A Stochastic Tessellation for Modelling and Simulating Colour Aluminium Grain Images[1]

Thomas C. M. Lee

Department of Statistics, The University of Chicago

November 21, 1997

Revised: May 12, 1998

Postal : 5734 University Avenue, Chicago, Illinois 60637

Phone : (773) 702 4837

Fax : (773) 702 9810

Email : tlee@galton.uchicago.edu

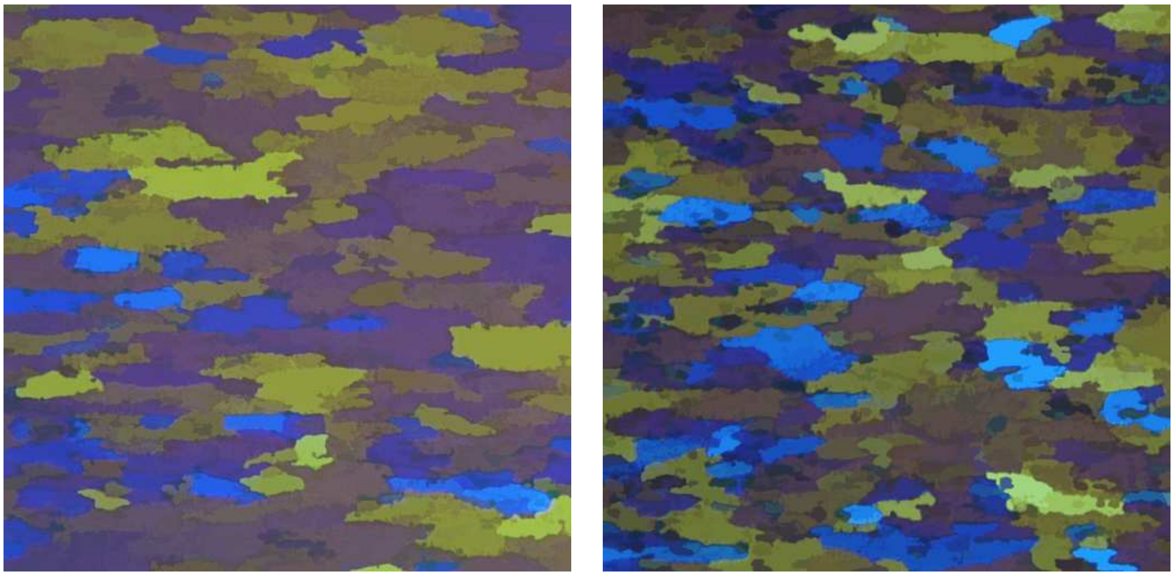
[1]Key words: aluminium grains, colour modelling, Johnson–Mehl modification, noise modelling, roughness, shape modelling, stochastic growth, tessellation, three dimensional modification

A Stochastic Tessellation for Modelling and Simulating Colour Aluminium Grain Images

Thomas C. M. Lee

Summary

The work described in this article is motivated by the need for modelling and simulating some colour aluminium grain images. One possible approach for modelling these colour aluminium grain images is to approximate the images by trivariate piecewise constant functions corrupted by additive and correlated noise. Such a piecewise constant function modelling approach is naturally divided into three parts: shape, colour and noise. The main focus of this article is to present a new tessellation model which can be applied to model the shape characteristics of the aluminium grain images. One special feature of the new tessellation model is that, it is capable of generating tessellation patterns having rough boundaries. Some empirical properties and variants of the tessellation model are also described. Then, as a secondary focus, this article describes models and methods for modelling the observable colour values and noise of the aluminium grain images. Both visual and empirical comparisons between real and simulated images suggest that our modelling is reasonable.



(a) (b)

Figure 1: *Two real aluminium grain images.*

1 Introduction

Most of the material to be presented in this article is motivated by the need for modelling highly complex tessellated images such as those displayed in Fig. 1. These images show grains in rolled aluminium coils, viewed through polarized light. These images are two of a series provided to the CSIRO Division of Mathematics and Statistics by Comalco Aluminium Ltd several years ago. Note the elongated nature of many of the grains. This is due to the fact that the coils have been rolled (i.e., flattened with a roller). As part of its quality control procedures, Comalco was interested in obtaining an understanding of the variability within and between coils of quantities such as grain area, perimeter and orientation.

