Robust and Unsupervised Perceptual Grouping of Curves of Dots

José Lezama[1], Gregory Randall[2], Jean-Michel Morel[3], Rafael Grompone von Gioi[3]
[1]Duke University, USA. [2]Universidad de la República, Uruguay. [3]École Normale Supérieure de Cachan, France.

1 Introduction

The Gestalt school of psychology proposed the existence of a short list of grouping laws governing visual perception. Among them, the law of *good continuation* can be stated as “All else being equal, elements that can be seen as smooth continuations of each other tend to be grouped together” [6] (Fig. 2). In the computational domain, attention to the Gestalt laws has been given since the early days of computer vision. D. Lowe was among the first to state the importance of incorporating the Gestalt principles of co-linearity, co-curvilinearity and simplicity for perceptual grouping algorithms [5]. Various computational formalizations of the *good continuation* principle have been proposed ever since, most notably the tensor voting approach [2, 3].

In this work[1], we propose a new model and algorithm for the perceptual grouping by good continuation using a simple model that favors local symmetries, and with a detection control based on the non-accidentalness principle. This allows the method to be general in the sense that it can capture smooth curves of any shape and scale, and is robust to outliers and noise. It is also unsupervised because detections are given by their statistical significance, which requires only a single parameter, namely the number of false detections that would be allowed in an image of random noise.

The proposed algorithm consists of two main steps: building candidate chains of points, and validating them. Candidate chains of points are built by considering triplets of points formed by joining nearest neighbors. Once valid triplets have been obtained, a graph representation is produced where each node corresponds to a triplet. A classic path finding algorithm is run on this graph to obtain paths between all pairs of triplets. Finally, the paths found are validated as non-accidental or rejected using thresholds obtained with the *a contrario* approach [1].

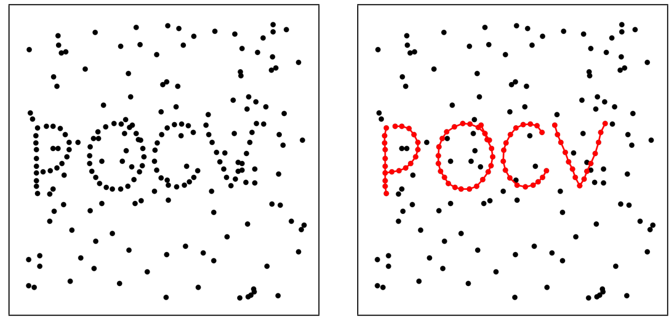


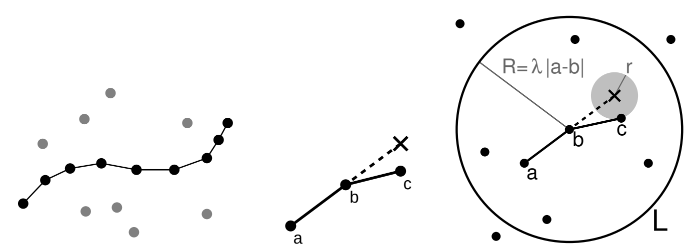
Figure 2: Result of our unsupervised perceptual grouping algorithm. **Left:** 80 points forming the word “POCV” plus 100 random points. **Right:** Detected sets of points in “good continuation”. The method automatically discovers the number of salient structures, and is able to distinguish structure from noise.

2 Mathematical Model

Let us consider a set of N planar points. The aim is to find a mathematical model that can predict when an ordered subset of points lies on a smooth curve that is perceptually salient relative to the background of the other points, Fig. 1(a). Each ordered subset of points (a sequence of points) will be called a *chain*; each set of three consecutive points in a chain will be called a *triplet*. The proposed model is based on the simple idea that the better the symmetry of the triplets, the better the saliency of the sequence.

The evaluation of a chain of points is based on the *non-accidentalness principle*, proposed as the rationale underlying perceptual thresholds. In a nutshell, an observed structure is relevant if it would rarely occur by chance.

[1]A previous version of this work was presented in [4].



(a) (b) (c)

Figure 1: Definition of the *good continuation* event. **(a)** A candidate *chain* is defined by an ordered sequence of points. **(b)** Three consecutive points in a chain define a *triplet*. The ideal symmetric point is represented by X. **(c)** The symmetry precision of a triplet is measured by the distance r from the third point c to the ideal point X, relative to the area of the local observation window R.

