity grouping law. The histograms for width of dark and light modes are shown on figures 23 and 24, respectively. The modes detected impose that the three bar widths be equal (dark strokes), related to the Gestalt similarity law, and that the two interspaces be equal (light strokes), related to the proximity law.

This concludes the list of partial gestalts that are currently implemented. In the following, we show different *Nachtänze* for the same figure, starting from a stripped-down one including only line segment chaining, and gradually adding partial gestalts so we can see the contribution of each one. The first example is shown on figure 25. The full animations are available at <http://iie.fing.edu.uy/~jirafa/nachtanz/>. As we can see, the figure continues to be formed by three black elements gradually degenerating into amorphous shapes.

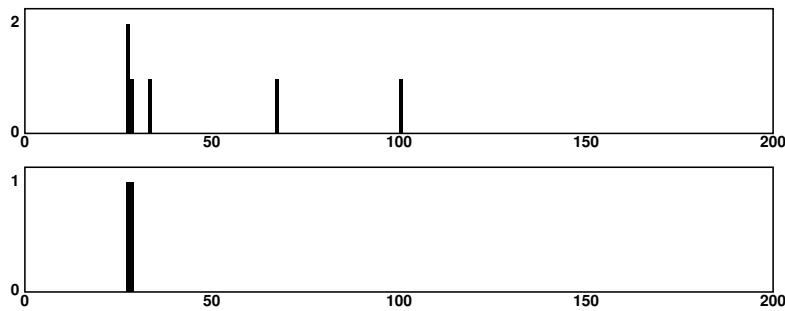


Figure 23: Top: histogram of dark strokes widths. As we can see, three of them are almost equal and correspond to the width of the three bars. The other three correspond to bars' heights and are different. Bottom: meaningful mode found.

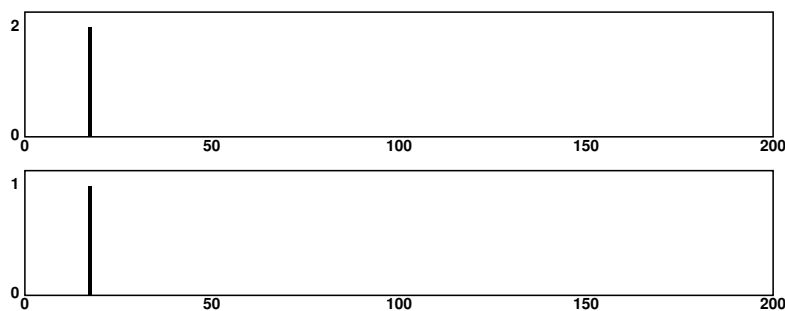


Figure 24: Top: histogram of light strokes width. Bottom: meaningful mode found.

For our next example, on figure 26 we added the restrictions imposed by the dark and light strokes, parallelism being thus enforced. The black elements become parallelograms, producing a perceptual difference to the original bars. With the addition of the angles mode partial gestalt, on figure 27, the parallelograms turn to rectangles and the

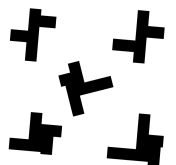


image gradually takes shape. But still, the misalignment of the bar bases is noticeable, as well as the uneven spacing.

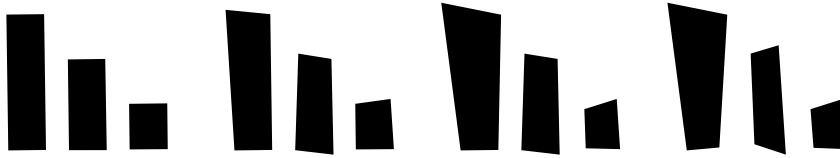


Figure 25: *Nachtanz* of the image on figure 2 only using the line segment chaining partial gestalt.

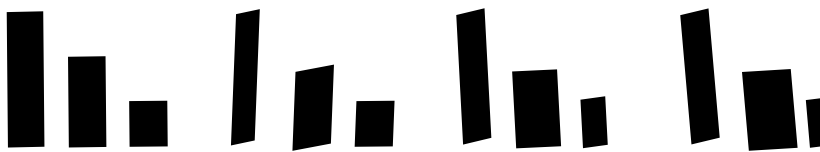


Figure 26: *Nachtanz* of the image on figure 2 only using the line segment chaining, and the dark or light stroke partial gestalt.



Figure 27: *Nachtanz* of the image on figure 2 imposing the restrictions defined by the following partial gestalts: line segment chaining, dark and light strokes, and modes of angles between chained line segments.



Figure 28: *Nachtanz* of the image on figure 2 resulting from the imposition of the partial gestalts: line segment chaining, dark and light strokes, modes of angles between chained line segments, and vertices alignments.

Adding the vertices alignment partial gestalt we obtain quite good results in figure 28, but it is evident that the difference in the bars width is missing, as well as the equal spacing. Both constraints are added with the modes of dark and light stroke widths. This completes the set of partial gestalts dealt with in this paper and produces the *Nachtanz* shown on figure 29.

We strongly encourage the reader to watch the corresponding animations at <http://iie.fing.edu.uy/~jirafa/nachtanz/>, as they give a much clearer impression of the “Gestaltic dance” as systematically exploring the degrees of freedom still left in

the picture. Still, there are some objections that prevent us to say that the figure is truly perceptually indistinguishable to the original. There are some extreme cases, like the very thin bars, that pop up among the others. Also, the height order inversion, as can be seen on the second row of figure 29 seems to change the essence of the image. The strictly decreasing order of the bar heights is probably a perceptually relevant feature which should be included as a new partial gestalt. In this sense, the gestalt dance is definitely a way to identify new missing Gestaltic relations, not actually mentioned in the literature.

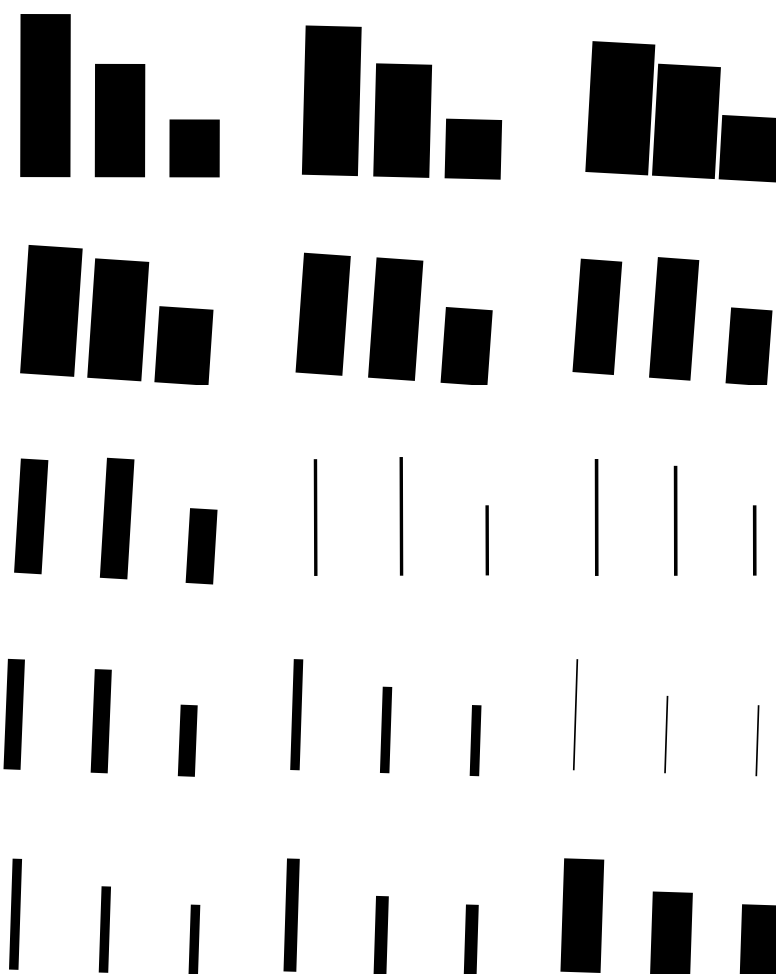
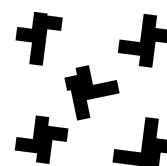


Figure 29: *Nachtanz* of the image on figure 2 imposing the full set of partial gestalts analyzed on this paper: line segment chaining, dark and light strokes, modes of angles between chained line segments, vertices alignments, and modes of dark and light strokes widths.

