HAL
open science

Appearance-guided Synthesis of Element Arrangements by Example

Thomas Hurtut, Pierre-Edouard Landes, Joëlle Thollot, Yann Gousseau, Rémy Drouilhet, Jean-François Coeurjolly

To cite this version:

Thomas Hurtut, Pierre-Edouard Landes, Joëlle Thollot, Yann Gousseau, Rémy Drouilhet, et al.. Appearance-guided Synthesis of Element Arrangements by Example. NPAR: 7th International Symposium on Non-Photorealistic Animation and Rendering, ACM, Aug 2009, New Orleans, United States. pp.51-60. inria-00391649v1

HAL Id: inria-00391649

https://inria.hal.science/inria-00391649v1

Submitted on 4 Jun 2009 (v1), last revised 29 Nov 2011 (v2)

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.

Appearance-guided Synthesis of Element Arrangements by Example

T. Hurtut, P.-E. Landes, J. Thollot
INRIA* - LJK† CNRS

Y. Gousseau
Telecom ParisTech - LTCI‡ CNRS

R. Drouillhet, J.-F. Coeurjolly
LJK† CNRS - SAGAG

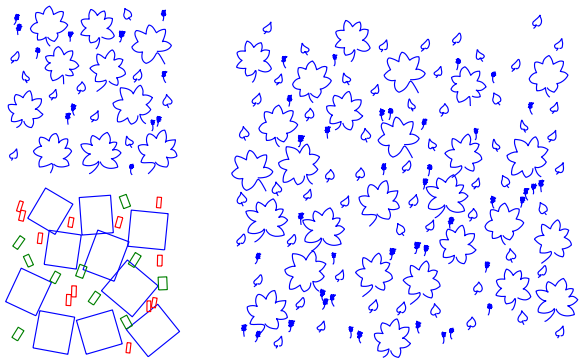


Figure 1: *Given a reference arrangement composed of vector elements* (top left)*, our analysis scheme divides the raw element set into appearance categories* (bottom left)*. Spatial interactions based on appearance can be learned by statistical modeling and exploited to yield visually similar arrangements* (right)*.*

Abstract

We present a technique for the analysis and re-synthesis of 2D arrangements of stroke-based vector elements. The capture of an artist's style by the sole posterior analysis of his/her achieved drawing poses a formidable challenge. Such by-example techniques could become one of the most intuitive tools for users to alleviate creation process efforts. Here, we propose to tackle this issue from a statistical point of view and take specific care of accounting for information usually overlooked in previous research, namely the elements' very appearance. Composed of curve-like strokes, we describe elements by a concise set of perceptually relevant features. After detecting appearance dominant traits, we can generate new arrangements that respect the captured appearance-related spatial statistics using multitype point processes. Our method faithfully reproduces visually similar arrangements and relies on neither heuristics nor post-processes to ensure statistical correctness.

Keywords: Vector texture synthesis, by-example synthesis, NPR.

1 Introduction

Automated stroke-based rendering systems are common in non photo-realistic rendering (NPR). Successful systems used to generate NPR depictions of 3D scenes or photographs are mainly based on heuristics or hard-coded rendering rules and it is up to the artist to take advantage of them to convey his/her own style. Fewer techniques, on the other hand, attempt to automatically learn it instead. In such approaches, the artist provides the system with an example, typically an eventually partially-finished drawing, which has to be analyzed in a way to grasp part of the user's style. The information extracted by this analysis is then used to automatically synthesize new examples visually similar to the original. Such approaches constitute very intuitive tools for artists to handle cumbersome and repetitive tasks, such as creating filling patterns. The main challenge of these techniques is to identify from a limited input what can be assimilated to style and capture it in a way that allows further synthesis. Moreover, to ensure a satisfactory variety of styles, priors that could restrict the scope of supported examples need to be avoided as much as possible.

We focus here on the synthesis of stroke-based elements arrangements. By arrangements, we mean distributions over the 2D plane of visual primitives that do not obey any placement rules or geometric constraints. In such cases, statistics over distances between elements are of primary importance and greatly characterize the input distribution. Our primary goal is then to faithfully reproduce these statistics in order to generate new resembling arrangements.

We claim that, more than sole spatial considerations, the distributed elements' appearance has to be investigated. More specifically, we believe that reproducing pair-wise occurrences of specific visual cues is mandatory to confer output arrangements the same "feel" as the given example. We therefore propose to model and take into account the elements' appearance in the synthesis. To achieve this, we concentrate on stroke-based elements defined as a set of of path-following strokes. Each stroke is a vector curve allowing us to take advantage of studies in line perception [Julesz 1986] to yield an effective element description.

Note that we assume that already-built elements are provided by the user, not individual strokes. They can either be directly drawn by an artist as a whole (this is the case of the examples shown in the paper), but could also be the result of a stroke clustering pre-process similar to the one proposed by [Barla et al. 2006b].

1.1 Related Work

Our technique exploits the vector elements' appearance to guide the synthesis of new arrangements. Related issues arise in various research fields in Computer Graphics, from raster texture synthesis to NPR stroke-based rendering systems. To provide users with intuitive manipulation handles, we favor example-based approaches over procedural techniques. Since texture synthesis is a rich Computer Science field, our review will focus on example-based methods only, before exploring line appearance encoding.

1.1.1 Pixel-based Texture Synthesis

Raster texture synthesis is very inspiring as it focused on example-based approaches early on, see for instance [Heeger and Bergen 1995]. In that case, elements are mere pixels and many successful techniques consider their colors as the realizations of a hidden Markov Random Field (MRF). Their objective is to simulate further sampling to generate new visually-close textures. Most techniques non-parametrically sample their input and use pixel neighborhood matching as an efficient way to implicitly capture its local behavior [Efros and Leung 1999; Wei and Levoy 2000; Ashikhmin 2001]. However, both the appearance and relative placements of such elements are quite limited. Though extra features can be embedded for improved matching [Wu and Yu 2004; Lefebvre and Hoppe 2006], pixels can only be assigned colors and are to follow the lattice structure imposed by the raster grid.

In our case, we aim at producing new arrangements of richer ele-

*ARTIS project-team
†Grenoble Universities, Laboratoire Jean Kuntzmann, CNRS UMR 5224
‡CNRS UMR 5141

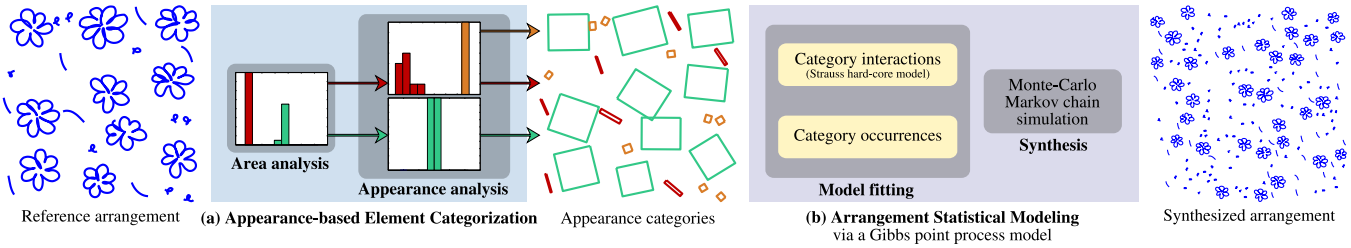


Figure 2: *Overview of our method. First, our* ***appearance-based element categorization*** (a) *examines the reference elements' shape and divides input elements into appearance categories. We perform this analysis by two successive grouping steps: according to the elements' area, and then to their visual appearance. Once the categories are established, we carry on with the* ***arrangement statistical modeling*** *of the user's input* (b). *Its goal is to capture spatial interactions within, and between categories. For that aim, we infer the parameters of a Gibbs point process model from the categorized example. We can then generate new arrangements by establishing new realizations from the fitted model using Monte-Carlo Markov chain sampling.*

ments distributed over the 2D plane without any prior placement.

