University of Groningen

Algorithm that mimics human perceptual grouping of dot patterns

Papari, G.; Petkov, N.

Published in:
BRAIN, VISION, AND ARTIFICIAL INTELLIGENCE, PROCEEDINGS

Publication date:
2005

Algorithm That Mimics Human Perceptual Grouping of Dot Patterns

G. Papari and N. Petkov

Institute of Mathematics and Computing Science,
University of Groningen,
P.O.Box 800, 9700 AV Groningen, The Netherlands

Abstract. We propose an algorithm that groups points similarly to how human observers do. It is simple, totally unsupervised and able to find clusters of complex and not necessarily convex shape. Groups are identified as the connected components of a Reduced Delaunay Graph (RDG) that we define in this paper. Our method can be seen as an algorithmic equivalent of the gestalt law of perceptual grouping according to proximity. We introduce a measure of dissimilarity between two different groupings of a point set and use this measure to compare our algorithm with human visual perception and the k-means clustering algorithm. Our algorithm mimics human perceptual grouping and outperforms the k-means algorithm in all cases that we studied. We also sketch a potential application in the segmentation of structural textures.

1 Introduction

One of the remarkable properties of the human visual system is its ability to group together bits and pieces of visual information in order to recognize objects in complex scenes. In psychology, there is a long tradition of research devoted to perceptual grouping. This tradition has been largely formed and is still very much influenced by gestalt laws that name factors, such as proximity or similarity in orientation, colour, shape, or speed and direction of movement which play a role in perceptual grouping (see e.g. [1-4]). The sole naming of these factors, however, is not sufficient for a quantitative analysis of an image aimed at grouping features together automatically, i.e. without the involvement of a human observer. Such automatic grouping is essential for the computerized recognition of objects in digital images.

In the current paper we describe an algorithm for the grouping of points (Section 2) and demonstrate that this algorithm delivers results that are in agreement with human visual perception (Section 3). To quantify the degree of agreement we introduce a measure of dissimilarity between two possible groupings of the points of one set. We also compare the performance of our algorithm in mimicking human visual perception with the performance of the k-means clustering algorithm, commonly used in computer vision (Section 4). Our algorithm is closely related to the gestalt law of proximity but it goes beyond that law in that it has predictive power. In Section 5 we refer to some related previous work, draw conclusions and outline future work on possible applications in computer vision.

2   Grouping Algorithm

2.1   Voronoi Tessellation and Delaunay Graph

Given a set $S = \{p_1, \ldots, p_N\}$ of N points in the plain, we partition the plain in cells $C_1, \ldots, C_N$ (bound or unbound) such that the points which belong to cell C_j, associated with point $p_j \in S$, are closer to p_j than to any other point $p_k \in S$, $k \neq j$:

$$q \in C_j \quad \Leftrightarrow \quad d(q, p_j) \leq d(q, p_k), \quad \forall q, \quad \forall p_j, p_k \in S . \qquad (2.1)$$

Such a partition of the plain is called the *Voronoi Tessellation* (VT) or *Voronoi diagram* related to S. The dual of the Voronoi tessellation, called the *Delaunay Graph* (DG), is obtained by connecting all pairs of points of S whose Voronoi diagram cells share a boundary (Fig. 1). More details about Voronoi tessellation and Delaunay graph can be found in the standard literature (see for example [5]).
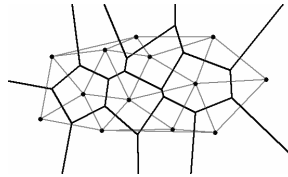

Fig. 1. Voronoi tessellation and Delaunay graph of a set of points

2.2   Reduced Delaunay Graph

Perceptually, a *group* of points in S is a subset of points, which are much closer to each other than to the other points of S. However, the concept of "much closer" is subjective and it is not well defined mathematically. We propose an algorithm which partitions S in disjoint subsets, and show that this partitioning corresponds to human visual perception of groups. We compute the Delaunay Graph *DG* of the set S and eliminate some edges from it to obtain a new graph that we call the Reduced Delaunay Graph; then, we regard the connected components of that graph as groups. This is illustrated by Fig. 2. In order to choose which edges must be removed, for each edge *pq* of the DG we first compute the distance *d(p, q)* (Fig. 2a); then we normalize it with the distances of *p* and *q* to their respective nearest neighbours:

$$\xi(p, q) = \frac{d(p, q)}{\min_{x \in S}\{d(p, x)\}}; \qquad \xi(q, p) = \frac{d(q, p)}{\min_{x \in S}\{d(q, x)\}} . \qquad (2.2)$$