In order to show how critical the detection of some structure can be, we made a modification of the image in figure 2, by slightly rotating the small black bar, see figure 30. This change is enough to free the small black bar from its parallelism constraints and to break the baseline alignment: four vertices are not enough to define the align-



ment. Nevertheless, the two large black bars are still parallel, and all the three bars remain rectangular and share the same width.

Our second example of *Nachtanz* applies to the image on figure 31. This image is still simple, but perceptually more complex than the previous one. We apply the same set of partial gestalt detectors and show two different *Nachtänze*. On the first one, figure 32, the restrictions from all the partial gestalts are imposed, except for the vertex alignments that was excluded on purpose. The result is suggesting, but clearly not satisfactory as a perceptual modification of the original image. The angle modes impose right angles on the modified crosses, and the dark stroke width modes make the bars of all crosses share the same width. Also, due to the parallelism imposed by light strokes, the four external crosses are “synchronized”.

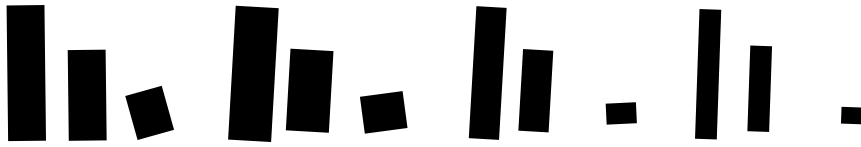


Figure 30: *Nachtanz* of a modification of the image on figure 2. The full set of partial gestalts described on this paper was used. The example shows how part of the structure is lost by a simple modification.

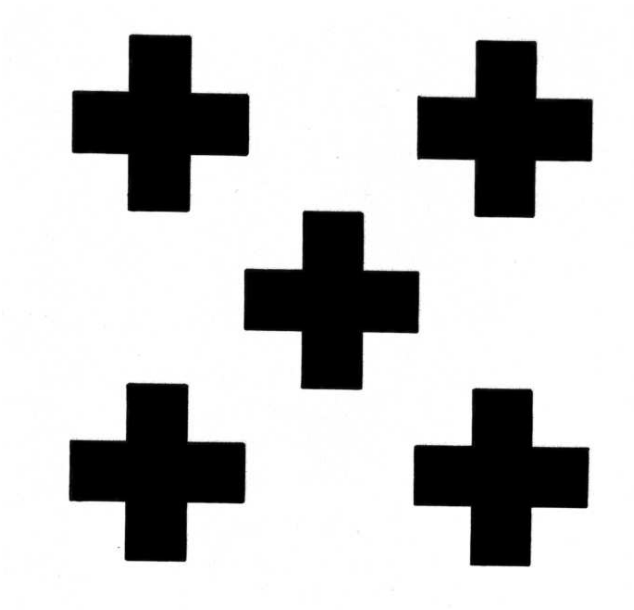


Figure 31: Scan of figure 5.20 from [25].

The restrictions imposed by vertex alignments are very strong, and their effect on the *Nachtanz* are marked as can be seen on figure 33. The figure is now perceptually structured and only three degrees of freedom are left: the global spacing between the central and the external crosses, the global angle of the figure, and the width of the elements of the crosses. All the three vary on figure 33.

Figure 34 shows the actual set of vertex alignments found on figure 31. Here, the alignment detector failed to detect some of the relevant alignments, for example the vertical alignments between the lateral crosses. However, due to the strong redundancy and collaboration of the partial gestalts present on the figure, the *Nachtanz* on figure 33 shows the correct perceptual structure. This is a first example of very simple gestalt collaboration.

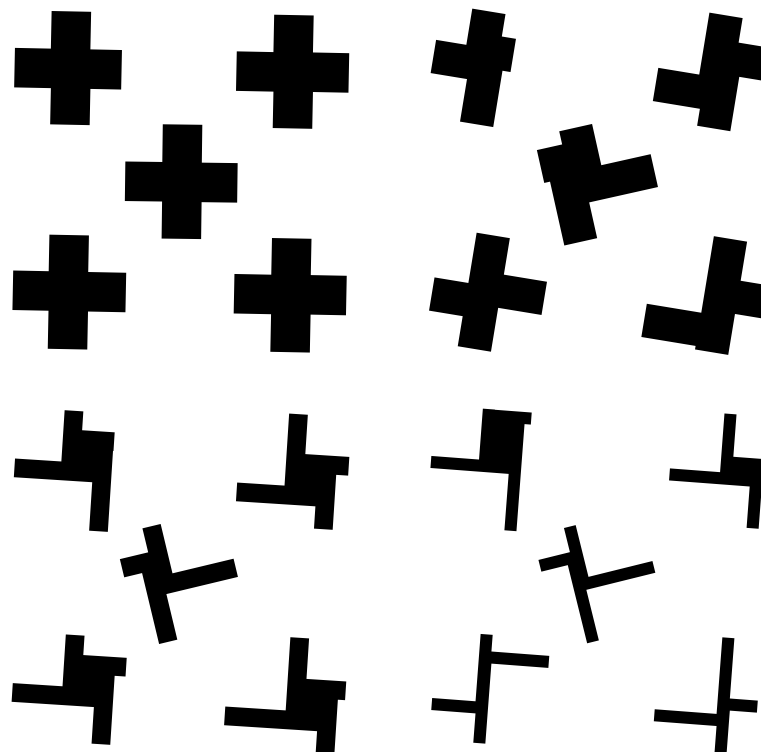
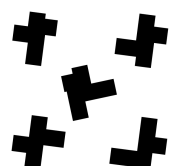


Figure 32: *Nachtanz* of the image on figure 31 with restrictions from the following partial gestalts: line segments, line segments chaining, dark and light strokes, modes of angles between chained line segments, and modes of dark and light strokes. The vertex alignment gestalt wasn't included on purpose.

In the next experiment, illustrated by Figure 35, the perceptual complexity increases notably, with the presence of amodal (or subjective) contours. Here we reach the limits of our current set of partial gestalts. Figure 36 shows the alignments detected by our algorithms. The four detections along the amodal contours help, in conjunction with the angle modes, to keep the structure of the subjective white square on the *Nachtanz*, somewhat an achievement. Nevertheless, relevant alignments are still missing, like the ones relating the black vertical and horizontal black bars, and the alignments of the thin lines forming the “square” with black border perceptually present on second plane.