1.1.2 Patch-based Texture Synthesis

Motivated by the need to capture visual structures lost by the independent process of pixels, texture synthesis involving wider elements, namely pixel patches, were proposed. Few of them however take care of explicitly capturing and handling their relative spatial arrangements.

An interesting example is the case of the texture particle representation [Dischler et al. 2002]. The input bitmap texture is decomposed into a set of small blob-like elements, coined *particles*. Their relative placement is captured by the distances between neighboring particles' axis-aligned bounding boxes. Neighborhood relationships are determined by successive morphological dilation operations until contact between elements is established. Re-synthesis is achieved via a seeding procedure that uses non-parametric sampling for additions of new patches.

Earlier work by [Guo et al. 2001] further completes the analysis process as they acquire their elements, introduced as *textons*, by visual learning. They infer the parameters of a texture model defining the input image as the composed realizations of two stochastic *texton processes*. Elements' appearance, density and spatial arrangements are embedded in this unified model whose configuration likelihood is described by a Gibbs distribution. The parameters maximizing it are estimated by gradient ascent, the overall arrangement evolving according to a Markov chain process. Though powerful, their method requires the evaluation of many parameters and the output texton set gets visually relevant only after hundreds of iterations. We still aim at using similar statistical tools since they provide us with an elegant way of enforcing appearance-based statistics over the output. Not only transposing those techniques to vector elements, we also propose faster solutions.

Other techniques, dedicated to near-regular textures, strive to explicitly identify the underlying lattice structure in the input in order to obtain meaningful building elements [Liu et al. 2004a; Liu et al. 2004b]. Regularity between peaks of auto correlation is investigated and *tiles* – minimal set of patches whose periodic repetition defines the texture – are extracted accordingly. Their insight is that the number of possibilities of tiling the 2D Euclidean plane is limited to the finite number of wallpaper groups. However, such approaches are difficult to generalize to non-regular arrangements of vector elements such as those we want to re-synthesize.

1.1.3 Extension to Vector Primitives

Many generative NPR systems use *strokes* as their basic rendering building blocks. Such inputs (stipples, curves or brush strokes) can be handled in vector form. Compared to pixel patches, this representation grants a more subtle description of the elements' content. This enables us to find new approaches extending example-based raster synthesis procedures.

First attempts consist of parametric approaches. [Jodoin et al. 2002] first deal with the synthesis of hatching patterns by modeling 1D hatch sequences with an explicit MRF to reproduce local pair-wise distances between successive elements. The statistical modeling is elegant but difficult to extend to automatic 2D drawing analysis. Similarly, [Barla et al. 2006a] propose a method to synthesize 2D arrangements of both points and lines and enforce specific statistics on element in a corrective step.

As in texture synthesis, efficient non-parametric sampling techniques were devised, like in Barla and co-authors' subsequent work [Barla et al. 2006b]. Their main contribution is to yield an intermediate input representation by building elements out of strokes using proximity and continuation constraints. For re-synthesis, they first generate, for a given density, a 1D or 2D set of seed points. Input elements are then pasted to those locations by local neighborhood matching. The employed neighborhood system is the Delaunay triangulation over the elements' barycenters and additional perceptual measures determine the matching. Though used during element building, appearance attributes do not contribute during the synthesis step and supported distributions are uniform due to the Llyod relaxation performed on the seed points. [Ijiri et al. 2008] propose a similar, more synthesis-oriented method. New arrangements are created incrementally and rule-based heuristics ensure the correctness of the ongoing triangulation. Again, elements' visual attributes do not influence the distribution itself and most of Barla's perceptual matching considerations are gone for the sake of interactivity.

Our inputs are similar to Barla's or Ijiri's with subtle differences though. We directly have already-built elements at our disposal contrary to the former, while our elements are not explicitly labeled contrary to the latter. Our approach is also different from their work as we formalize arrangement analysis and re-synthesis as a statistical learning problem.

1.1.4 Line Appearance Encoding

One of our contributions is to use dominant element appearance traits as soft constraints influencing the synthesized distribution it-

self. All previously mentioned techniques only account for spatial considerations to determine their output arrangements. Even methods whose inputs allow relevant appearance analysis overlook this valuable information.

Element's appearance encoding is thus of primary importance. NPR research in style transfer accounts for that concern, a proper representation of the strokes' visual attributes largely contributing to the transfer success. Freeman's work on line drawing stylization uses an implicit representation of line appearance by using training data sets of lines and finding nearest neighbors in the target style set [Freeman et al. 2003]. The user's line drawing style is captured by the WYSIWYG NPR rendering system by encoding over-sketch as offsets relative to the line base path [Kalnins et al. 2002]. Style is then encoded as an explicit MRF which allows further 1D synthesis. [Hertzmann et al. 2002] extend their analogy framework to polyline stylization by example and match neighborhoods of the curve's points by comparing point positions and tangent magnitudes. Finally, [Brunn et al. 2007] capture line style as the details functions yielded by a wavelet-like decomposition of the lines.

In this paper, we dispose of compound elements composed of several path-following strokes. This representation of our input allows us to propose more elaborated measurements inspired by line perception studies and use those as relevant features for appearance categorization.

1.2 Contributions

Proposing new approaches for both arrangement analysis and synthesis, the contributions of our method are two-fold:

1. We propose a new algorithm to categorize the elements of a given arrangement using perceptually motivated measures.
2. Based on these measures, we use a multitype point process model to perform synthesis. We chose a model adapted to the capture and restitution of appearance statistics evaluated between, and within element categories.

The main advantages of our method are that it does not require any assumption concerning the input arrangement's distribution and that it performs accurate handling of the elements' appearance. We provide a detailed overview of our arrangement synthesis method's work-flow in Figure 2.

2 Appearance-based Element Categorization

The first step of our method aims at categorizing the example's elements according to their appearance. If some elements exhibit a similar appearance thorough the input, we want to recognize them as belonging to a same category. Elements considered as unique in the example will be grouped in an outlier category. Note that this step corresponds to an automated solution for the manual labeling of [Ijiri et al. 2008]. Our result could, therefore, be used as an input for their algorithm.