Quoting D. Lowe, “we need to determine the probability that each relation in the image could have arisen by accident, $P(a)$. Naturally, the smaller this value is, the more likely the relation is to have a causal interpretation” [5]. The *a contrario* framework [1], a formalization of this principle, is used to provide automatic detection thresholds, compatible with perception, and to handle noise points. Given a random model for the data, the *a contrario* methodology consists in evaluating the expectation of the occurrence of an error as small as the one observed, relative to an ideal structure. If this expectation is small, the event is considered perceptually meaningful.

The probability of observing a chain of points where all of its triplets have a given degree of symmetry is evaluated in a random background model assuming that the points in the image were randomly distributed. The imperfection of a triplet translates into the distance r between the observed third point and its ideal symmetric position, given the local context. To provide scale-invariance, the local context is given by a circular local window L with radius R, where $R = \lambda \cdot |a-b|$ is proportional to the triplet size, Fig. 1(c).

Given that n points were observed in L (not counting the two points defining the triplet), the *a contrario* model assumes that the n points result from a spatial uniform Poisson process in L. Under these assumptions, the error of each triplet is translated into probabilistic terms. Let us call ρ the distance between the ideal point X and its nearest point in L under H_0. The error associated to a triplet is

$$e = \mathbb{P}(\rho \le r) = 1 - \left(1 - \frac{r^2}{R^2}\right)^n. \tag{1}$$

Consider a chain $\mathcal{C}$ of k points $a_1, a_2, \ldots, a_k$. The error e_i of each of the $k-2$ triplets (a_i, a_{i+1}, a_{i+2}) can be evaluated by Eq. (1), and the worst case value, $e_{\max} = \max\{e_1, e_2, \ldots, e_{k-2}\}$, is associated to the whole chain. The probability of all errors being lower than $e_{\max}$ is $\mathbb{P}(E_{\max} \le e_{\max}) = e_{\max}^{k-2}$. Notice that this is not the probability of observing the exact chain $\mathcal{C}$, but the probability of observing, under H_0, chains whose triplets have all error $e_{\max}$ or less relative to ideal symmetric triplets.

The Number of False Alarms (NFA) [1] for a chain of points in *good continuation* is defined as

$$\mathrm{NFA}(\mathcal{C}) = N_{tests} \cdot \mathbb{P}(E_{\max} \le e_{\max}) = bN\sqrt{N} \cdot e_{\max}^{k-2}. \tag{2}$$

The NFA is an upper bound on the expected number of chains with the same error as $\mathcal{C}$ or smaller, to be observed by chance in the *a contrario* model H_0. A large NFA means that such an event is to be expected under the *a contrario* model and therefore is irrelevant. On the other hand, a small

NFA corresponds to a rare event and therefore arguably a meaningful one. The number of tests N_{tests} counts the chains considered as potential good continuations. Given an observed candidate chain of points, the algorithm considers the latter event as an ε-meaningful *good continuation* when the corresponding NFA is lower than $\varepsilon = 1$.

3 Algorithm

For each of the N input points, its b nearest neighbors are explored to form a pair. For each of the $N \cdot b$ pairs, the symmetric point is computed. The two points closest to the symmetric point are used to form candidate triplets.

To find the grouping of triplets into chains, a graph representation of the triplets is constructed where a pair of triplets is considered *adjacent* when they share two points in such a way that they can form a chain of four points. The Floyd-Warshall algorithm is used to find the shortest path joining every pair of triplets. Each path found is a candidate chain that is finally evaluated using the NFA, Eq. (2). Chains with an NFA lower than a meaningfulness threshold $\varepsilon = 1$ are kept as detections.

Once all the *good continuation* events are found, we are interested in keeping only non-redundant detections. Note that a *good continuation* event might mask another smaller event contained in itself (e.g. a subset of the points in a meaningful chain can be also meaningful). We shall say that an event A masks an event B, if $\mathrm{NFA}_A < \mathrm{NFA}_B$ and the chains share at least two points. A non-redundant list of detections is obtained by ordering the detections by NFA and discarding the masked detections.

The algorithm requires two parameters. The number of nearest neighbors b used for exploration, and λ, the ratio of the local window size to a triplet's size[2].

4 Evaluation

Figure 3 presents example results of the algorithm. A first experiment is to verify that under the *a contrario* hypothesis the detector finds no meaningful structure. To this aim, the first two rows show the result of applying the detector to images with randomly distributed points. The second experiment presents normally distributed random points. This experiment suggests that the *a contrario* hypothesis of a local Poisson process is general enough to model points that are unstructured at a local scale.