Figure 37 shows the corresponding *Nachtanz*. Part of the structure is lost. The large black cross perceived on the original image behind the white square breaks down, as well as the square with black border. This changes the global interpretation of the figure. However, the fact that the amodal white square on the foreground keeps its



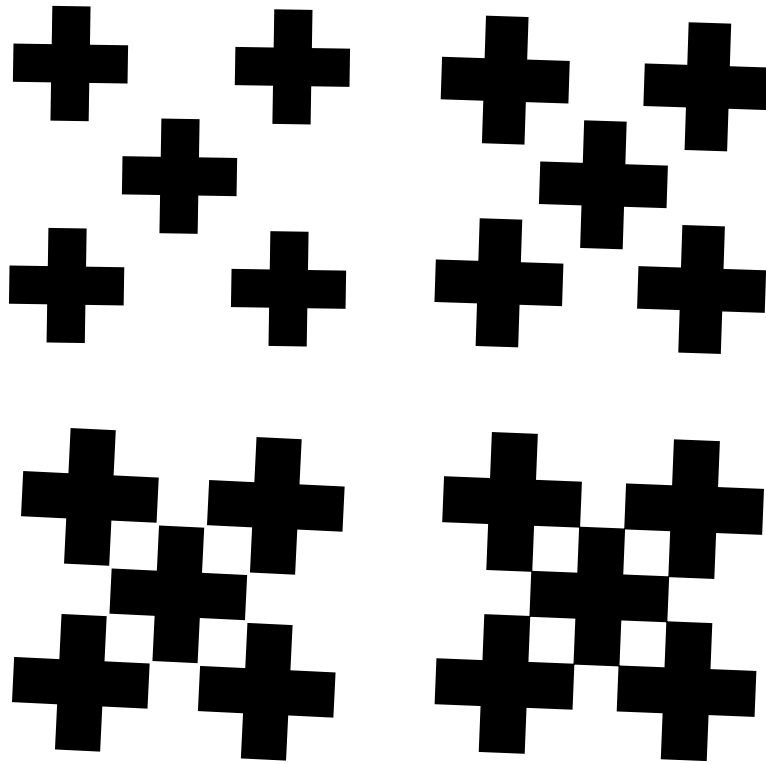


Figure 33: *Nachtanz* of the image on figure 31 imposing the full set of partial gestalts analyzed in this paper.

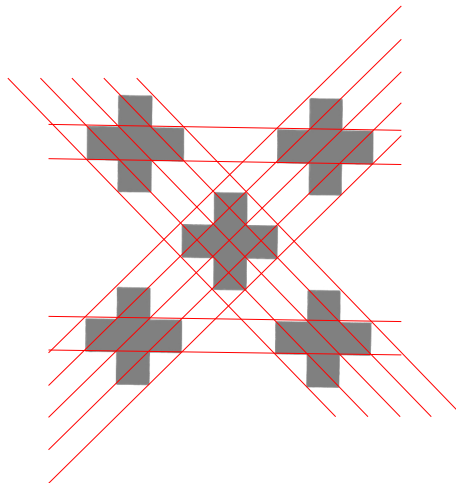


Figure 34: Vertex alignments detected on figure 31.

presence can be considered as a success due to the collaboration of partial gestalts.

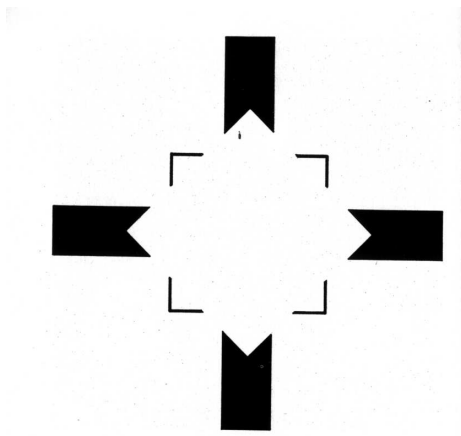


Figure 35: Scan of figure 10.03 from [25].

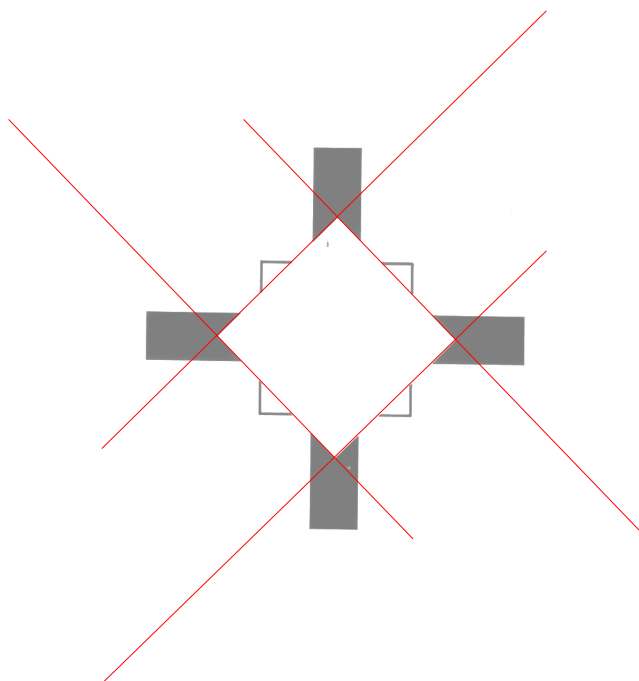
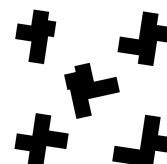


Figure 36: Vertex alignments detected on the figure 35.

The failure of the previous *Nachtanz* is due to a partial failure of the alignment detector. The next experiment shows a case where a new kind of partial gestalts is needed. The image shown on figure 38 is simple only in appearance. Indeed, it involves the concept of inclusion of the white squares inside the black ones. Inclusion is part of a list of partial gestalts that still needs to be enforced, along with symmetry and convexity.

Figure 39 shows the resulting *Nachtanz*. The inclusion is preserved on the left hand



squares due to the parallelism imposed by the dark strokes. But there are no strokes to retain the inner square on the right hand. Also, we would have expected that both squares on the left hand remain centered. It is not the case, and this shows a double detection failure. A detection of the alignment of the diagonal vertex or a detection of the constant width of black stroke would have been sufficient. Note, however, that the four squares remain squares: this is the action of the angles modes and of the modes of width. A careful observation of figure 38 reveals a dark spot at the center of the figure, between the left and right black figures. This defect is large enough to produce the detection of small line segments, as can be seen on the first frame of the *Nachtanz* of figure 39. But since no structure is associated with them, they move freely on the following frames.