The first step to extract such quantitative information is to segment the images into their constituent grains. For example, see Fig. 2 for a semi–automatically obtained segmentation of the image displayed in Fig. 1(a). In order to evaluate and compare *quantitatively* the

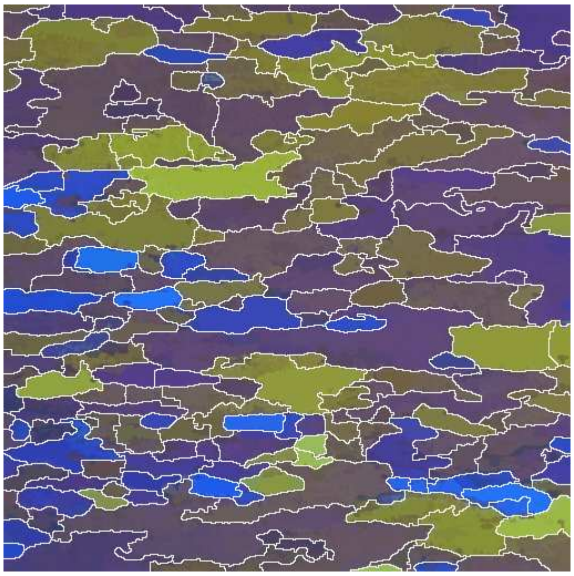


Figure 2: *A semi–automatic segmentation of the image displayed in Fig 1(a).*

performance of different segmentation procedures, it is useful to develop models and methods which can be used to simulate the important characteristics of these aluminium grain images.

A close examination of the images suggests that a colour aluminium grain image can be well approximated by a trivariate two–dimensional (2D) piecewise constant function with noise, possibly correlated, superimposed. That is, grains and grain boundaries can be represented respectively by connected regions and discontinuity points of a trivariate 2D piecewise constant function. Such a piecewise constant function model can be simulated, in a natural manner, by the following three steps:

1. **Shape**: Generate a spatial *tessellation*. Each region of the generated tessellation will be treated as a single aluminium grain, and boundaries of the tessellated regions will be treated as grain boundaries.

2. **Colour**: Given a tessellation, assign a *simulated mean colour value* to each of the generated "empty" regions. Such simulated mean colour values should resemble the observed

mean colour values of the real aluminium grain images.

3. **Noise**: Given a "colour–assigned" tessellation, add suitable (possibly *correlated*) noise to it. Such noise, from a visual sense, should corrupt a noise–free "colour–assigned" tessellation in a similar manner to the noise in the real aluminium grain images.

The main contribution of this article is the proposal of a new tessellation model which can be applied to model the shape characteristics of the aluminium grain images (part of the material has been presented in an earlier conference; see Lee & Cowan 1994). In particular, this new model incorporates quite naturally the rough boundaries that are clearly visible in the aluminium grain images. In addition, we also describe methods for modelling and simulating the colour and noise characteristics of the aluminium grain images.

The rest of this article is organized as follows. Section 2 presents the new tessellation model. One key component of the model is a *stochastic growth mechanism*, and Section 3 compares this growth mechanism with other growth mechanisms appearing in the literature. Some properties of the tessellation model are presented in Section 4, while some possible variants of the tessellation model are described in Section 5. Section 6 demonstrates how the tessellation model can be applied to simulate the shape characteristics of the aluminium grain images. Sections 7 and 8 describe, respectively, methods for modelling the colour and noise characteristics of the aluminium grain images. Then, in Section 9, real and fully simulated images are empirically compared. Finally Section 10 concludes this article.

2 The Proposed Tessellation Model

This section presents our tessellation model. We mention again that the real grain boundaries are visually rough, and that the proposed tessellation model is capable of generating tessellation patterns with rough boundaries.

2.1 Why Existing Tessellation Models are not Appropriate?

Many common tessellation models have been proposed in the literature for the modelling of image data. However, we believe that they are not as appropriate as our proposed model. First, many of the common models, such as the Voronoi, Delaunay and Poisson line tessellations, are inappropriate as their boundaries consist of straight lines (i.e., visually smooth); see, for example, Ahuja & Schachter (1983), Stoyan, Kendall & Mecke (1987), Okabe, Boots & Sugihara (1992), Schwertel & Stamm (1997) and references given therein.

There is one class of models, namely coverage models, which can be applied to generate tessellations with rough boundaries. These tessellation models are also known as "bombing" models (e.g., Ahuja & Schachter 1983), dead–leaves models (e.g., Matheron 1968, Jeulin 1979, Jeulin 1980 and Jeulin 1993) and falling–leaves tessellations (e.g., Cowan & Tsang 1994). The idea is that flat objects are successively dropped onto an image, with those dropped later covering those dropped earlier. Then, when the process finishes, visible object boundaries form a tessellation. Thus a tessellation with rough boundaries can be obtained if the boundaries of those flat objects are rough. However, this object dropping mechanism bears no relation to the physics (see below) of the formation of the real aluminium grains, so we regard it as not as appropriate as our proposed model, which does follow the physics to some extent.