Note that in general $\xi(p, q) \neq \xi(q, p)$. In this way, we assign two ratios, $r_1(e) = \xi(p, q)$ and $r_2(e) = \xi(q, p)$, to each edge e of the DG (Fig.2b).

Next, we reduce the two above mentioned numbers to a single quantity, computed as their geometric average:

$$r(e) = \sqrt{r_1(e) \cdot r_2(e)} . \qquad (2.3)$$

and remove from the DG every edge e for which $r(e)$ is larger than a fixed threshold r_T. We call the remaining graph *Reduced Delaunay Graph* (*RDG*).

More precisely, RDG = (V_{RDG}, E_{RDG}) is a graph whose vertex set V_{RDG} is the same as the vertex set of the DG, and whose edge set E_{RDG} contains those edges $e \in E_{DG}$ for which $r(e)$ is less than or equal to a given threshold vlaue r_T:

$$V_{RDG} = V_{DG}; \qquad E_{RDG} = \left\{ e \in E_{DG} \,\middle|\, r(e) \le r_T \right\}. \tag{2.4}$$

Finally, we regard the connected components of the RDG as *groups*.

This is illustrated in Fig. 2, where the value chosen for r_T is 1.65. The two normalized distances assigned to edge $p_1 p_2$, for instance, are $\xi(p_1, p_2) = 4.95$ and $\xi(p_2, p_1) = 2.48$. Their geometrical average is 3.5, larger than r_T, so $p_1 p_2$ must be removed. The other edges that must be removed are shown in dashed line in Fig.2c. The remaining edges (shown in solid line) produce two connected components, $\{p_1 p_3 p_4\}$ and $\{p_2 p_5 p_6\}$, which we regard as groups.

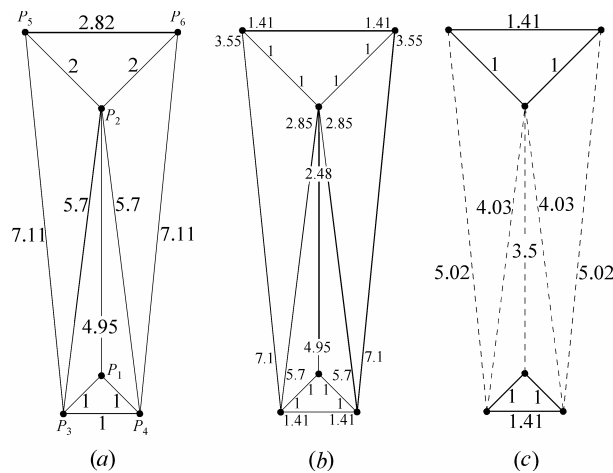**Fig. 2.** (*a*) A set of points and its Delaunay graph. The numbers assigned to the edges are equal to the distances between the corresponding vertices. (*b*) A pair of normalized distances is assigned to each edge of the DG. (*c*) The pair of normalized distances assigned to an edge is replaced by a single number, computed as their geometric mean. The dash lines represent the edges removed from the DG; these are the lines that are assigned numbers larger than the threshold (1.65). The vertices and the remaining edges define the reduced Delaunay graph; its connected components are regarded as groups.

3   Results

In this section we apply our algorithm to several sets of points (see Figures 3-5); in all these examples we take a threshold value $r_T = 1.65$ that is empirically chosen.