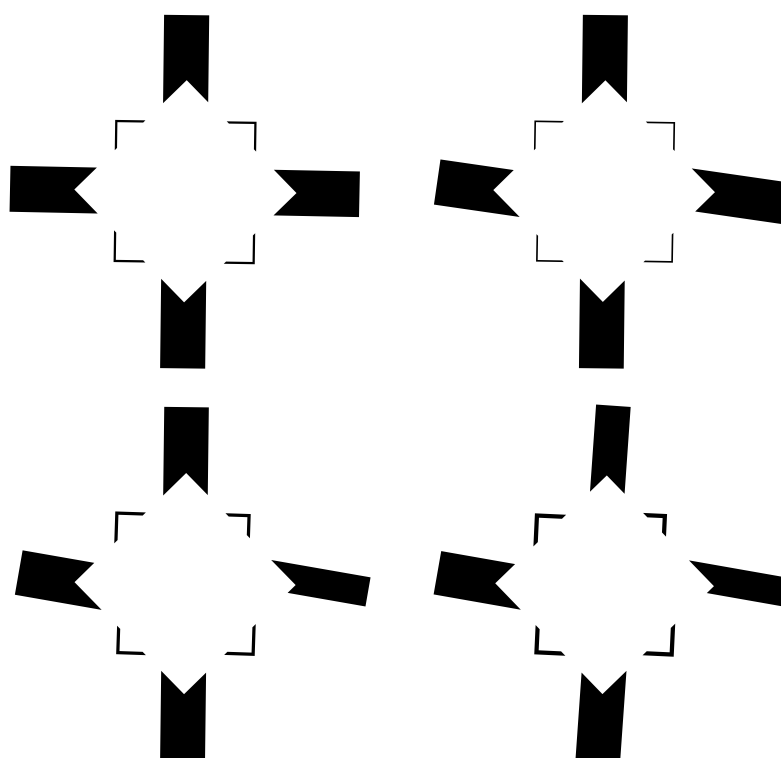


Figure 37: *Nachtanz* of the figure 35 involving the full set of partial gestalts described on this paper: line segments, line segments chaining, dark and light strokes, modes of angles between chained line segments, modes of dark and light strokes, and vertex alignments.

A second example of partial gestalt still missing is shown on figures 40 and 41. The figure is very simple, but the smooth curve stops from being so on the *Nachtanz*. The reason is that in our present set, no analysis tool is used to cope with convexity or smooth curves.

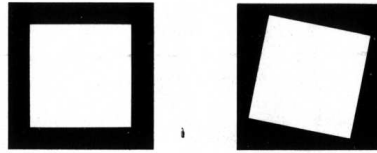


Figure 38: Scan of figure 1.47 from [25].

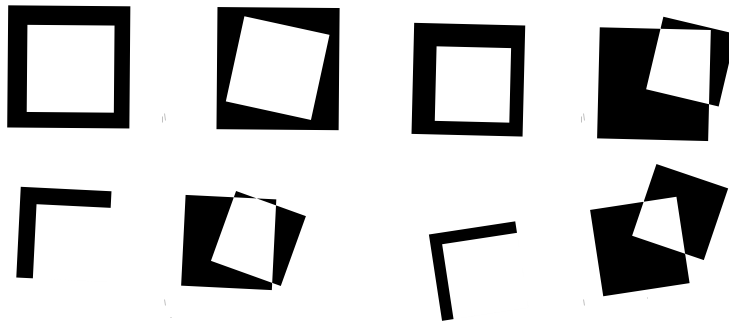


Figure 39: *Nachtanz* of the figure 38 involving the full set of partial gestalts described on this paper: line segments, line segments chaining, dark and light strokes, modes of angles between chained line segments, modes of dark and light strokes, and vertex alignments.



Figure 40: A curved shape.

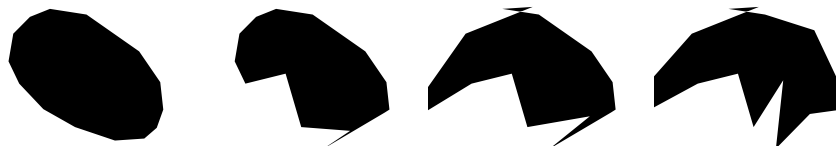
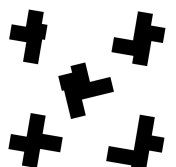


Figure 41: *Nachtanz* of the image on figure 40.



5 Conclusion

The Gestalt method did not deal with real 3D scenes or images of 3D scenes, but with graphical simple images. But it conjectured that these images constituted a complete and sufficient test bed for human vision. This was and is legitimate, in exactly the same way as it is legitimate to test a complex system with well controlled, elementary and complete impulses, like engineers or neurophysiologists do. Our purpose is to analyze this whole Gestalt corpus fully automatically. This corpus is composed of scanned images, so they are real images after all. To the best of our knowledge, there is no automated analysis in computer vision of this impressive and inventive Gestalt corpus. Certainly an algorithm able to analyze all images of this corpus will not analyze any other real image. For example there are no real textures in these geometric images, which is their main drawback. On the other hand, the goal of the paper was to demonstrate how complex and rich the discussion of the interaction of Gestalt laws becomes, even on the simplest figures of the corpus.

The authors are aware that these results will have to be confirmed on a consistent and broad enough set of digital images. The experiments herewith give a strong evidence that the sketched procedure will work soon on the whole Gestalt school figure thesaurus, therefore permitting at the very least a whole assessment of Gestalt theory at its end in the eighties of the past century. An application to the vectorization of scans of graphical images might be a practical outcome.

Some recent theories of perception stress the intimate relation between action and perception, stating that they form an inseparable *sensorimotor coupling* [33, 39, 15]. In these views, perception serves action by linking the expected sensory information to a given action. According to O'Regan & Noë [39] “*To see a bottle, for example, is to explore visual-motor contingencies, such as transformations in the appearance of the bottle as one moves in relation to it.*” Our proposal is a shy expression of these principles: each partial gestalt detector imposes restrictions to the possible modifications of the image. Citing again from [39]: “*seeing is not directly related to having a retinal image, but to being able to manipulate the retinal image.*” In this sense, the Gestaltic dance shows some (limited) abilities gained in the manipulation of the image by the detected structures.

A Implementation Detail: how each grouping law is automatically enacted on digital images

A.1 Line Segment Chaining

The fundamental condition needed to link two line segments is that they belong to the same edge. But defining edges is tricky, as shown in [16] and [17] for the particular case of straight edges, i.e., line segments. Accordingly, a chaining algorithm should follow the same ideas, defining the geometric structures of vertex and curve elements, and then verifying if the image gradient is compatible with such elements.

The simple algorithm we propose here is based on the assumption that level-lines are shared along edges. This is not true in general, as shown on figure 42. However, it is reasonably true when limited to a local region. Figure 43 shows some of the level-lines on a detail of the image of figure 13; one can see how the assumption is approximately satisfied.

The criterion is simple: two line segments (obtained by the fully automatic algorithm LSD [17]) are chained when they share image level-lines within the local zone of

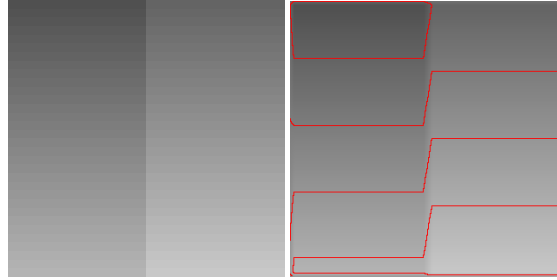


Figure 42: An image of a vertical edge over a soft gradient background (left) and some of its level-lines superposed to it (right). Image edges and level-lines are different concepts. Level-lines are shared along edges only locally.

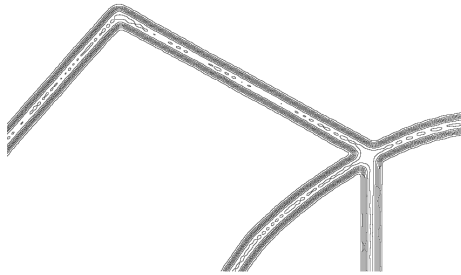


Figure 43: Level-lines of part of the image in figure 13. Note that some level-lines are shared by unrelated edges near the X crossing.