The reasons behind our appearance-driven element categorization are the following. According to the Gestalt law of similarity grouping, the Human Visual System tends to mentally perform perceptual categorization and build groups from isolated elements. Once those ensembles are established, strong visual interactions can arise. Not only elements can be perceived as interacting with the other members of its group, but interactions can also occur at the group level. This phenomenon is illustrated on Figure 1 where elements are visually split into three main appearance-based categories, namely the elongated, the cross-like, and the smaller strokes. Since all those elements do not overlap and are mixed quasi-regularly, inter-category interactions are considered repulsive here. On the other hand, intra-category interactions are different and could be described as follows. If considered only with respect to the other members of their own category, the elongated and small strokes seem regularly placed. For the cross-like elements, however, this placement rule does not apply. This distinction between inter- and intra-category visual interactions is mandatory to devise a good capture of the arrangement's visual attributes. We propose a method that can account for it.

2.1 Stroke-based Element Description

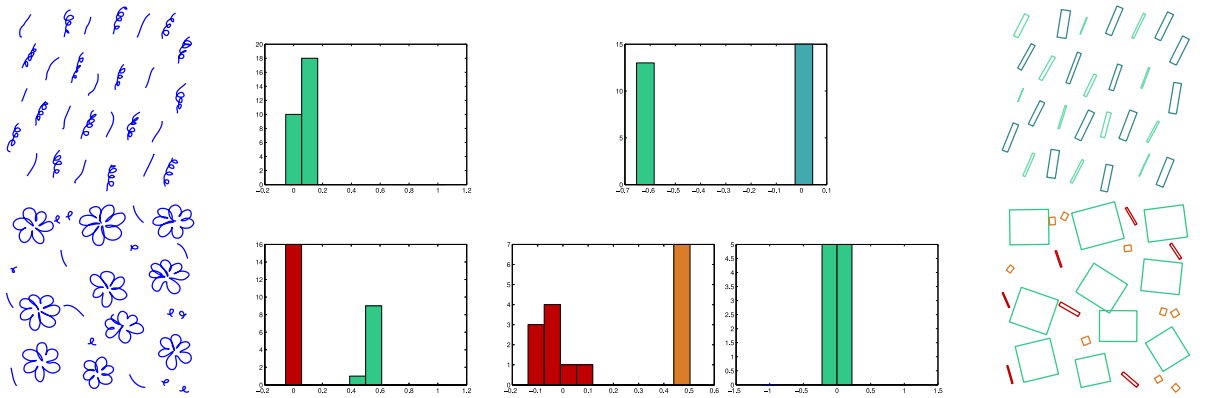
[Julesz 1986] studied human perceptual discrimination of textures composed of stroke-based elements, which he called *textons*. In his theory, discriminative features include the element's principal orientation, as well as its number of crossings and extremities. We use these features as our elements' appearance descriptors. Besides, Julesz's textons all shared the same size. To account for that, we add two features to each element's description: its area and elongation. In practice, elements' orientation, elongation and area are estimated on their fitted bounding box. Crossings and extremities are measured directly on the strokes constituting the element.

Next section tackles the issue of grouping together elements that meaningfully share similar characteristics in this feature space. This brings us to directly compare features that capture drastically different visual characteristics. This question is common to all clustering problems in heterogeneous spaces. Before comparison, features are normalized on $[0, 1]$. Care must be taken that the $[0, 1]$ interval still covers enough visual variation for each characteristic. Orientation is normalized by 2π. Elongation is defined as the ratio of the element's major axis over its minor axis and is normalized by 3. Elements whose elongation before normalization exceeds 3 are tagged as *very elongated* and their associated normalized elongation value is limited to 1. Similarly, elements whose area is larger than 5% of the reference arrangement's area are considered as *large* and attached a normalized surface value of 1. Since we have vector elements at our disposal, we can accurately estimate the curvature of their constitutive strokes. We embed this valuable shape information into our description by counting the number of points of strong curvature along the elements' curves. This feature intuitively corresponds to the number of perceived extremities and is normalized by 10. Lastly, we account for the number of crossings within each element and, as for the extremities, normalize it by 10.

In summary, this gives the following feature set:

Element features	Normalization constant
[Ijiri et al. 2008] Area	5% wrt reference arrangement surface
Principal orientation	2π
Elongation	3
Number of extremities	10
Number of crossings	10

Our description is highly discriminative and focuses on the lines' geometrical shape. Yet notice that the proposed line representation is by no means final, and incorporating other features could be possible. One needs to carefully choose those as the more features are added, the more observations in the input must be provided in order to devise meaningful statistics over a more highly-dimensional feature space. Correlation between descriptive components should also be as low as possible to reduce redundancy. For instance, embedding elements' colors in our descriptor set would thus require special care, such as palette extraction, to avoid the classical "curse of dimensionality" issue.



a) Reference arrangement b) Area histogram c) Appearance histogram d) Appearance categories

Figure 3: Categorization process *Given a reference arrangement* (a), *the modes of the area histogram are first detected* (b). *In this example, one mode is found for the top row, while two are identified for the bottom row corresponding to the small and large elements. For each resulting area category, modes of the appearance histogram are then computed* (c). *Here, two modes are detected for the first row, the most discriminative feature being the number of crossings. The bottom row displays the appearance histogram of small elements where two MM-modes appear (blue and green modes). Only one MM-mode has been detected on the appearance histogram (red mode) for the larger element set. The resulting categories are finally shown with corresponding colors* (d).

2.2 Detection of Meaningful Feature Modes

Our goal here is to categorize elements sharing common visual characteristics once represented in the previously introduced feature space. Our approach is based on two important perception considerations. First, visual perception argues that size is the first information to be perceived for visual recognition tasks. Consequently, we need to bring together elements whose area is roughly the same. Second, as Julesz observed in his studies, it often happens that, depending on the observed elements, not all the descriptive features participate to the perceptual categorization process. Not only useless for categorization, the remaining non-discriminative features also add noise in our sparse feature space which suffers from the usually low number of elements provided by the user. Identifying noise-inducing features is then crucial for ensuring a robust appearance-driven analysis.

Our categorization scheme thus falls into two stages. First, we categorize elements according to their area. Second, for each of the obtained groups considered individually, we perform another categorization step according to the elements' dominant appearance, computed via dimensionality reduction on the remaining four appearance features. Figure 3 illustrates this two-step scheme.

Element grouping according to element area and appearance is established by detecting relevant modes of those two features' density that we approximate by histograms evaluated over the reference arrangement. Because of the lack of prior concerning the number of expected categories, we rely on the *a contrario* method proposed in [Desolneux et al. 2003] as our mode-seeking procedure. We recall this method in the following section to make the paper self-contained.

***A contrario* Methods** *A contrario* approaches have been successfully applied to many Computer Vision problems among which the analysis of histogram modes. The main insight is to rely on a general perception law called the Helmholtz principle which states that an event is perceptually meaningful if it is unexpected. More formally, if the expectation of its occurrences is low under a random assumption.