In the cases shown in Fig.3*a-b* the RDG is connected and all points are grouped together in one single group; this corresponds to the human visual perception of these

dot patterns; the examples of Fig.3*c-f* are point configurations in which distinct groups, arranged both regularly and randomly, are formed. In the case shown in Fig.4*a* the points arranged in a circle and the points inside that circle form separate groups. Fig.4*b-c* present examples of sets in which there are isolated points. Fig. 4*d-f* illustrate the ability of our algorithm to find clusters of complex shape that are not necessarily convex.

Fig.5a-b show examples of point sets in which groups of points are immersed in a group that covers the whole image. In Fig.5a, for instance, we perceive a set of crosses immersed in a square grid, and it is exactly what the RDG reveals. Similarly, in Fig.5*b* we perceive small squares surrounded by a grid of octagons and big squares. In Fig.5c the RDG reveals the lines traced by the points. This example has a certain relation to the gestalt law of good continuation. The arcs of the RDG not only allow identifying groups but also reveal the order of continuation.

4   Quantitative Comparison to Human Observers

We claim that our algorithm mimics the grouping properties of the human visual system. To quantify this claim, we now compare the grouping results obtained with this algorithm with the groupings defined by human observers. For this purpose we first introduce a measure of dissimilarity between two possible groupings of the points of one set.

4.1   Dissimilarity Between Two Partitions of a Point Set

Let $C_1 = \{U_1, \ldots, U_N\}$ *be* a collection of disjoint subsets of a point set S, such that the union of these subsets is identical with S. Each such subset defines a group of points, and C_1 defines a possible grouping or partitioning of the points of S into different groups. Let $C_2 = \{W_1, \ldots, W_M\}$ be another such grouping of S. We introduce the following quantities:

$$\alpha_{i,k} = \frac{\text{card}(U_i \cap W_k)}{\text{card}(S)}; \quad u_i = \frac{\text{card}(U_i)}{\text{card}(S)}; \quad w_k = \frac{\text{card}(W_k)}{\text{card}(S)}. \tag{4.1}$$

In analogy with probability theory, we call α_{ik} *joint probability* and u_i and w_k marginal probabilities; we also say that C_1 and C_2 are independent if $\alpha_{ik} = u_i w_k$. We also define the following *formal entropies* [6]:

$$H(C_1) = -\sum_i u_i \ln u_i; \quad H(C_2) = -\sum_k w_k \ln w_k; \quad H(C_1, C_2) = -\sum_{i,k} \alpha_{i,k} \ln \alpha_{i,k}. \tag{4.2}$$

We now define the dissimilarity coefficient $\rho(C_1, C_2)$ between the two partitions C_1 and C_2 as follows:

$$\rho(C_1, C_2) = \frac{H(C_1, C_2) - \min\{H(C_1), H(C_2)\}}{\max\{H(C_1), H(C_2)\}}. \tag{4.3}$$
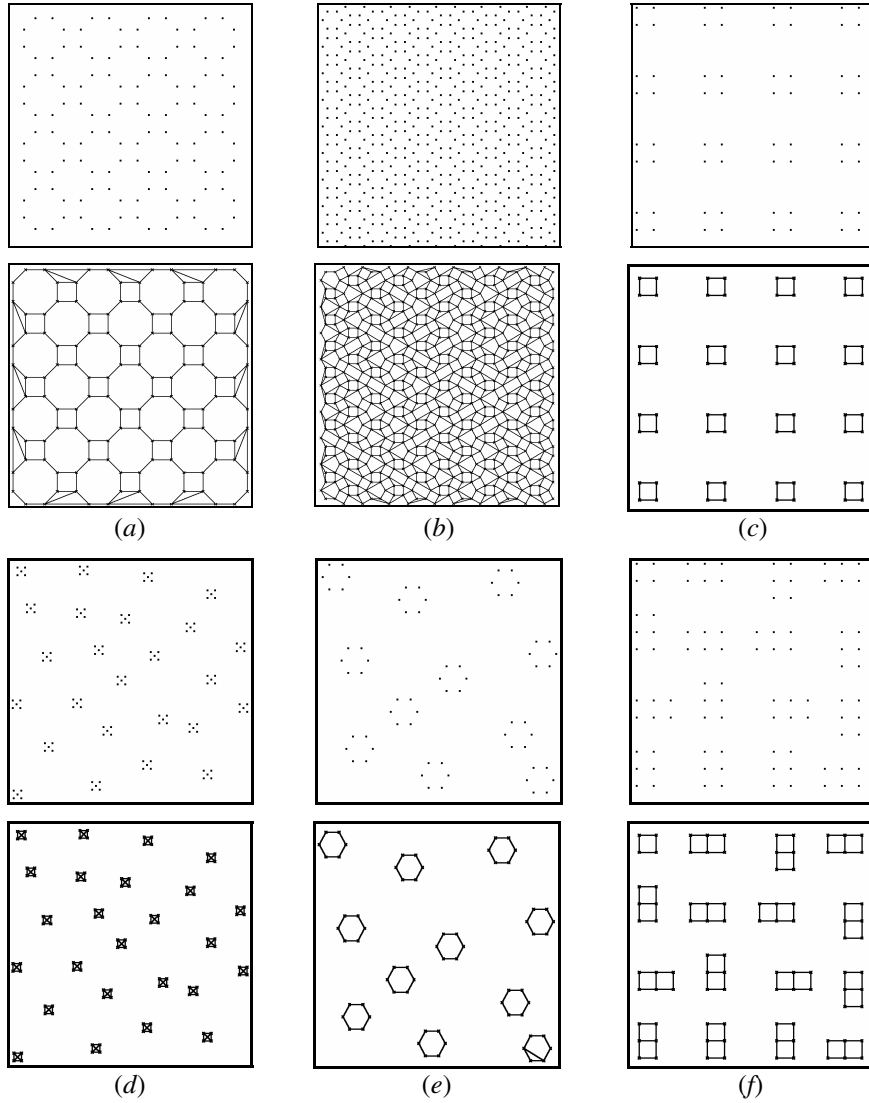
Fig. 3. Six point sets (first and third rows) and their corresponding RDGs (second and fourth rows). In cases (*a-b*) the RDGs are fully connected and all points belong to the same group. In cases (*c-f*) the points are clustered in groups, both regularly and randomly arranged, which are correctly detected by the respective RDGs.
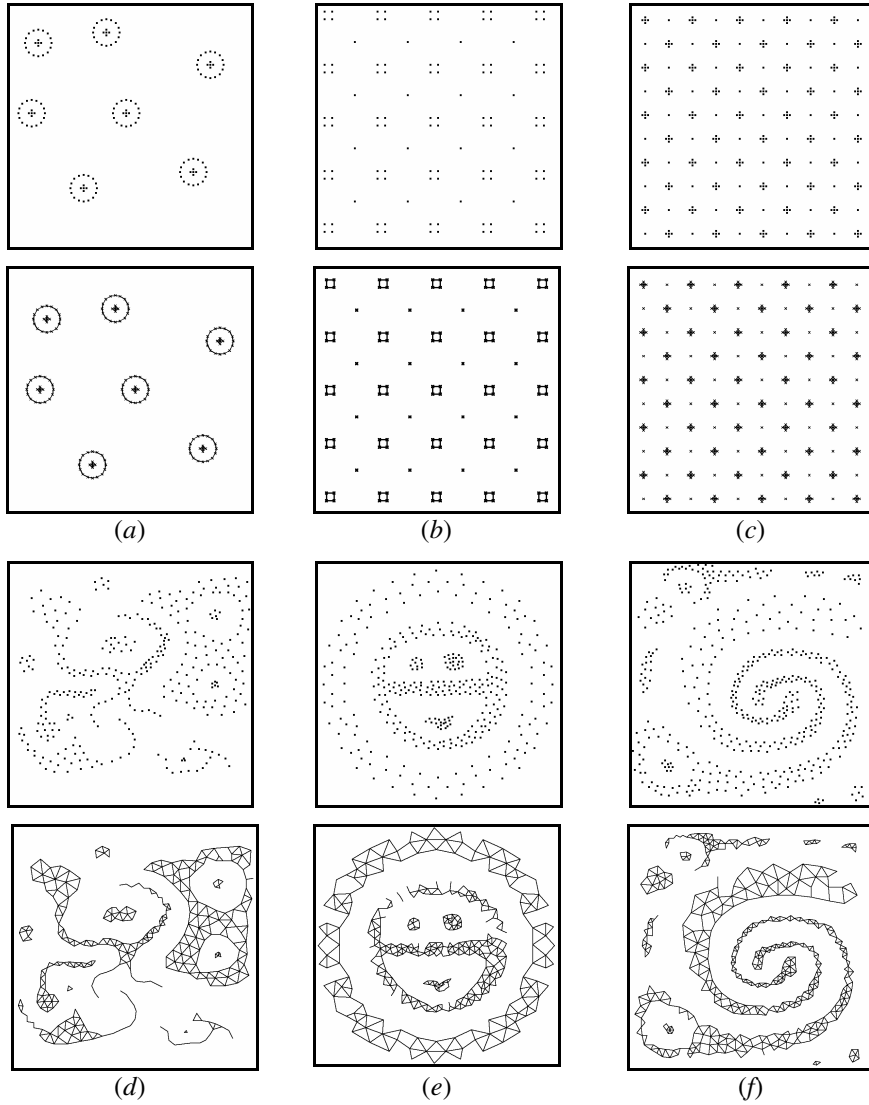