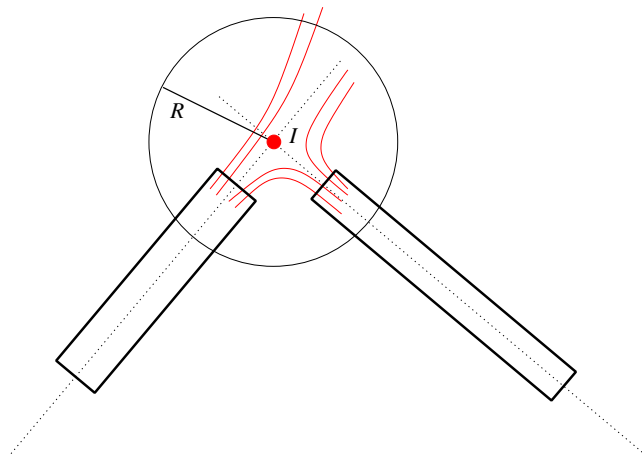
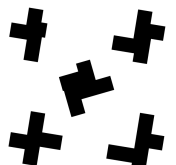


Figure 44: Here we propose a simple line segment chaining criterion: two line segments (the rectangles) are chained when they share level-lines (red curves) within the local zone defined by the intersection point I and a radius R .



their intersection, see figure 44. The level-lines considered are the ones spanning from the short side of the rectangle defining the line segment. The local zone is defined by the intersection point I of the line segments, and a fixed radius R . The full shared level-line path going from one line segment (starting at the border of its defining rectangle) to the next one should be contained in the radius R . When a chaining is accepted, the line segment linking endpoints are updated to their intersection point and become a common endpoint. Any further modification to one should also affect the other.

Our implementation uses the FLST algorithm developed by Monasse et al. [37], that provides the image level-lines of a bilinear interpolation of the image. Our algorithm depends strongly on the parameter R , clearly an unsatisfactory aspect. In our tests R was set to 10 pixels.

The result on figure 14 shows the advantages and disadvantages of this procedure. On the one hand, it shows an acceptable result given the simplicity of the algorithm, as it provides the fundamental edge structure of the image. On the other hand, the algorithm fails in some cases, even if the image is relatively simple. In curved edges, the end of the rectangle of a line segment may hang off the edge, precluding the evaluation of the right level-lines and leading to failure. Another unsatisfactory case is when two line segments are almost collinear, and their intersection is far from them, failing to satisfy the locality region criterion. Finally, on natural images with complex level-lines structures, unrelated line segments close to each other can be mistakenly chained.

A.2 Alignment of Points

Following the computational Gestalt theory, we use here a variant of the alignment detectors proposed in [7] and in section 3.2.2 of [8].

We start with a model for N non-structured points, where no alignment is expected to be present. Then, alignments are defined as large deviations from that model. In the non-structured model, the points follow a uniform and independent random law. Each point is independent of the others and uniformly distributed on the image domain. Now, a precision for the alignment is needed that will determine how far from perfectly straight an alignment can be. A reasonable family of tests, or possible alignments, is determined by a set of lines that a perfect alignment could follow. For each test, a strip is defined by the perfect alignment line and the precision used, see figure 45.

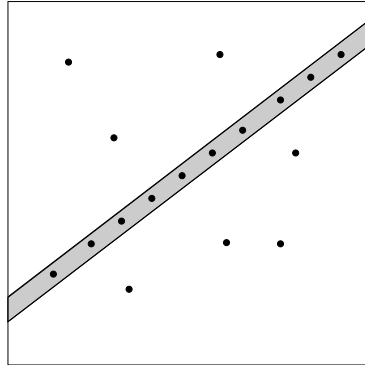


Figure 45: A strip defining the alignment and a set of dots.

Given the area $A(S)$ of the strip S and the area of the domain A_D , the probability

that a uniformly distributed point be placed in the strip is given by

$$p(S) = \frac{A(S)}{A_D},$$

and, as each point is independent of the others, the probability of observing at least $k(S)$ points in the strip out of a total of N is given by the binomial tail,

$$b(N, k(S), p(S)).$$

Now we can define the NFA for a tested strip S in the usual way as

$$\text{NFA}(S) = N_{\text{tests}} \cdot b(N, k(S), p(S)),$$

and when $\text{NFA}(S) \leq \varepsilon$ an ε -meaningful alignment is found.

There are two natural ways of choosing the family of tests to be performed. One is to use all the lines that cover the domain, provided that the central line of one alignment is not covered by the precision strip of another one. A second option is to test the lines determined by two observed points. The latter is the one we will use in the following; in this setting the number of tests is given by

$$N_{\text{tests}} = \frac{N(N-1)}{2}.$$

Now, the count of points in the strip should be modified in order to exclude the two points that define the strip. The NFA becomes

$$\text{NFA}(S) = \frac{N(N-1)}{2} \cdot b(N-2, k(S)-2, p(S)).$$

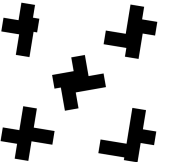
We still need to say how to set the precision, that is, the width of the strips. In some cases the problem can naturally provide it. For example, if we knew that point coordinates had been quantized, then it would be misleading to use finer precisions than the quantization step. In most cases, however, it is enough to try multiple precisions. Formally, this is equivalent to repeating each test with a different precision, so the number of tests N_{tests} must be multiplied by the number of precisions used. The final NFA is,

$$\text{NFA}(S) = \alpha \cdot \frac{N(N-1)}{2} \cdot b(N-2, k(S)-2, p(S)),$$

where α is the number of precisions tested. The range of precisions could be determined automatically using as criterion, for example, the largest p that can lead to a meaningful alignment using all the points, and the smallest p needed to validate a perfect alignment of three points. In practice this range is too large. In our experiments we used a fixed range: the coarser precision was selected to produce a probability $p(S)$ equal to 0.005, and $\alpha = 8$ different values were tried, each one a factor $\sqrt{2}$ smaller than the previous one; the finest precision corresponds to $p(S) = 3.1 \times 10^{-4}$.

This algorithm produces somewhat redundant detections, see figure 46. Different tests can be simultaneously positive for what is basically the same alignment. For example, a very meaningful alignment can be detected for various precisions. Also, more than one strip defined by different points belonging to the same alignment will produce valid detections, see figure 46 center.

An *exclusion principle*, in the same spirit as the one already used in line segment detection, see [8, 16], has been added to reduce redundancy. Each point is only allowed



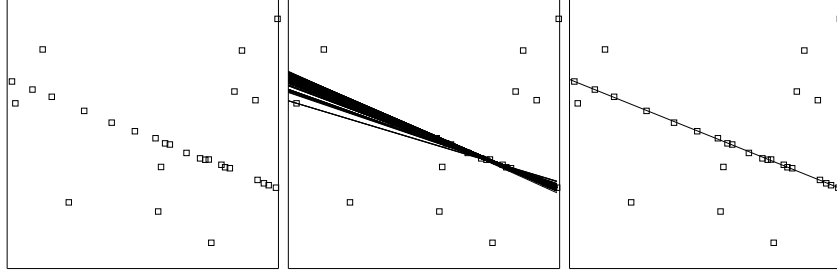


Figure 46: Left: set of points. Center: All the meaningful alignments found. Right: Meaningful alignments after applying the exclusion principle.

to be counted as belonging to one alignment, the one with the best (smallest) NFA value that contains the point. Iteratively, the alignment with the smallest NFA value is selected and validated as a detection. Then, the NFA value for the remaining alignments is recomputed, but without counting the points already used by validated alignments. The NFA value of many alignments will increase (i.e., they become less meaningful) and some of them will have $NFA > \epsilon$ and will be rejected. This process is repeated until all the alignments are validated or rejected. Figure 46 right shows one example.