2.2 Seed–and–Grow Strategy

It is known that the formation of those aluminium grains follows some sort of growth process. Therefore it is natural to build our tessellation using a "seed–and–grow" strategy. That is, we start with a number of seeds (or nuclei, or centres) and grow grains outward from the seeds.

Eventually the grains meet each other and those points of contact define grain boundaries, which define a tessellation of the space. The Voronoi tessellation is one example of this: all seeds grow with uniform growth rate in all directions simultaneously, and boundaries are formed when two growing regions meet. As a result of the uniform growth rate, these boundaries are smooth. In order to produce rough boundaries, the proposed tessellation employs a growth mechanism having a *stochastic growth rate.*

Besides producing rough boundaries, another major characteristic of the proposed tessellation is that it is defined on a d-dimensional regular grid (e.g., square, triangular or hexagonal) rather than the traditional $\mathbb{R}^d$, for the purpose of facilitating computer simulation. We first restrict ourselves to the case of $d = 2$ and the grid is square with a 4–connected neighbour system. Extensions to higher dimensions and other regular grids are straightforward; see Section 5.2.

2.3 Initial Pixel Allocation

The tessellation building process is divided into two stages: *initial pixel (or seed) allocation* and *pixel–labelling* (or *seed–growing*). The second pixel–labelling stage can be further divided into two components: *pixel choosing* and *pixel assigning.* In this article a digital version of a two dimensional homogeneous Poisson process is used to perform the initial pixel allocation. We could certainly use other spatial point processes for this initialization, but we will only consider the Poisson case here.

2.4 Pixel–labelling

Suppose n initial pixels are generated (and hence there are n grains in the final realization) and we group these n pixels into n sets: $G_1, G_2, \ldots, G_n$. Thus initially each G_i has exactly one pixel. The pixel labelling mechanism labels one pixel at a time according to the stochastic rule described below, and the whole process stops when all pixels have been labelled.

At any moment during the labelling process, the domain of interest (i.e., the image to be tessellated) is partitioned into three sets L, C and N, and every pixel of the image belongs to exactly one of these three sets. L, C and N represent "**L**abelled", "**C**andidates for labelling" and "**N**ot candidates for labelling" respectively, and they are defined as:

$$\begin{aligned} L &= \bigcup_{i=1}^{n} G_i, \\ C &= \{x \notin L \mid nbr(x) \cap L \neq \emptyset\}, \\ N &= \{x \notin L \cup C\}, \end{aligned}$$

where $nbr(x)$ denotes the set of all 4–connected neighbours of the pixel x. Hence before the first labelling, L is the set of all initial pixels. With this setting, labelling a pixel is equivalent to *choosing* an $x \in C$ and *assigning* it to exactly one of the G_i's. Of course after this pixel assignment we have to remove x from C and perform necessary updates on all related sets.

The stochastic labelling mechanism is as follows. First we further partition C into 4 sets C_1, C_2, C_3 and C_4, and define n_1, n_2, n_3 and n_4 as:

$$\begin{aligned} C_i &= \{x \in C \mid \text{ number of labelled neighbours of } x \text{ is } i\}, \\ n_i &= \text{ number of elements in } C_i, \end{aligned}$$

for $i = 1, \ldots, 4$. We also associate each C_i with a tuning parameter λ_i. Now, for any $x \in C_i$ and conditional on L, we assume

$$P(x \in C_i \text{ is the next pixel for labelling} | L) = \frac{\lambda_i}{\sum_{j=1}^{4} n_j \lambda_j}, \qquad (1)$$

where $P(E)$ is the probability of the event E. Equation (1) merely says that, the probability that an $x \in C_i$ will be chosen as the next pixel for labelling is proportional to λ_i. If $\lambda_i = i$

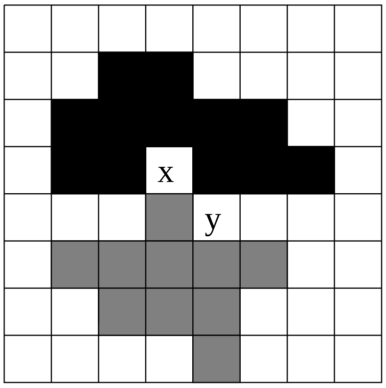


Figure 3: *White squares are unlabelled pixels. Using the random rule, pixels x and y will be labelled black with probabilities 3/4 and 1/3 respectively. Using the majority rule, x will be labelled black while y will be labelled grey.*

for $i = 1, \ldots, 4$, it translates to: the probability that a particular pixel will be chosen as the next labelling pixel is proportional to how many labelled neighbours that this pixel has. Note that Equation (1) is the key equation which drives our stochastic growth mechanism, and that those λ_i's are parameters of our tessellation.