Fig. 4. Other six point sets (first and third rows) and their corresponding RDGs (second and fourth rows). In case (*a*) the points that form a circle belong to a different group than the point inside that circle. Cases (*b-c*) are examples of RDGs that have isolated points; in examples (*d-f*) the shape of the clusters is complex and not necessarily convex.
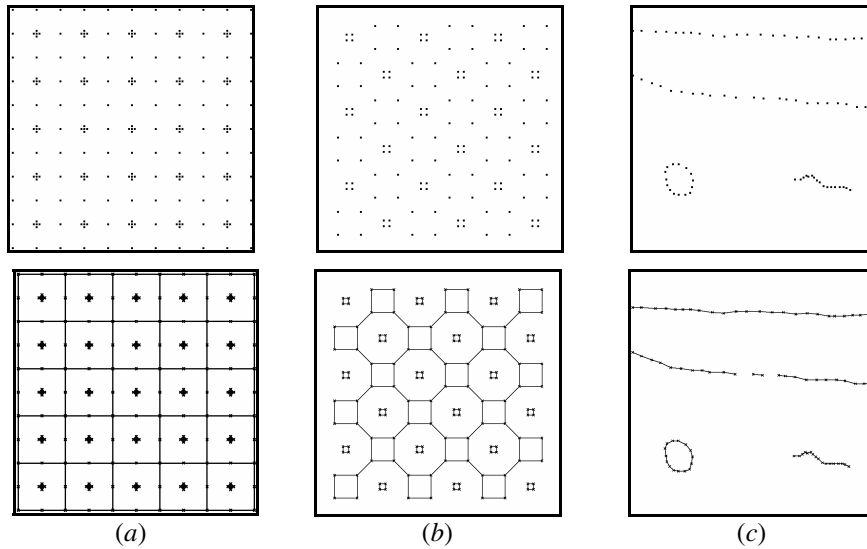
Fig. 5. Point sets (above) and their corresponding RDGs (below). (*a-b*) A system of groups is immersed in a group that covers the whole image. (*c*) A case related to the gestalt principle of good continuation. The edges of the RDG reveal the connectivity of that continuation.