A.3 Strokes

The key notion is how to define parallelism. The problem arises because no parallelism is perfect and we need a criterion to decide if two line segments can be regarded as parallel or not. We performed some experiments with different notions of parallelism. For example, a natural formulation would be to consider the angle between line segments as the discriminating variable and to use an *a contrario* setting to decide if two line segments are parallel when their angle is surprisingly similar, deviating from what could be expected on an unstructured model. Unfortunately, the criterion of the angle between lines is not consistent with the notion of human perception of parallelism, as depicted in figure 47. It is more likely that human perception is able to measure the unchanged width, as suggested by Gestalt theory.



Figure 47: The difference in angle between the lines on the left and on the right are the same. Nevertheless, the lines on the left seem more parallel than the ones on the right.

The criterion we will use is very simple, see figure 48. Given two line segments s_1 and s_2 , the mean line L is traced, with the mean orientation of line segments s_1 and s_2 . Then, the overlapping of the projections of s_1 and s_2 to L is evaluated. If the overlapping is null, then s_1 and s_2 do not form a stroke. Otherwise, the distances d_1 and d_2 are evaluated, from line segment s_1 to s_2 , and measured orthogonal to line L at the extremes of the overlapping. Line segments s_1 and s_2 form a stroke if $|d_1 - d_2| < d_{TH}$.

We need to set the threshold d_{TH} . It is clear that this threshold should be dependent on the size of the image and on what angle in the retina represents the pixel size. But we do not have a principled theory for this yet. In our test we just used a fixed threshold of 1 pixel.

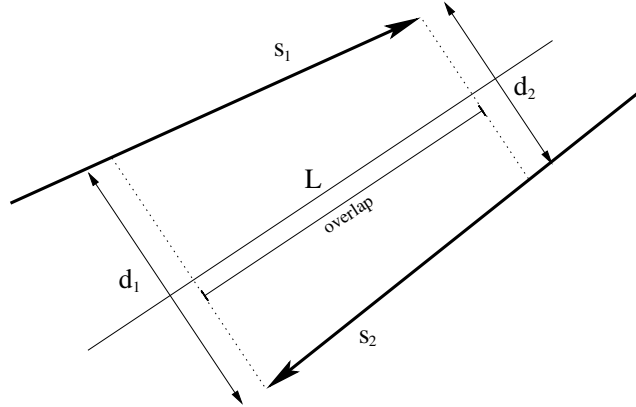


Figure 48: Definition of a stroke.

The complete algorithm to detect strokes will test all couples of line segments, and reject the ones that are not roughly in opposite orientation. Then, the mean line L , the overlapping and the distances d_1 and d_2 are computed. The stroke is finally validated or not. Note that the method can lead to dark or light strokes, depending on the order of the opposite line segments. A final step is needed to reject strokes intersected by other line segments. Actually, in our definition, only intersecting line segments *parallel* to the stroke could provoke a rejection.

A.4 Nachtanz Implementation

This section gives the energy terms for each one of the restriction types we deal with. For the experiments presented here we need to be able to impose line segment chaining, alignment of points, strokes, equal angle between groups of points, and equal width on groups of strokes. We want to achieve this with a reduced set of control point restriction primitives. Chained line segments are imposed just by using a common control point for the shared endpoint. We need a control point alignment primitive. For strokes we need a parallelism primitive and we must prevent the stroke from changing contrast and from decreasing the overlapping zone; this is imposed by a primitive for keeping control points in a half-plane defined by two control points. Finally, we have especial primitives for angles and stroke width.

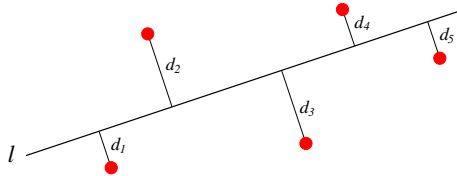
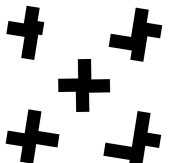


Figure 49: The energy associated to the alignment of points is the sum of the squared distance of points to the least square fitting line.

If we want to impose an alignment restriction to n points p_1, \dots, p_n , the energy is



given by the sum of squared distances to the best fitting line, see figure 49,

$$E(r_{\text{alignment}}) = \min_l \sum_{i=1}^n d^2(p_i, l),$$

where l is a fitting line and $d(p, l)$ is the distance from point p to the line l . The correction term associated to point p_i is given by $-\frac{\partial E}{\partial p_i}$. However, to simplify the formulation, we will make a quasi-static assumption and suppose that the least square fitting line l is not (significantly) modified by p_i and use

$$CT(r_{\text{alignment}}, p_i) = -2d(p_i, l)\vec{v}$$

as correcting factor to point p_i , where \vec{v} is the a normal vector from line l to the point p_i .

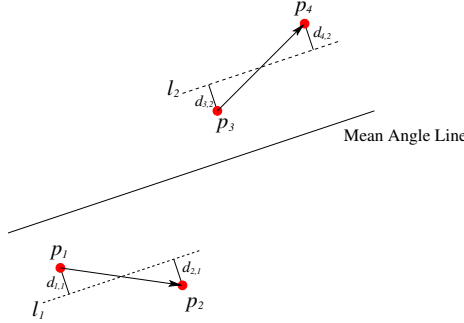


Figure 50: The energy associated to parallelism is the sum of the squared errors from points to the mean angle fitting lines.

We turn now to parallelism. As stated in section A.3, we will measure the parallelism error by the distance of points to its mean angle fitting lines, see figure 50. Similarly to the previous case, the energy term is defined by

$$E(r_{\text{parallelism}}) = d^2(p_1, l_1) + d^2(p_2, l_1) + d^2(p_3, l_2) + d^2(p_4, l_2),$$

where l_1 is the mean angle line fitting points p_1 and p_2 , and l_2 is the mean angle line fitting points p_3 and p_4 , and $d(p, l)$ is the distance from point p to line l . Again, we will make a quasi-static assumption and use

$$CT(r_{\text{parallelism}}, p_i) = -2d(p_i, l_j)\vec{v}$$

as correction term for point p_i , where \vec{v} is a normal vector from line l_j to point p_i .

We get to the half-plane restriction type. A set of points p_1, \dots, p_n is restricted to stay in the half-plane H defined by two other control points p_a and p_b . In analogy to previous cases, the energy is defined as the squared distance from the points out the half plane to the border line l defined by p_a and p_b :

$$E(r_{\text{half-plane}}) = \sum_{i=1}^n d^2(p_i, l) \mathbb{1}_{p_i \notin H}.$$

The correction term for point p_i is given by

$$CT(r_{\text{half-plane}}, p_i) = -2d(p_i, l) \mathbb{1}_{p_i \notin H} \vec{v},$$

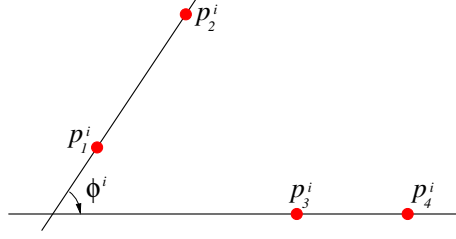


Figure 51: Measure of angle between two pairs of control points.

where \vec{v} is a normal vector from line l to point p_i .