In the case of histogram analysis, the random assumption is that the descriptor values are i.i.d. uniformly in the L histogram bins $\{b_1, \ldots, b_L\}$. Let us consider an interval noted $S_{i,j} = \{b_i, \ldots, b_j\}$ with $i \leq j$. The prior probability $p_{i,j}$ that an element has its feature descriptor in $S_{i,j}$ is then $p_{i,j} = (b_j - b_i + 1)/(b_L - b_1 + 1)$. Following [Desolneux et al. 2003], we define $S_{i,j}$ as an ε-*meaningful interval* if

$$N\mathcal{B}(p_{i,j}, N_E, k_{i,j}) < \varepsilon \tag{1}$$

where $N = L(L+1)/2$ is the number of possible connected sets of bins; N_E is the number of input elements; $k_{i,j}$ denotes the number of elements in $S_{i,j}$, and $\mathcal{B}$ is the tail of the binomial distribution:

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

The quantity $N\mathcal{B}(p_{i,j}, N_E, k_{i,j})$ can be interpreted as the expectation of the bins from $S_{i,j}$ to occur by pure chance. If this estimate is very low, such bins constitute a meaningful interval. The ε parameter has been shown to cause a logarithmic ε-dependency on meaningfulness, making such approaches robust with respect to their unique parameter [Desolneux et al. 2000]. When set to 1, this leads to the following intuitive interpretation: bins appearing at least once in a random situation are considered as meaningful events.

In order to separate modes inside meaningful intervals, we can similarly define ε-*meaningful gaps* within the distribution histogram as the intervals containing fewer points than the expected average. We say that $S_{i,j}$ is an ε-meaningful gap if

$$N\mathcal{B}(1 - p_{i,j}, N_E, N_E - k_{i,j}) < \varepsilon$$

A *meaningful mode* is defined as a meaningful interval that does not include of meaningful gaps. Lastly, in order to forbid the case of non-disjoint meaningful modes, a meaningful mode is said to be *maximal* if it does not contain, and is not contained in another mode showing greater meaningfulness. Maximal meaningful modes are mentioned as MM-modes in the rest of the paper.

Categorization Algorithm Using *a contrario* histogram mode detection, we then obtain the following categorization algorithm.

First, we compute the MM-modes of the element area histogram estimated over the complete arrangement. This provides us with a preliminary set of categories. Any connected interval of bins which does not belong to an MM-mode is considered as an outlier area category.

Second, for each of these categories (including the possible outlier categories), we perform dimensionality reduction on the four remaining appearance features by Robust Principal Component Analysis [Hubert et al. 2002]. We then identify the MM-modes of these features after their projection onto their first principal component. Each found MM-mode defines an appearance category. Similarly to step one, for each appearance histogram, any connected interval of bins which does not belong to an MM-mode is considered as an outlier appearance category.

In the context of example-based methods, we consider arrangements that have a rather low number of elements, typically below a hundred, and thus a low number of distinct appearance categories. This restricts the precision of the histograms we can analyze. In all our experiments, the distribution of features is estimated on 10 bins, but the discretization scheme can be made more accurate as more input elements are provided by the artist. Likewise, if an area category contains less than 10 elements, we do not split it any further.

3 Statistical Arrangement Modeling

This section presents the statistical process that models the spatial arrangement of categorized elements. Once the parameters of this model are learned on the reference arrangement, the synthesis step consists in running realizations of this model at the desired scale, shape, or density needed by the user.

3.1 Multitype Point Process Model

With the input's appearance categories at hand, we now investigate the elements' relative positions from the perspective of their visual aspect. For that aim, we propose to capture their spatial arrangement via a multitype point process, a statistical model dedicated to the analysis of interactions between a finite set of typed categories. By considering pair-wise element distances as interactions between our established categories, we implicitly grasp the underlying correlation between the elements' appearance and their spatial organization. This model accounts for the interactions gathered over the whole input and supports a wide range of distributions, from stochastic to near-regular.

In our specific case, we assimilate the point data resulting from a realization of this model to the set of the N_E input elements $\mathbf{x} = \{x_1, \ldots, x_{N_E}\}$. Given an element $x_i \in \mathbf{x}$, we associate its corresponding appearance category label m_i to it, m_i being taken from the N_C possible categories labels stored in the label set $\mathcal{M}$. It should be noted that, since $N_C < N_E$, the labels m_i, m_j, may refer to the very same appearance category even though they are related to two distinct elements x_i and x_j.

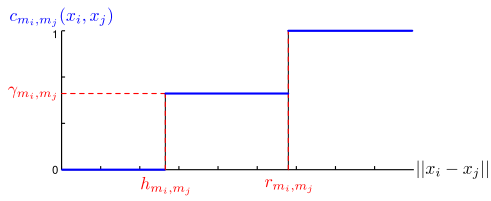
A way to construct a point process model is to write down its probability density function (PDF) with respect to a Completely Random Situation. Such point processes are called *Gibbs point processes* and offer many advantages. Manipulating their PDF to make them account for intricate interactions is easy and further simulation is ensured by well-known Monte-Carlo Markov Chain algorithms. Since we focus here on pair-wise interactions between element categories, we can define our model's PDF, noted $f(\mathbf{x})$, as follows (see [Ripley 1981] for further details):

$$f(\mathbf{x}) \propto \left[\prod_{x_i} d_{m_i}(x_i)\right] \left[\prod_{x_i \neq x_j} c_{m_i,m_j}(x_i, x_j)\right] \tag{2}$$

where $d_m(\cdot)$ is the occurrence probability function of elements from the m category and $c_{m,m'}(\cdot,\cdot)$ is the interaction probability function between the m and m' categories.

A good rule of thumb for statistical modeling is to exploit models whose number of parameters does not exceed the number of observed data. Here, we thus use the simple *Strauss hard-core interaction* which directly relates interaction probability between appearance categories to the Euclidean distance between their elements:

$$c_{m_i,m_j}(x_i, x_j) = \begin{cases} 0 & \text{if } ||x_i - x_j|| < h_{m_i,m_j} \\ \gamma_{m_i,m_j} & \text{if } h_{m_i,m_j} \leq ||x_i - x_j|| < r_{m_i,m_j} \\ 1 & \text{otherwise} \end{cases}$$



The explicit definition of the interaction probability function of a category pair $m, m' \in \mathcal{M}$ then requires the estimation of three constant parameters noted $h_{m,m'}$, $r_{m,m'}$, and $\gamma_{m,m'}$. The first two are distance thresholds, called *hard-core distance* and *trend threshold* respectively. The last one is a scalar in range $[0, 1]$ defining the *interaction strength* and its tuning enables us to model a variety of arrangements from regular to random. Since our proposed interaction functions are symmetric, we just need to evaluate $3N_c(N_c + 1)/2$ interaction parameters to completely define our statistical arrangement model, with N_C being the total number of appearance categories.

3.2 Estimation of the Model Parameters

The multitype Strauss hard-core model is a generic descriptive model that can reproduce various spatial arrangements. This diversity is embedded in the parameters that need to be estimated from the input arrangement by likelihood maximization.