It can easily be shown that ρ is symmetric, positive, equal to zero iff $C_1 = C_2$, and that it satisfies the triangular inequality; therefore, it defines a metrics in the space of the groupings. ρ has also the property to be always between 0 and 1, with $\rho = 1$ if and only if the two concerned groupings are independent. Fig.6 illustrates that the concept of dissimilarity between two groupings is related to their correlation.

4.2 Results

We used the dissimilarity coefficient defined above to compare the results of our grouping algorithm with the perceptual grouping as done by human observers. For each of the images named in the first column of Table 1, we asked eight human observers (male, age varying between 25 and 48) to group the points of the corresponding set. Originally, we used a larger number of point sets but then we excluded sets for which the observers produced different groupings; the set shown in Fig. 4d is one such case that is included here for illustration; the maximum dissimilarity coefficient of two groupings produced by two observers for this set was 0.026. For all other point sets included here the groupings produced by different observers are identical.

In all the cases but the ones presented in Fig.4d and Fig.5c our algorithm produces groupings that are identical with the groupings defined by the human observers. Consequently, the dissimilarity coefficients of the groupings obtained by the algorithm and by perceptual grouping are 0 in these cases, Table 1. One case in which a

difference arises is shown in Fig.7. As can be seen, this difference is small which is well reflected in the small value of the corresponding dissimilarity coefficient. For the case shown in Fig.4d the dissimilarity coefficient is computed as the average of the dissimilarity coefficients between the algorithmic grouping on one hand and each of the human perceptual groupings on the other hand. The average value obtained in this way is smaller than the maximum value of dissimilarity between the perceptual groupings by two different observers.

Table 1 shows also the results achieved by another algorithm - the k-mean clustering [7]. Since this algorithm requires a number of clusters to be specified, this number has been selected to be equal to the number of groups drawn by the human observers. In all cases, the RDG algorithm outperforms the k-means algorithm in its ability to mimic human perception. The performance difference is especially large for point sets in which the groups (as defined by human visual perception) are not convex and have complex form, Fig.7.
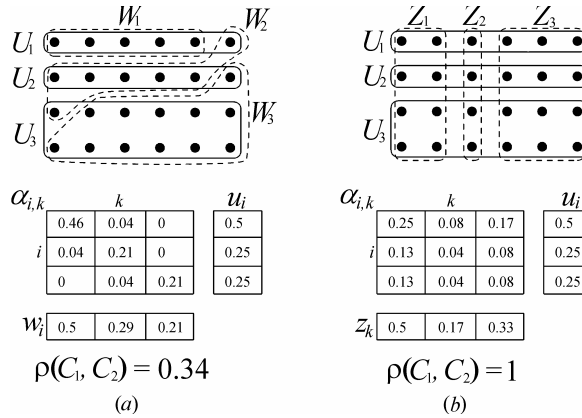

Fig. 6. (*a*) The groupings/partitions $C_1 = \{U_1, U_2, U_3\}$ and $C_2 = \{U_1, U_2, U_3\}$ defined by the continuous and dashed lines, respectively, are strongly correlated and consequently their dissimilarity $\rho(C_1, C_2)$ is small. (*b*) The partitions C_1 and $C_3 = \{U_1, U_2, U_3\}$ are totally uncorrelated and their dissimilarity coefficient is maximal.
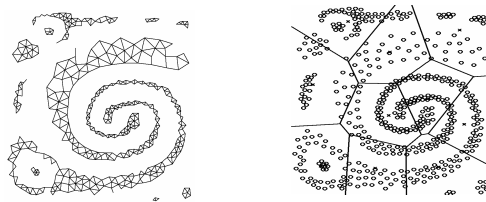

Fig. 7. (left) The RDG grouping algorithm mimics human perception while (right) the k-means algorithm clusters points in a very different way

Table 1. Dissimilarity coefficient values between perceptual grouping by humans and two grouping algorithms, RDG and k-means

Point set	ρ (*RDG* to human)	ρ (k-mean to human)
Fig. 3*c*	0	0.08
Fig. 3*d*	0	0.04
Fig. 3*e*	0	0.05
Fig. 3*f*	0	0.05
Fig. 4*a*	0	0.24
Fig. 4*d*	0.007	0.48
Fig. 4*e*	0.21	0.85
Fig. 4*f*	0	0.68
Fig. 5*c*	0.17	0.30

5 Summary and Conclusion

Attempts to identify perceptually meaningful structure in a set of points have been made in previous works. Voronoi neighbourhood is used in [8]. A discussion in [9] motivates a graph-approach to detect Gestalt groups. In [10] the following concept is introduced: two points p and q of a set S are considered *relatively close* if $d(p,q) <$ max$\{d(p,x),\ d(q,x)\}$, $\forall x \in S$; linking all points of S which are relatively close, we obtain what is called the Relative Neighbourhood Graph [11]. In [11] it is shown by some examples that points are linked in a way that is perceptually natural for human observers. However, this algorithm cannot be used to find groups of points, which is the goal of the current paper. A nice algorithm, presented in [12], repeatedly splits the convex hull of the point set, until clusters are found; but it is unable to find clusters which are one inside the other.