After having chosen the next pixel x, say, for labelling from the C_i's using (1), we have to assign this chosen pixel x to exactly one of the G_i's. If all labelled neighbours of x belong to the same G_i, then x will be assigned to this same G_i. If those labelled neighbours of x come from more than one G_i (in this case x is a boundary pixel), then we could assign x to G_i with probability proportional to the number of neighbours of x already assigned to this G_i. We call this boundary formation rule the "random rule". We could also have a "majority rule": assign x to the set G_i which contains most of those labelled neighbours of x. See Fig. 3 for an illustrative example.

Repeat this pixel choosing and assigning mechanism until all pixels have been labelled (or equivalently, when C and N both become empty). This completes the algorithm but a further step of rescaling the vertical axis of such tessellations with respect to the horizontal axis helps to produce simulated images which appear to be similar to images of rolled aluminium grains. Fig. 4 and Fig. 5 provide realizations of the proposed tessellation and the growth process respectively.

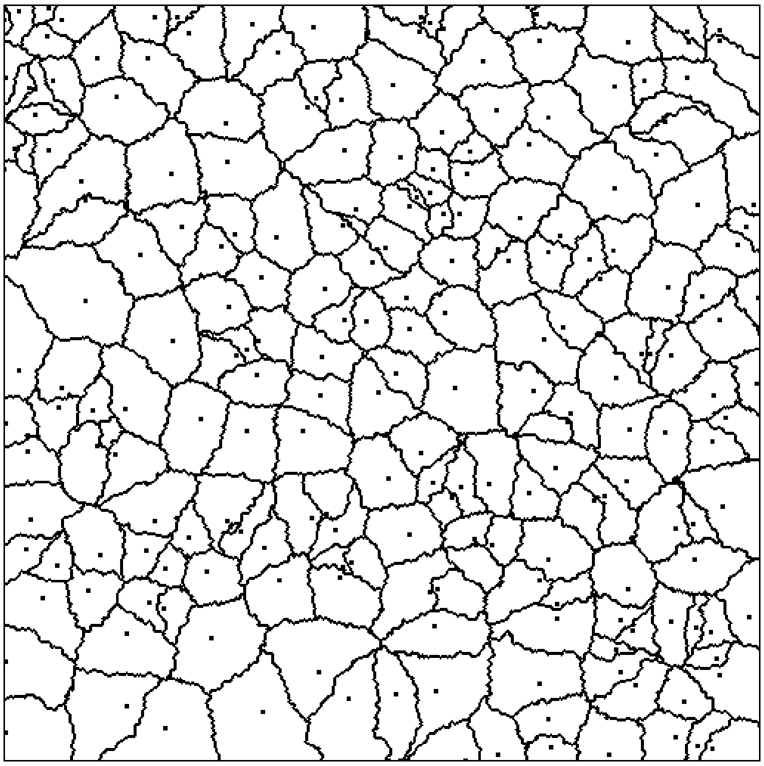
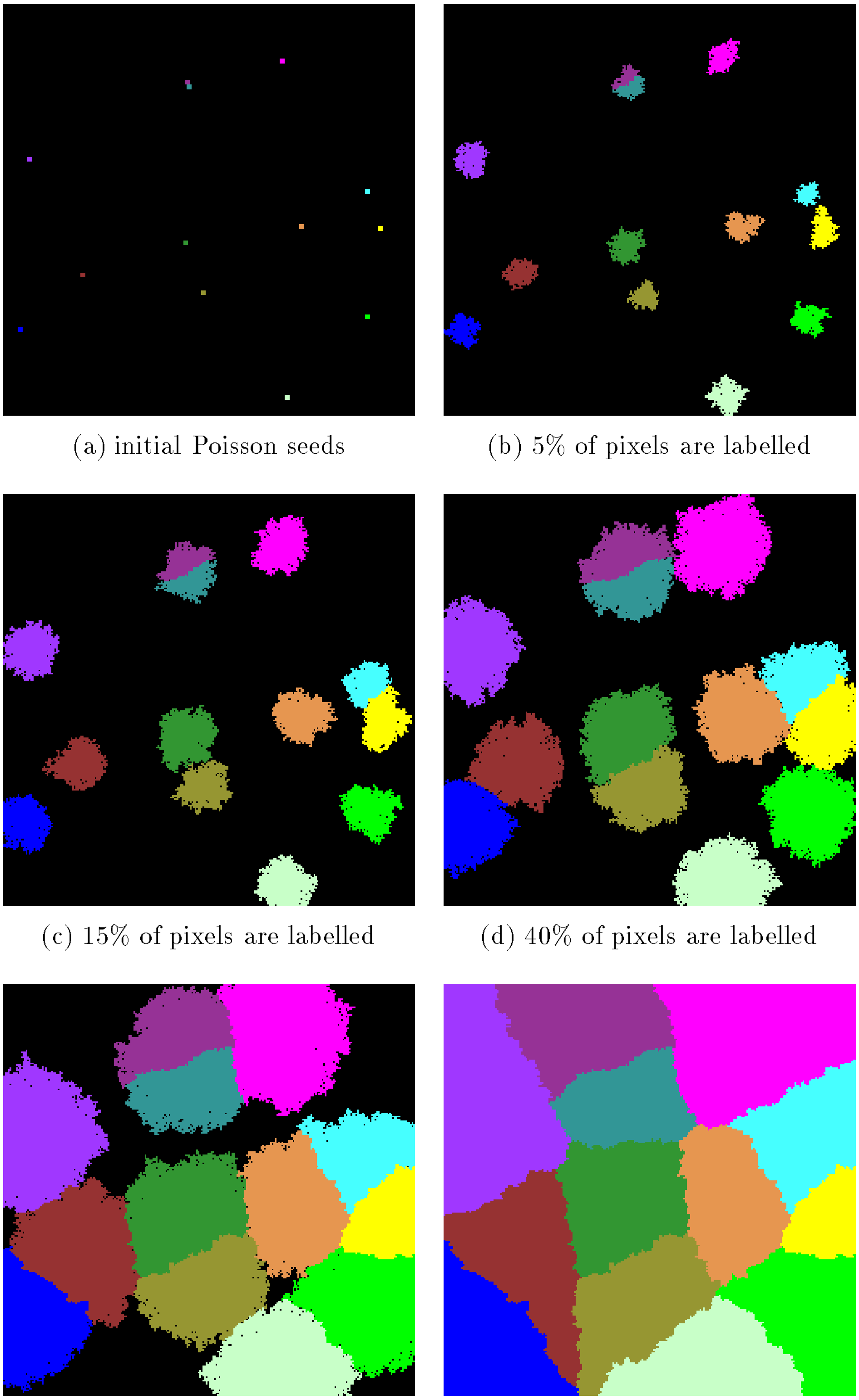