Given the limited set of provided elements, we need to make an important simplifying assumption to ensure a tractable statistical fitting. We suppose the reference arrangement is *stationary* which intuitively comes down to presuming that the artist draws homogeneously over the reference surface. Our re-synthesis still allows the creation of inhomogeneous element distributions. This simplification allows us to treat the categories' occurrence probability functions $d_m(\cdot)$ as constants during the estimation of the parameters. We denote this set of constants Δ.

Moreover, the statistical approach we adopt to estimate the parameters of the multitype point process is hazardous for extremely small categories. In practice, we assume that the user did not draw groups of similar elements containing less than three elements.

3.2.1 Hard-core Distances $h_{m,m'}$ Between Category Pairs

Given a pair of appearance categories $m, m' \in \mathcal{M}$, the hard-core distance obtained by likelihood maximization estimation $h^*_{m,m'}$

corresponds to the minimum distance between pairs of elements picked from the specified categories:

$$h^*_{m,m'} = \min_{\substack{x_i,x_j \\ m_i=m \\ m_j=m'}} ||x_i - x_j||$$

3.2.2 Trend Distances $r_{m,m'}$ Between Category Pairs

Maximizing the trend distances' likelihood estimator is more involving and intuitively corresponds to finding the circular window radius from which the reference arrangement is seen to be the most regular.

To compute that radius value, we use Ripley's $\mathcal{L}$ function which quantifies the deviation of the arrangement, when investigated at a specified scale, relative to a Completely Random Situation [Ripley 1981]. Here follows its formulation :

$$\mathcal{L}_{m,m'}(r) = \sqrt{\frac{\mathcal{K}_{m,m'}(r)}{\pi}} - r$$

$\mathcal{K}_{m,m'}(r)$ is the expected number of elements from the m category lying at a distance r of a randomly picked element of the m' category. As such, it gives an estimate of the element density evaluated at a given scale of a category with respect to another and is normalized in a way that a purely random distribution yields a constant value $\mathcal{L}_{m,m'}(r) = 0$ for all r. Distributions which exhibit more regularity present a negative $\mathcal{L}_{m,m'}$ profile. We thus look for r^*, the first value for which $\mathcal{L}_{m,m'}$ reaches a local minimum. This attests that regularity occurs with maximal amplitude at that scale.

3.2.3 Interaction Strengths $\gamma_{m,m'}$ and Category Occurrence Probabilities d_m

The estimation of the remaining models parameters involves the maximization of the PDF of our model evaluated over the reference arrangement $f(\mathbf{x})$. Finding the optimal parameter sets $\Gamma^* = (\gamma^*_{m,m'})$ and $\Delta^* = (d^*_m)$ comes down to find the best "explanation" by our statistical model of the observed input. However, as Equation (2) suggests, $f(\mathbf{x})$ is defined up to a normalization constant whose explicit evaluation is intractable. To circumvent this problem, we instead maximize the following log pseudo-likelihood involving ratios of f:

$$\sum_{x_i} \log\left(\frac{f(\mathbf{x})}{f(\mathbf{x}/\mathbf{x_i})}\right) - \frac{1}{N_c}\sum_{m=1}^{N_c}\int_{W_\mathcal{R}} \frac{f(\mathbf{x}\cup u_m)}{f(\mathbf{x})}du \qquad (3)$$

where $W_\mathcal{R}$ corresponds to the input drawing window and u_m to an element from the mth appearance category. The involved PDF ratios can be understood as such: given the fixed element distribution $\mathbf{x}$, they quantify the conditional probability of observing an element at a specified location u. The first term of Equation (3) favors locations where observed elements actually lie, while the second term penalizes all the other locations within the drawing window W. The integral is usually estimated using a grid on $W_\mathcal{R}$ where locations u are the centers of each grid cell weighed by its surface. In our experiments, we use a regular grid.

This formula was first proposed by [Besag 1977] and later extended by [Jensen and Møller 1991]. Its suitability to a wide range of Gibbs point processes has been recently proved by [Billiot et al. 2008]. It admits a unique extremum in the (Γ, Δ) parameter space which we find using a Newton-Raphson approach.

3.3 Synthesis by Markov chain Monte-Carlo

As stated in Section 3.1, one noticeable strength of Gibbs point process models is their easy simulation using Markov chain Monte Carlo methods. This interesting property provides us with a convenient means to generate new arrangements that apparently obey the same stochastic process as the provided input. Since we cannot directly sample from $f(\mathbf{x})$, we construct a Markov chain whose set of vertices coincides with the set of elements from the reference arrangement $\mathbf{x}$ and whose equilibrium distribution is to converge to our fitted model's PDF $f(\mathbf{x})$.

We can now compute new realizations of our statistical model over a synthesis window $W_\mathcal{S}$, namely new element arrangements, by using a variant of the Metropolis-Hastings algorithm adapted to point processes [Geyer and Møller 1994].

- randomly initialize output arrangement $\mathbf{x}_0 = \mathbf{x}$ s.t. $f(\mathbf{x}) > 0$
- for time-steps t from 1 to T do:
  - alter current arrangement $\mathbf{x}_t$ using one of the two following equiprobable perturbations:

  Element birth:
    - add an element u at random location in $W_\mathcal{S}$
    - assign random category label m_u to u
    - candidate arrangement $\mathbf{x}' = \mathbf{x}_t \cup \{u\}$
    - compute acceptance rate $R_b = \frac{f(\mathbf{x}')}{f(\mathbf{x}_t)}\frac{n_t}{A_t}$

  Element death:
    - pick random element u from $\mathbf{x}_t$
    - candidate arrangement $\mathbf{x}' = \mathbf{x}_t/\{u\}$
    - compute acceptance rate $R_d = \frac{f(\mathbf{x}_t)}{f(\mathbf{x}')}\frac{A_t}{n_t}$
  - if $R_{b/d} > 1$, then accept perturbation ($\mathbf{x}_{t+1} \leftarrow \mathbf{x}'$)
  else accept perturbation ($\mathbf{x}_{t+1} \leftarrow \mathbf{x}'$) with a probability $R_{b/d}$
  otherwise keep current arrangement unchanged ($\mathbf{x}_{t+1} \leftarrow \mathbf{x}_t$)

Figure 4: *Arrangement synthesis by Metropolis-Hastings sampling*

In the pseudo-code provided Figure 4, we denote $\mathbf{x}_t$ the state of our Markov chain at time-step t. For a specified number T of iterations ($T = 10^5$ in our experiments), we slightly perturb $\mathbf{x}_t$ by introducing or removing one element and obtain a new candidate state for the chain $\mathbf{x}'$. These elementary perturbation events, respectively coined the *birth* or *death*, are effectively taken into account if they satisfy an acceptance rate criterion. Acceptance rates for births and deaths, called R_b and R_d, depend of the ratio of the model's PDF evaluated over $\mathbf{x}_t$ and $\mathbf{x}'$, as well as the current arrangement's area A_t and element number n_t.