We introduced the concept of a RDG that we use to group points. We demonstrated that this algorithm mimics human visual perception. For this purpose we introduced a quantitative measure of dissimilarity between two possible groupings of the points of a set. The dissimilarities between groupings produced by our algorithm and perceptual groupings done by human observers are zero or very small for a number of point sets that we studied. In contrast, the popular k-means clustering algorithm groups points in a way that is quite different from visual perception. Our algorithm is very simple, totally unsupervised, and is able to find groups of complex shape.

In future work we will apply this algorithm to computer vision problems, such as segmentation of structural textures. Such textures cannot be treated adequately with traditional filter based approaches (see e.g. [13]). Our idea is to first reduce a texture (Fig.8*a*) to one or more sets of points that indicate the positions of structural elements, using morphological filters, and to group them using the RDG approach, Fig.8*b-c*.
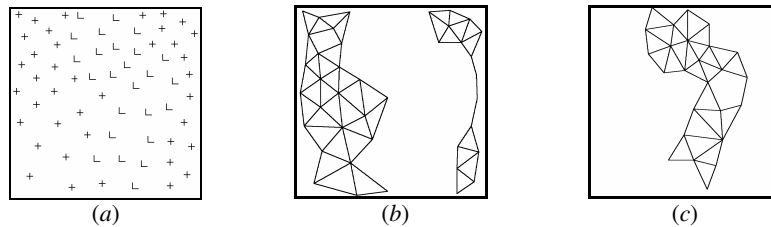
Fig. 8. (*a*) A texture in which the regions are defined by different structural elements. (*b*) The positions of the + elements are identified by a morphological filter selective for the + shape, and the points are connected in a RDG to identify regions of + elements. (*c*) The regions that contain L elements are identified in a similar way using another morphological filter.

References

1. Moore, P., Fitz, C.: Gestalt Theory and Instructional Design. *J. Technical Writing and Communication*, Vol. 23, No.2 (1993) 137-157.
2. Wertheimer, M.: Untersuchungen zur Lehere von der Gestalt II. *Psychologische Forschungen*, Vol. 4 (1923) 301-350.
3. Koffka, K.: Principles of Gestalt Psychology. R. & K. Paul London (1935).
4. Koheler, W.: Gestal Psychology. Mentor Books New York (1947).
5. De Berg, M.: Computational Geometry. Springer (2000).
6. Cover, T. M., Thomas, J. A..: Elements of Information Theory, Wiley series in Telecommunication. Wiley and Son New York (1991).
7. Hartigan, J. A.: Clustering Algorithms. Wiley & Sons New York (1975).
8. Ahuja, N.: Dot pattern processing using Voronoi neighborhoods, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 4, No 3 (1982) 336-343.
9. Zhan, T. C.: Graph-theoretical methods for detecting and describing gestalt clusters, *IEEE Trans. On Computers*, Vol C-20, No 1 (1971) 68-86.
10. Lackford, P. A.: Regionalization, Theory and alternative algorithms. *Geographical Analysis*, Vol. 1 (1969) 196-212.
11. Toussaint, G. T.: The relative neighbourhood graph of a finite planar set, *Pattern Recognition*, Vol. 12 (1980) 261-268.
12. Garai, G., Chaudhuri, B. B.: A Split and merge procedure for polygonal border detection of dot pattern, *Image and Vision Computin,* Vol.17 (1999) 75-82.
13. Grigorescu, S.E., Petkov, N., Kruizinga, P.: Comparison of texture features based on Gabor filters, *IEEE Trans. on Image Proc.*, Vol. 11, No. 10 (2002) 1160-1167.