Figure 4: *A realization of the proposed tessellation with Poisson intensity 0.0008,* $r_1 = r_2 = 1$ *(*r_1, r_2 *are defined in Section 4.1) and the random rule for boundary formation. Dots inside grains are initial seeds generated by the Poisson process. Since in our formulation boundaries are composed of horizontal and vertical "edges" between pixels, the actual boundaries cannot be exactly displayed in images. The way we deal with this is to display all pixels which have at least one of their four edges as part of a boundary. Size of image:* 512×512*.*



(a) initial Poisson seeds

(b) 5% of pixels are labelled

(c) 15% of pixels are labelled

(d) 40% of pixels are labelled

(e) 70% of pixels are labelled

(f) final tessellation

Figure 5: *A realization of the proposed growth process.*

We close this section with the following remark about those grains that intersect the image boundaries. The shapes of these grains are "biased" in the sense that, during the growth process, no grains with seeds outside the image are growing towards the centre of the image and interact with these "boundary grains". One simple way to adjust for this is first to obtain a large initial tessellation and then take a middle subset of it as the final pattern. However, throughout this article, we do not perform this adjustment.

3 Comparison with some Growth Models

Eden (1961) proposed a model in order to mimic the growth process of an organism which starts with a single cell (i.e., there is only one initial pixel). Eden's growth process is the same as the growth process described here when all λ_i's are equal.