The simulation output is a spatial distribution of category labelled elements. We finalize our synthesized arrangement by directly pasting onto each output element's location a reference element randomly picked from the correct appearance category.

4 Results and Discussion

We now present some results and put our technique into perspective with previous methods before discussing its current shortcomings.

4.1 Experimental Results

Examples of categorization and synthesis are shown Figure 5. Those examples attest that the Strauss hard-core process can reproduce various kinds of element distributions, from fairly regular to completely random (e.g., Figures 5-a and 5-e respectively). Thanks to the global multitype optimization procedure, distances between elements are adjusted according to the interactions occurring within

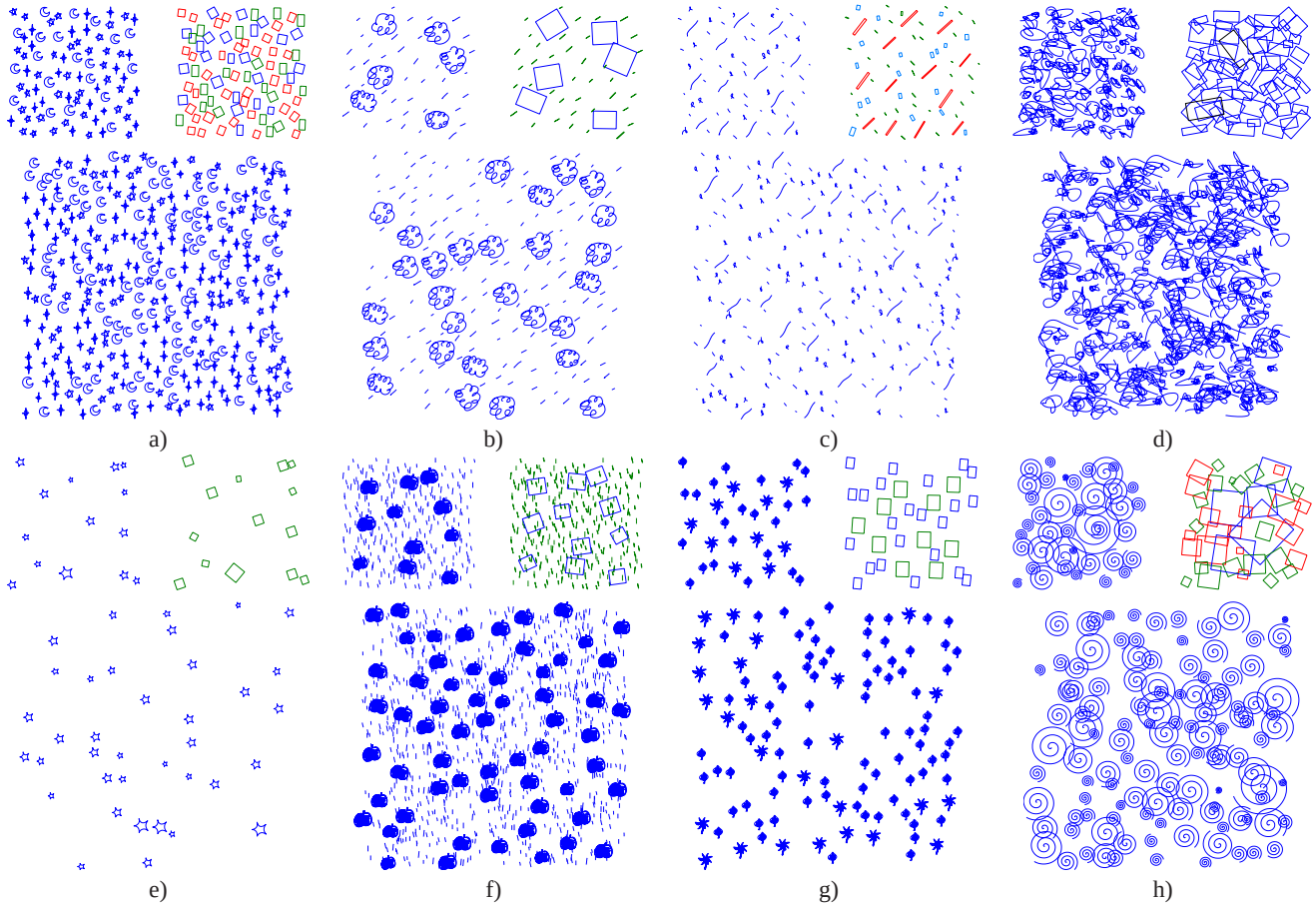


Figure 5: *Synthesis results attesting the variety of spatial distributions that our model can handle. The multitype Strauss hard-core model captures distributions ranging from fairly regular to random,* (a) *and* (e) *respectively. This diversity can be observed inside each category of elements. For instance, apples in* (f) *are regularly distributed, whereas the background is randomly arranged. Similarly, interactions between categories can also vary from repulsive to random,* (g) *and* (h) *respectively.*

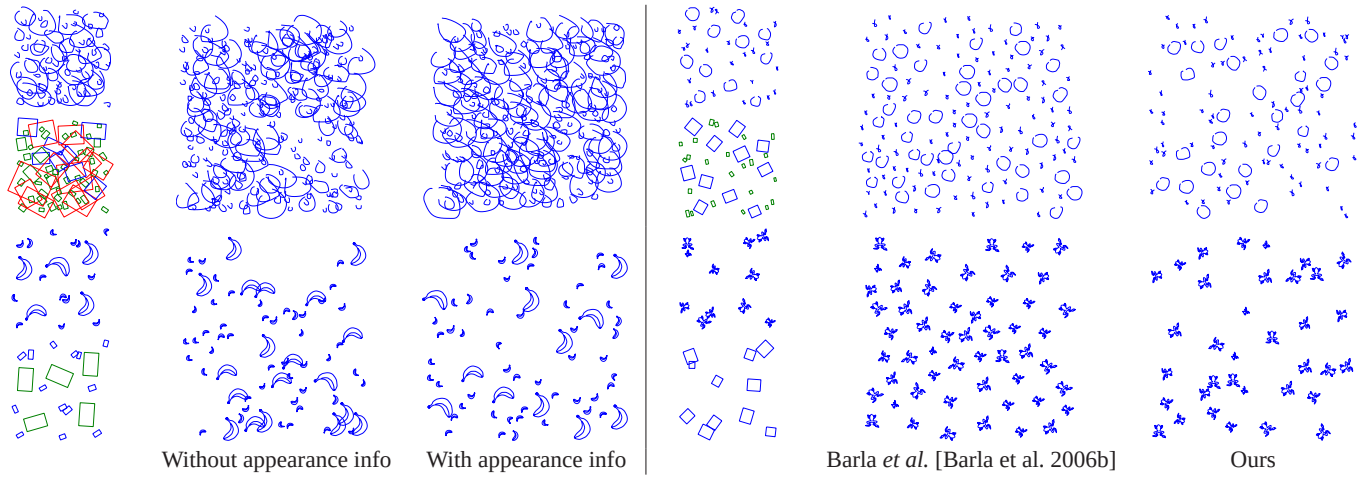


Figure 6: Comparisons (Left) *These examples motivate our appearance-driven categorization. All arrangements are generated by a Strauss interaction model, the ones in the center via a monotype point process overlooking interactions between elements of different appearance. Note the apparition of holes or overlaps which are absent from the input.* **(Right)** *The upper reference is the combination of a more and less regularly distributed sets of elements (the circles and crosses respectively). The approach in [Barla et al. 2006b] enforces regularity over the output while our approach captures the element interactions and preserves them. The lower example consists of a non-uniform arrangement which cannot be represented by the distribution obtained after Lloyd relaxation. A similar phenomenon would arise if using the seeding procedure by [Ijiri et al. 2008], as close seeds are merged below some distance threshold and additional seeds are created in empty regions.*