Our next restriction is the one related to angles. For a group of four points p_1, p_2, p_3, p_4 , we will measure the angle between the line $\overline{p_1 p_2}$ and line $\overline{p_3 p_4}$. For a set of $4n$ points, $p_1^1, p_2^1, p_3^1, p_4^1, p_1^2, \dots, p_3^n, p_4^n$, we want to impose that the angles of all groups are equal. Figure 51 shows the relations for group number i .

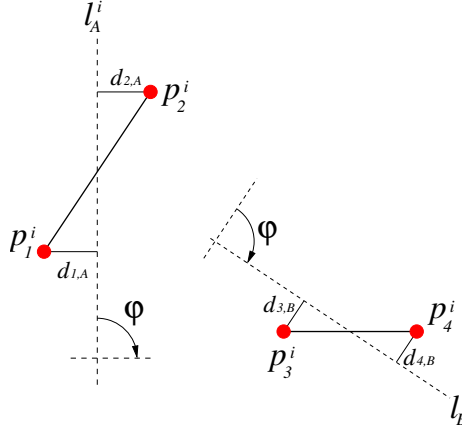


Figure 52: Relations involved in the equal angle energy term.

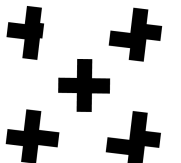
We will note φ the mean angle of the n angles ϕ^i , and the angle that all the groups should share. Now, for each group i we will define two lines l_A^i and l_B^i , see figure 52. Line l_A^i cross the middle point between p_1^i and p_2^i and has an angle φ with the line $\overline{p_3 p_4}$. In other words, points p_1^i and p_2^i should belong to line l_A^i when the restriction is satisfied. Analogously, p_3^i and p_4^i should belong to line l_B^i . With this, the energy term is defined by

$$E(r_{\text{angle}}) = \sum_{i=1}^n \left(d^2(p_1^i, l_A) + d^2(p_2^i, l_A) + d^2(p_3^i, l_B) + d^2(p_4^i, l_B) \right).$$

As in the previous cases, the correction term is obtained by $-\frac{\partial E}{\partial p_i}$ and, again, we will use the quasi-static assumption to simplify the solution: We will assume that φ is not (significantly) modified by applying the correction term to point p_j^i . Then we have,

$$CT(r_{\text{angle}}, p_j^i) = -2d(p_j^i, l_X^i) \vec{v},$$

where l_X^i is l_A^i for j to 1 or 2, and l_B^i otherwise, and \vec{v} is a normal vector from line l_X^i to point p_j^i .



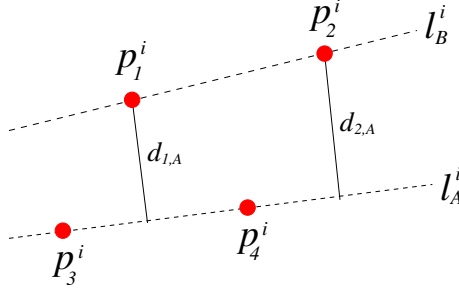


Figure 53

The last restriction to impose is equal stroke width. We want to impose equal width to n strokes, so we will refer to $4n$ points, four for each stroke. For stroke i , points p_1^i and p_2^i define one side of the stroke, and points p_3^i and p_4^i form the other side, see figure 53. We will call l_A^i the line defined by p_3^i and p_4^i , and l_B^i the line defined by p_1^i and p_2^i . These two lines should be parallel, but this can be false in intermediate steps and should be contemplated by the way width is measured. A simple solution is to define the width as the mean of the distance from each of the points of the strokes to the line of the opposite side. Then,

$$\text{width}^i = \frac{d(p_1^i, l_A^i) + d(p_2^i, l_A^i) + d(p_3^i, l_B^i) + d(p_4^i, l_B^i)}{4}.$$

We will call ξ the mean width of all strokes. Now, the energy term for equal stroke width is defined by

$$E(r_{\text{width}}) = \sum_{i=1}^n \left((d(p_1^i, l_A^i) - \xi)^2 + (d(p_2^i, l_A^i) - \xi)^2 + (d(p_3^i, l_B^i) - \xi)^2 + (d(p_4^i, l_B^i) - \xi)^2 \right).$$

Using the same quasi-static assumption, we will assume that ξ is not (significantly) modified by applying the correction term to point p_j^i and we get the correction term

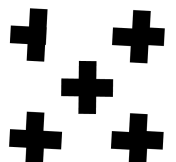
$$CT(r_{\text{width}}, p_j^i) = -2(d(p_j^i, l_X^i) - \xi)\vec{v},$$

where l_X^i is l_A^i for j to 1 or 2, and l_B^i otherwise, and \vec{v} is a normal vector from line l_X^i to point p_j^i .

References

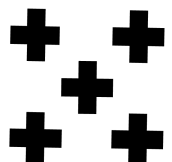
- [1] F. Attneave. Some informational aspects of visual perception. *Psychological review*, 61(3):183, 1954.
- [2] H.B. Barlow and B.C. Reeves. The versatility and absolute efficiency of detecting mirror symmetry in random dot displays. *Vision Research*, 19(7):783–793, 1979.
- [3] J. Beck, A. Rosenfeld, and R. Ivry. Line segregation. *Spatial Vision*, 4, 2(3):75–101, 1989.
- [4] E. Bienenstock, S. Geman, and D. Potter. Compositionality, MDL priors, and object recognition. *Advances in neural information processing systems*, pages 838–844, 1997.

- [5] F. Boselie. Local versus global minima in visual pattern completion. *Perception & Psychophysics*, 43(5):431–445, 1988.
- [6] J. Delon, A. Desolneux, J.L. Lisani, and A.B. Petro. A nonparametric approach for histogram segmentation. *IEEE Transactions on Image Processing*, 16(1):253–261, 2007.
- [7] A. Desolneux, L. Moisan, and J.M. Morel. A grouping principle and four applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003.
- [8] A. Desolneux, L. Moisan, and J.M. Morel. *From Gestalt Theory to Image Analysis, a Probabilistic Approach*, volume 34 of *Interdisciplinary Applied Mathematics*. Springer, 2008.
- [9] W.D. Ellis, editor. *A Source Book of Gestalt Psychology*. Humanities Press, 1967 (originally 1938).
- [10] J. Feldman. Regularity-based perceptual grouping. *Computational Intelligence*, 13(4):582–623, 1997.
- [11] J. Feldman and M. Singh. Information along contours and object boundaries. *Psychological Review*, 112(1):243, 2005.
- [12] D.J. Field, A. Hayes, and R.F. Hess. Contour integration by the human visual system: Evidence for a local “association field”. *Vision Research*, 33(2):173–193, 1993.
- [13] J.M. Fulvio, M. Singh, and L.T. Maloney. Precision and consistency of contour interpolation. *Vision research*, 48(6):831–849, 2008.
- [14] D. Geiger, K. Kumaran, and L. Parida. Visual organization for figure/ground separation. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR ’96)*, page 155, 1996.
- [15] J.J. Gibson. *The Senses Considered as Perceptual Systems*. Houghton Mifflin Company, 1966.
- [16] R. Grompone von Gioi, J. Jakubowicz, J.M. Morel, and G. Randall. On straight line segment detection. *Journal of Mathematical Imaging and Vision*, 32(3):313–347, November 2008.
- [17] R. Grompone von Gioi, J. Jakubowicz, J.M. Morel, and G. Randall. LSD: A fast Line Segment Detector with a false detection control. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(4):722–732, April 2010.
- [18] Rafael Grompone von Gioi. *Inverse Geometry: Graphical Interpretation of Images*. PhD thesis, CMLA, ENS Cachan, 2010.
- [19] S. Grossberg and E. Mingolla. Neural dynamics of form perception: Boundaries completion, illusory figures, and neon color spreading. *Psychological Review*, 92(2):173–211, 1985.
- [20] S. Grossberg and E. Mingolla. Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations. *Attention, Perception, & Psychophysics*, 38(2):141–171, 1985.