Williams & Bjerknes (1972) studied the growth of a tumour and suggested a model for describing the phenomenon. This model, sometimes called the contact process, has also been investigated for example by Harris (1974), Durrett (1980) and Griffeath (1981). As in Eden's model, only one initial pixel is considered. In their model, labelled pixels can be "delabelled" and the relative probability of labelling (or delabelling) is governed by a parameter κ. When $\kappa = \infty$ (i.e., labelled pixels cannot be "delabelled"), the probability for a particular unlabelled pixel to be labelled next is proportional to the number of its labelled neighbours. This is the same as the growth process described here when the λ_i's are in the ratio 1, 2, 3 and 4. On the other hand, if κ is very small, an object may actually shrink rather than grow.

Richardson (1973) introduced a model in which more than one unlabelled pixel can be labelled in the same time step. Again, only one initial pixel is considered. In this model, each candidate for labelling (i.e., unlabelled pixels having at least one labelled neighbour) will be labelled in the next time step with probability p. So if there are N candidates for labelling, we expect Np candidates will be labelled in the next time step.

More recently, a growth process called *probabilistic mixing* has been introduced by Thompson & Rosenfeld (1993). This probabilistic mixing growth combines both deterministic and stochastic growth processes. These authors also considered the case when the initial seed is not a single pixel, but is a *skeleton* — a set of connected pixels of width 1.

A somewhat different growth process, the diffusion limited aggregation (DLA) model, is discussed for example by Witten & Sander (1981) and Meakin & Sander (1985). The first step of this DLA model is to place a seed in the centre of the image. Then a second pixel (or particle) is randomly placed in the image and is allowed to "walk" in a random fashion. When this second particle hits the seed, it becomes part of the growing seed (grain). A third pixel is then released to walk randomly until it joins the growing seed, and then a fourth pixel is released and so on. The procedure stops when the number of pixels released hits a pre–set limit, and grains resulted from this growth process are often of "star–shape". Also, this DLA growth process has been extended to the case of multiple seeds; see Brémond & Jeulin (1994*a*) and Brémond & Jeulin (1994*b*).

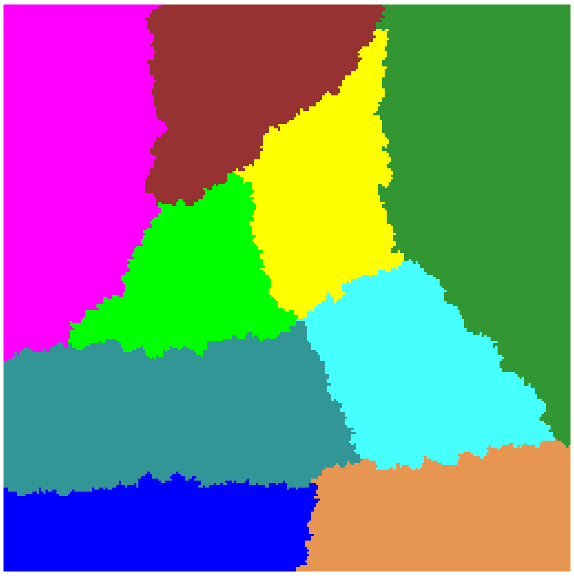
Clearly, there is some overlap between our model and some of those above. However, most of the above models consider a single initial seed, while our proposed growth model is capable of handling the growth of multiple seeds and the formation of region boundaries.

4 Some Properties of the Tessellation

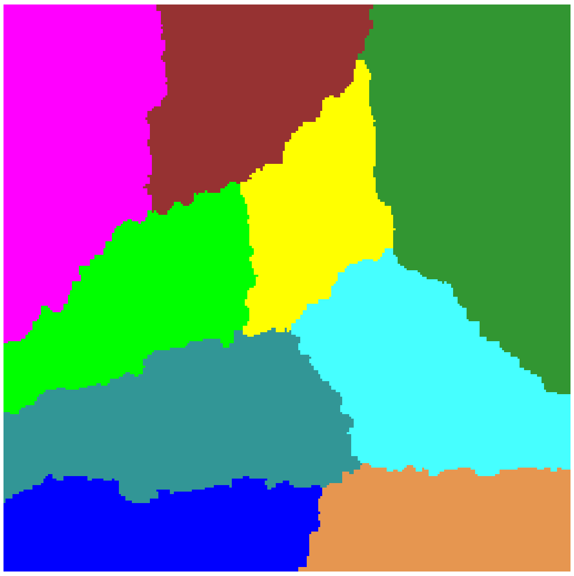
4.1 Number and Effect of Independent Parameters

At first sight, there are six independent parameters for the proposed tessellation: the Poisson intensity μ, an indicator parameter I for choosing the random or the majority boundary formation rule, and the four λ_i's. However, there are in fact only four independent parameters which have an influence on the final crystallization pattern: μ, I, $r_1 = \log(\lambda_2/\lambda_1)$ and $r_2 = \log(\lambda_3/\lambda_1)$. This follows, firstly because an $x \in C_4$ has a defined fate regardless of the value of λ_4, and so λ_4 is redundant, and secondly because the sum of all probabilities defined by (1) is one.