and between categories. This important property is especially visible on Figure 5-b where the large elements push back the smaller ones beyond the hard-core distance estimated between the two involved categories. Techniques relying on regular element distributions cannot respect such placement constraints, and in this case, element overlap would then occur.

Once our model's parameters have been estimated over the input, new arrangements can easily been synthesized onto various shapes, with possibly different element densities. A gray-level brush-like tool can then be used to intuitively draw arrangements of user-specified densities as illustrated on Figure 7.

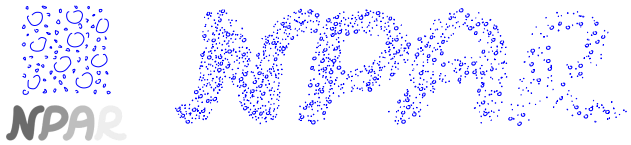


Figure 7: *Example of user-drawn brush strokes whose intensity grants effective control over the synthesized densities of elements.*

In terms of performances, both appearance categorization and statistical fitting are interactive, the bottleneck of the method being the Metropolis-Hastings sampling procedure used for re-synthesis. All examples provided in the paper take about 5 seconds to be categorized and generated on a standard PC. Improvements can be investigated to reduce this computational load. More sophisticated perturbations in the MCMC procedure, such as translation and rotation of elements, could be investigated [Green 1995]. We could also use a spatially discretized grid during sampling, since our application does not meet accuracy requirements as high as the classical uses of such statistical tools.

4.2 Comparisons With Related Work

Improved Handling of Appearance The results on the left side of Figure 6 demonstrate the importance of our appearance analysis step. Arrangements generated by a *monotype* Strauss hard-core model, which considers all elements as visually equivalent, are not fully satisfactory. Even though the overall spatial distribution of the elements' locations is captured, undue holes or overlaps occur and compromise the resemblance of the results with the provided reference.

[Barla et al. 2006b] do account for visual similarity to some extend, by notably defining a perceptual distance used to compare element neighborhoods defined over the Delaunay graph and pick the input element to stitch to a given output location. Their measure, however, only relies on the elements' bounding boxes, whereas our approach integrates more perceptual features enabling the distinction of elements of comparable bounding box. The handling of the elements' appearance by [Ijiri et al. 2008] serves the very same purpose of guiding neighborhood matching. They do not provide any automatic analysis method though, as they require the user to manually label the different elements. Finally, our technique directly correlates statistics between appearance and spatial locations of the elements over the whole input and not just local neighborhoods.

Supported Element Distributions As already presented Figure 5, our method can faithfully reproduce a wide spectrum of element distributions. We notably capture *non-uniform* distributions, which is an improvement over the approaches proposed by Barla *et al.* and Ijiri *et al.*. Re-synthesis results displayed on the right side of Figure 6 attest that fact. Those consist of sets of irregularly distributed elements whose spatial organization cannot be captured by the point distribution resulting from a Llyod relaxation used in [Barla et al. 2006b]. As such, Barla's output arrangements always seem to follow an underlying hexagonal lattice structure. The procedural technique suggested by Ijiri *et al.* also leads to similar results as they devise several growing rules – such as the seed merging and empty space-filling rules – that force the output Delaunay triangulation to be unskewed.

The special case of *strongly regular* distributions is different, as it is not properly supported by our current Strauss model. While this shortcoming also exists in [Barla et al. 2006b], specific care has been provided by [Ijiri et al. 2008] to handle such arrangements, but to the price of user intervention, such as the manual correction of the extracted Delaunay triangulation. In our case, this issue is related to the fact we only consider pair-wise interactions. We believe that considering statistics of interactions involving more than two elements could help us lift that limitation.

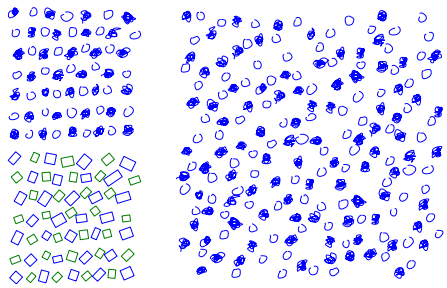


Figure 8: *This example illustrates the main limitation of the Strauss model. Since it only accounts to second order interactions, it cannot reproduce well strongly regular arrangements such as this reference whose elements are pasted onto on a rectangular lattice. More sophisticated models that uses higher order interactions could be investigated to push back this limitation.*

Automation vs. Versatility By locally modifying its element density parameters, our model inherently proposes some intuitive handles for the user to design the synthesized arrangements. For instance, it is straightforward to make our output distributions follow a specified path, typically drawn by the user via a intensity-varying brush tool. This allows the same kind of expressiveness as the *spray* and *boundary tools* in [Ijiri et al. 2008]. However, our local element density control is novel.

Further control on the elements' orientation, such as randomization or flow-guided harmonization, could also be added. Those would take the form of post-processing steps, however, since incorporating too many parameters would endanger the tractability of our statistical fitting. Considering other appearance features, like elements' size or color, could also lead to interesting results and is currently left to future work. Indeed, the focus of the present article is to increase the amount of automatically extractable information from a provided example and capture the interplay between the appearance and the spatial organization of a set of observed elements.

4.3 Limitations and Future Work

A group of perceptually similar elements can sometimes be over-categorized. This effect has yet no consequence on the synthesis step since it leads our model fitting to infer interactions between similar objects stored in different appearance groups. These interactions are reproduced in the synthesized arrangement, but remain unnoticed –as in the reference arrangement– because of the perceptual similarity of the elements (Figure 9-left). Actually, over-categorization on that specific example involves very small ele-

ments and a closer inspection on those tiny shapes does indicate differences between them. Such dissimilarities are hardly visible without explicit zooming though.

Another limitation of our approach is the spatial representation of the elements by their respective centroid. When elements are strongly elongated, this representation is not adapted and is misleading for the model. This is visible in Figure 9-right showing a hatching arrangement. The centroid distribution is well reproduced. Yet, interactions between centroids is not representative of interactions between the actual elements. This drawback could be circumvented by using the Hausdorff distance between bounding boxes instead of the Euclidean distance as the parameter for the interaction functions. Besides, the synthesis of such hatching patterns implies to answer another ill-posed problem. Would the artist expect the system to cover the output window with strokes picked from the example, or by directly adapt their length ?