The last four experiments show figures where curvilinear point structures are present and the algorithm correctly detects them. Note how the algorithm automatically determines the number of structures in each figure, is robust to noise, and handles the different scales, even when changes of scale occur inside a structure.

5 Conclusion

We propose a new model for the perceptual grouping law of good continuation based on local symmetries. Concatenations of triplets are validated as perceptually relevant by considering their expectation of occurrence in a random image. Our method is unsupervised, robust to noise and scale invariant, and it requires no parameter tuning.

[1] A. Desolneux, L. Moisan, and J.M. Morel. *From Gestalt Theory to Image Analysis, a Probabilistic Approach*. Springer, 2008.

[2] D. Gong and G. Medioni. Probabilistic tensor voting for robust perceptual grouping. In *CVPR Workshops*, 2012.

[3] G. Guy and G. Medioni. Perceptual grouping using global saliency-enhancing operators. In *ICPR92*, 1992.

[4] J. Lezama, G. Randall, J-M. Morel, and R. Grompone von Gioi. Good continuation in dot patterns: A quantitative approach based on local symmetry and non-accidentalness. *Vision Research*, 2015.

[5] D. Lowe. *Perceptual Organization and Visual Recognition*. Kluwer Academic Publishers, 1985.

[6] S. E. Palmer. *Vision Science: Photons to Phenomenology*. The MIT Press, 1999.

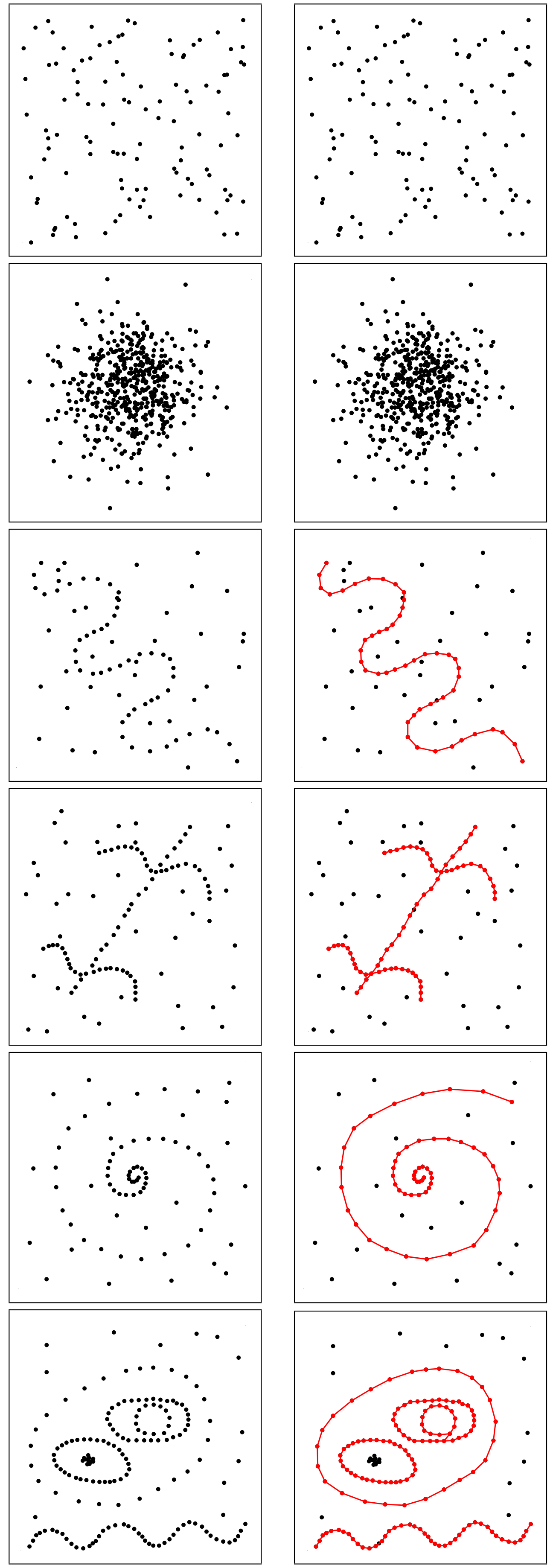


Figure 3: Example results. Left: input. Right: detected perceptually relevant curves in red. Note how the algorithm can detect any number of curves at different scales while producing no false detections in noise.

[2] All the results shown in this abstract use $b = 5$ and $\lambda = 4$.