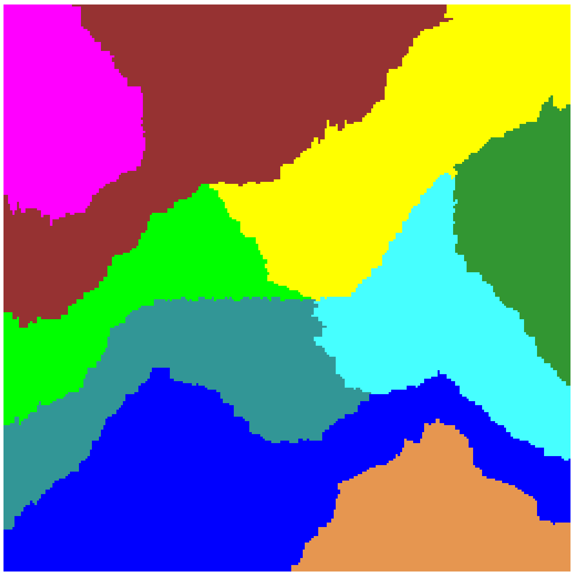
It is not hard to imagine that different combinations of r_1, r_2 and I would have different effects on the final crystallization pattern. Fig. 6 attempts to demonstrate this: it seems that I tends to have a bigger impact on the local behaviour of the region boundaries (the majority rule appears to give smoother boundaries; see the next subsection), while r_1 and r_2 tend to affect the overall shapes of the regions.



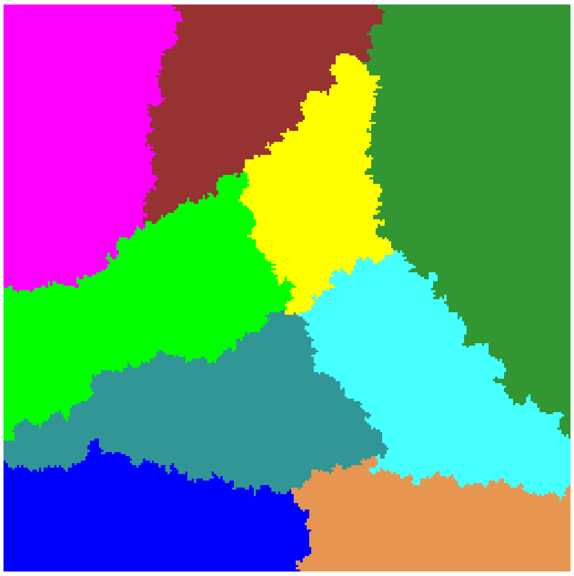
(a) random rule, $r_1 = r_2 = 0$



(b) majority rule, $r_1 = r_2 = 0$



(c) random rule, $r_1 = r_2 = 10$



(d) random rule, $r_1 = r_2 = -10$

Figure 6: *Effects of different parameter combinations. All four realizations were started with the same set of initial seeds.*

4.2 Boundary Roughness

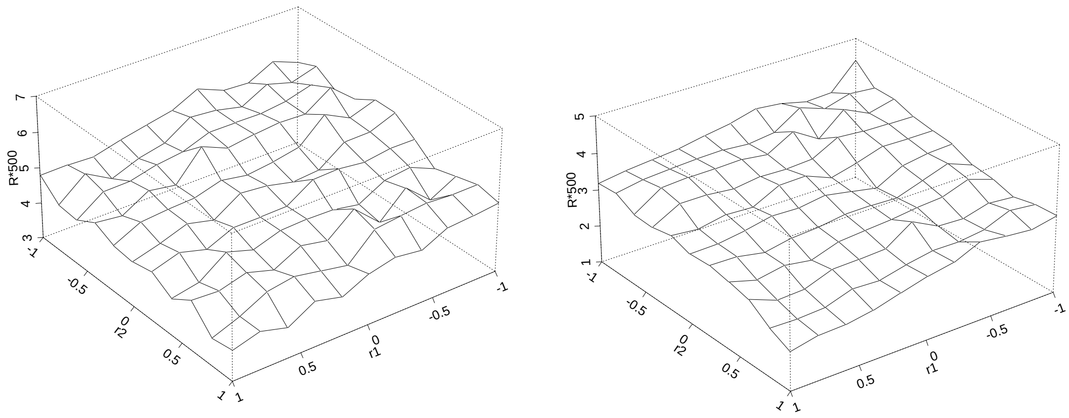
As suggested before, by changing r_1, r_2 and/or applying a different boundary formation rule, we can change the boundary roughness, and hence it is desirable to have some quantitative measure of the boundary roughness. Fractal dimension is *not* suitable for the following reason. If we look at the boundaries at the microscopic level, all these boundaries just consist of horizontal and vertical straight lines. That is, they are "smooth" and always have fractal dimension one (e.g., see Barnsley 1988) no matter how rough it appears to human eyes.