- [21] S. Grossberg, E. Mingolla, and W.D. Ross. Visual brain and visual perception: How does the cortex do perceptual grouping? *Trends in Neurosciences*, 20(3):106–111, 1997.
- [22] G. Guy and G. Medioni. Perceptual grouping using global saliency-enhancing operators. In *Pattern Recognition, 1992. Vol. I. Conference A: Computer Vision and Applications, Proceedings., 11th IAPR International Conference on*, pages 99–103. IEEE, 1992.
- [23] Feng Han and Song-Chun Zhu. Bottom-up/top-down image parsing with attribute grammar. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(1):59–73, January 2009.
- [24] B. Jenkins. Component processes in the perception of bilaterally symmetric dot textures. *Attention, Perception, & Psychophysics*, 34(5):433–440, 1983.
- [25] Gaetano Kanizsa. *Grammatica del vedere*. Società editrice il Mulino, 1980.
- [26] Gaetano Kanizsa. *Vedere e pensare*. Società editrice il Mulino, 1991.
- [27] P.J. Kellman and T.F. Shipley. A theory of visual interpolation in object perception. *Cognitive Psychology*, 23(2):141–221, 1991.
- [28] W. Köhler. *Gestalt Psychology*. Liveright, 1947.
- [29] M. Kubovy, A.O. Holcombe, and J. Wagemans. On the lawfulness of grouping by proximity. *Cognitive Psychology*, 35(1):71–98, 1998.
- [30] Y.G. Leclerc. Constructing simple stable descriptions for image partitioning. *International journal of computer vision*, 3(1):73–102, 1989.
- [31] D. Lowe. *Perceptual Organization and Visual Recognition*. Kluwer Academic Publishers, 1985.
- [32] D. Marr. *Vision*. Freeman and co., 1982.
- [33] H.R. Maturana and F.J. Varela. *The tree of knowledge: The biological roots of human understanding*. Shambala, revised edition, 1992.
- [34] J. McDermott and E.H. Adelson. The geometry of the occluding contour and its effect on motion interpretation. *Journal of Vision*, 4(10):9, 2004.
- [35] W. Metzger. *Gesetze des Sehens*. Verlag Waldemar Kramer, Frankfurt am Main, third edition, 1975.
- [36] W. Metzger. *Laws of Seeing*. The MIT Press, 2006 (originally 1936). English translation of the first edition of [35].
- [37] P. Monasse and F. Guichard. Fast computation of a contrast-invariant image representation. *IEEE Transactions on Image Processing*, 9(5):860–872, 2000.
- [38] D. Mumford and A. Desolneux. *Pattern theory: the stochastic analysis of real-world signals*. AK Peters, 2010.
- [39] J.K. O’Regan and A. Noë. A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24:939–1031, 2001.

- [40] H.K. Pao, D. Geiger, and N. Rubin. Measuring convexity for figure/ground separation. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 2, pages 948–955. IEEE, 1999.
- [41] Z. Pizlo, M. Salach-Golyska, and A. Rosenfeld. Curve detection in a noisy image. *Vision Research*, 37(9):1217–1241, 1997.
- [42] X. Ren, C. Fowlkes, and J. Malik. Figure/ground assignment in natural images. *Computer Vision–ECCV 2006*, pages 614–627, 2006.
- [43] J. Rissanen. *Stochastic complexity in statistical inquiry theory*. World Scientific Publishing Co., Inc., 1989.
- [44] N. Rubin. The role of junctions in surface completion and contour matching. *Perception*, 30(3):339–366, 2001.
- [45] Hans Rupp. Über optische Analyse. *Psychologische Forschung*, 4(1):262–300, 1923.
- [46] Sudeep Sarkar and Kim L. Boyer. Perceptual organization in computer vision: A review and a proposal for a classificatory structure. *IEEE Transactions on Systems, Man, and Cybernetics*, 23(2):382–399, 1993.
- [47] Alessandro Sarti, Ravi Malladi, and James A. Sethian. Subjective surfaces: A geometric model for boundary completion. *International Journal of Computer Vision*, 46(3):201–221, 2002.
- [48] A. Sha’ashua and S. Ullman. Structural saliency: The detection of globally salient structures using a locally connected network. In *Proceedings of the Second International Conference on Computer Vision*, pages 321–327, 1988.
- [49] M. Singh and D.D. Hoffman. Completing visual contours: The relationship between relatability and minimizing inflections. *Attention, Perception, & Psychophysics*, 61(5):943–951, 1999.
- [50] Yi-Zhe Song and Peter Hall. Stable image descriptions using gestalt principles. In George Bebis, Richard Boyle, Bahram Parvin, Darko Koracin, Paolo Remagnino, Fatih Porikli, Jrg Peters, James Klosowski, Laura Arns, Yu Chun, Theresa-Marie Rhyne, and Laura Monroe, editors, *Advances in Visual Computing*, volume 5358, pages 318–327. Springer Berlin / Heidelberg, 2008.
- [51] A. Treisman. The binding problem. *Current opinion in neurobiology*, 6(2):171–178, 1996.
- [52] W.R. Uttal. The effect of deviations from linearity on the detection of dotted line patterns. *Vision Research*, 13(11):2155–2163, 1973.
- [53] W.R. Uttal. *An autocorrelation theory of form detection*. Lawrence Erlbaum Associates, Publishers, 1975.
- [54] J. Wagemans. Skewed symmetry: A nonaccidental property used to perceive visual forms. *Journal of Experimental Psychology: Human Perception and Performance*, 19(2):364, 1993.
- [55] J. Wagemans. Detection of visual symmetries. *Spatial vision*, 9(1):9–32, 1995.



- [56] H. Wallach. Ueber visuell whargenommene bewegungrichtung. In *Psychologische Forschung*, 1935.
- [57] M. Wertheimer. Untersuchungen zur Lehre von der Gestalt. II. *Psychologische Forschung*, 4(1):301–350, 1923. An abridged translation to English is included in [9].
- [58] Andrew P. Witkin and Jay M. Tenenbaum. On the role of structure in vision. In Jacob Beck, Barbara Hope, and Azriel Rosenfeld, editors, *Human and Machine Vision*, pages 481–543. Academic Press, 1983.
- [59] Andrew P. Witkin and Jay M. Tenenbaum. What is perceptual organization for? *IJCAI-83*, 2, 1983.
- [60] D. Wouterlood and F. Boselie. A good-continuation model of some occlusion phenomena. *Psychological Research*, 54:267–277, 1992.
- [61] S. Wuerger, R. Shapley, and N. Rubin. On the visually perceived direction of motion by hans wallach: 60 years later. *Perception*, 25:1317–1368, 1996.
- [62] S.C. Zhu and A. Yuille. Region competition: Unifying snakes, region growing, and bayes/mdl for multiband image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 18(9):884–900, 1996.
- [63] Song-Chun Zhu. Embedding Gestalt laws in Markov random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(11):1170–1187, November 1999.
- [64] Song-Chun Zhu and David Mumford. A stochastic grammar of images. *Foundations and Trends in Computer Graphics and Vision*, 2(4):259–362, January 2006.