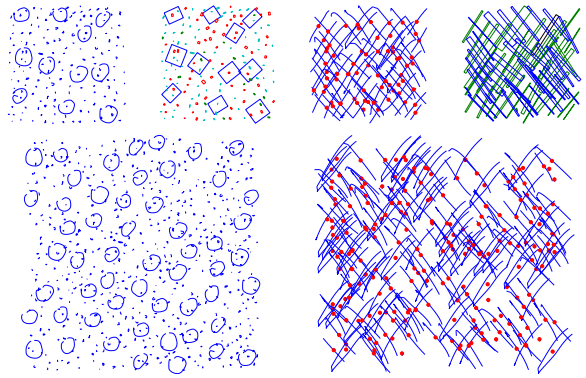


Figure 9: Limitations (Left) *Visually similar elements may sometimes end up in more categories than necessary, such as the small elements displayed here which fall into three distinct categories instead of one* (upper right). *We call that phenomenon* ***over-categorization****. It has close to no impact on the visual quality of our synthesis results, as the system then strives to reproduce unnoticeable interactions between categories containing resembling elements.* **(Right)** *Since elongated elements are not well represented by their sole centroids (red dots), our interaction model based on point-wise distances does not accurately account for the actual interactions between elements. It should noted that the distribution of the centroids is yet preserved.*

5 Conclusions

We have presented an example-based method to synthesize arrangements of vector elements that combines the appearance-guided categorization of the elements and the statistical modelling of the spatial interactions occurring within and between appearance categories. The categorization step is based on several perceptual principles and it could be profitably exploited in other methods such as the procedural approach of [Ijiri et al. 2008]. To the best of our knowledge, our statistical modelling for 2D element arrangements is novel. We believe that multitype point processes –and marked point processes in general– constitute interesting and flexible theoretical tools that could be further investigated.

Acknowledgements: we thank Xavier Descombes and Florence Forbes for their inspiring discussions.

References

ASHIKHMIN, M. 2001. Synthesizing natural textures. *I3D '01*, 217–226.

BARLA, P., BRESLAV, S., MARKOSIAN, L., AND THOLLOT, J. 2006. Interactive hatching and stippling by example. Tech. rep., INRIA.

BARLA, P., BRESLAV, S., THOLLOT, J., SILLION, F., AND MARKOSIAN, L. 2006. Stroke Pattern Analysis and Synthesis. *EUROGRAPHICS '06 25*, 3, 663–671.

BESAG, J. 1977. Some methods of statistical analysis for spatial data. *Bulletin of the International Statistical Institute 47*, 2, 77–92.

BILLIOT, J.-M., COEURJOLLY, J.-F., AND DROUILHET, R. 2008. Maximum pseudolikelihood estimator for exponential family models of marked gibbs point processes. *Electronic Journal of Statistics 2*, 234–264.

BRUNN, M., SOUSA, M. C., AND SAMAVATI, F. F. 2007. Capturing and re-using artistic styles with reverse subdivision-based multiresolution methods. *International Journal of Image and Graphics '07 7*, 4, 593–615.

DESOLNEUX, A., MOISAN, L., AND MOREL, J.-M. 2000. Meaningful alignments. *International Journal of Computer Vision 40*, 7–23.

DESOLNEUX, A., MOISAN, L., AND MORE, J. 2003. A grouping principle and four applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence 25*, 4, 508–513.

DISCHLER, J.-M., MARITAUD, K., LÉVY, B., AND GHAZANFARPOUR, D. 2002. Texture particles. *EUROGRAPHICS '02 21*, 401–410.

EFROS, A. A., AND LEUNG, T. K. 1999. Texture synthesis by non-parametric sampling. *ICCV '99*, 1033–1038.

FREEMAN, W. T., TENENBAUM, J. B., AND PASZTOR, E. 2003. Learning style translation for the lines of a drawing. *ACM Transactions on Graphics 22*, 1, 33–46.

GEYER, C., AND MØLLER, J. 1994. Simulation procedures and likelihood inference for spatial point processes. *Scandinavian Journal of Statistics*, 359–373.

GREEN, P. J. 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika 82*, 4, 711–732.

GUO, G.-E., ZHU, S.-C., AND WU, Y. 2001. Visual learning by integrating descriptive and generative methods. *ICCV '01 1*, 370–377.

HEEGER, D. J., AND BERGEN, J. R. 1995. Pyramid-based texture analysis/synthesis. *SIGGRAPH '95*, 229–238.

HERTZMANN, A., OLIVER, N., CURLESS, B., AND SEITZ, S. M. 2002. Curve analogies. *Eurographics workshop on Rendering '02*, 233–246.

HUBERT, M., ROUSSEEUW, P., AND VERBOVEN, S. 2002. A fast method for robust principal components with applications to chemometrics. *Chemometrics and Intelligent Laboratory Systems 60*, 1-2, 101–111.

IJIRI, T., MECH, R., IGARASHI, T., AND MILLER, G. 2008. An Example-based Procedural System for Element Arrangement. *EUROGRAPHICS '08 27*, 2, 429–436.

JENSEN, J., AND MØLLER, J. 1991. Pseudolikelihood for exponential family models of spatial point processes. *Annals of Applied Probability 1*, 445–461.

JODOIN, P.-M., EPSTEIN, E., GRANGER-PICHÉ, M., AND OSTROMOUKHOV, V. 2002. Hatching by example: a statistical approach. *NPAR '02*, 29–36.

JULESZ, B. 1986. Texton gradients: The texton theory revisited. *Biological Cybernetics 54*, 4, 245–251.

KALNINS, R. D., MARKOSIAN, L., MEIER, B. J., KOWALSKI, M. A., LEE, J. C., DAVIDSON, P. L., WEBB, M., HUGHES, J. F., AND FINKELSTEIN, A. 2002. Wysiwyg npr: Drawing strokes directly on 3d models. *SIGGRAPH '02*, 755–762.

LEFEBVRE, S., AND HOPPE, H. 2006. Appearance-space texture synthesis. *SIGGRAPH '06 25*, 541–548.

LIU, Y., COLLINS, R. T., AND TSIN, Y. 2004. A computational model for periodic pattern perception based on frieze and wallpaper groups. *IEEE Transactions on Pattern Analysis and Machine Intelligence 26*, 354–371.

LIU, Y., LIN, W.-C., AND HAYS, J. 2004. Near-regular texture analysis and manipulation. *SIGGRAPH '04 23*, 368–376.

RIPLEY, B. D. 1981. *Spatial statistics*. Wiley New York.

WEI, L.-Y., AND LEVOY, M. 2000. Fast texture synthesis using tree-structured vector quantization. *SIGGRAPH '00*, 479–488.

WU, Q., AND YU, Y. 2004. Feature matching and deformation for texture synthesis. *SIGGRAPH '04 23*, 364–367.