In order to investigate the effects of changing the parameter values on boundary roughness, we did the following. First, we measure the distance of the boundary of the grain from the centre of its initial pixel (or its centroid if the initial pixel is unknown) as a function of angle or arc–length. Call these distances the grain radii. It is usually convenient to choose "equally spaced" radii. Now suppose we have chosen n radii $l_1, \ldots, l_n$ and let $\bar{l}$ be their mean. Then we compute

$$R = \frac{\sum_i (l_{i+1} - l_i)^2}{\sum_i (l_i - \bar{l})^2}$$

as our roughness measure. Notice that R is a random function of r_1, r_2 and the boundary formation rule indicator parameter I (we assume that the Poisson intensity μ is small enough not to have any effect on the boundary roughness). Also notice that the use of R implicitly assuming that all the grains are of "star–shape" (of course this "star–shape" assumption is not always satisfied, but it is not a major issue for the experimental work of this subsection; see the caption of Fig. 7).

Simulations show that R (or more precisely, the expected value of R) heavily depends on the choice of the boundary formation rule: the random rule *always* gives a larger value of R than the majority rule with the same r_1 and r_2. Also, smaller values of r_1 and r_2 tend to produce a larger value of R with r_1 being more influential; see Fig. 7.

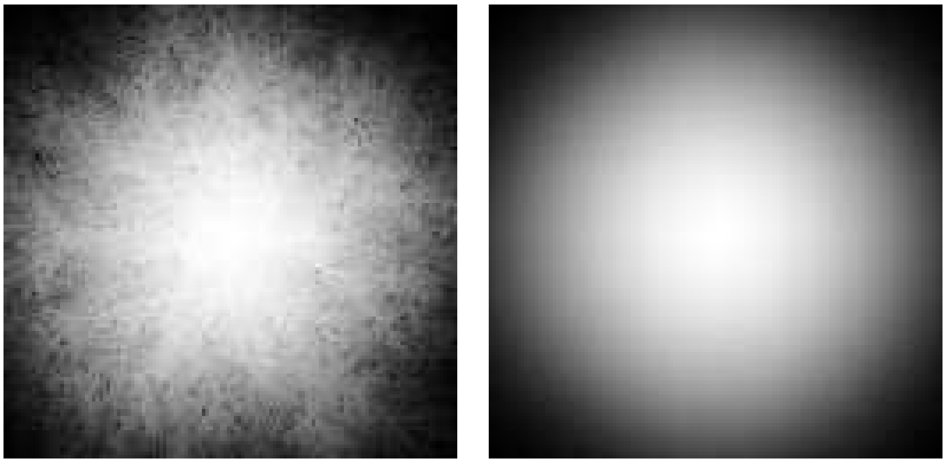


(a) random rule (b) majority rule

Figure 7: *Perspective plots demonstrating the effect of various r_1, r_2 and boundary formation rules on R. Each data point is the average of fifty different simulations. It is clear that, for both boundary formation rules, the roughness measure R is a decreasing function of r_1 and r_2. Another observation is: the random rule always gives a larger R than the majority rule. In order to speed up simulation times and minimize random errors, all simulations are done with the same nine regularly spaced initial pixels (since the seeds were regularly spaced, quite often the grains were all of "star–shape").*

4.3 Is the Tessellation Isotropic?

Since we are using a 4–connected system, we only have two growth directions: horizontal and vertical, and hence the stochastic growth mechanism and the final tessellation pattern are not isotropic. Somewhat surprisingly, our simulations suggest that both the growth process and the tessellation are isotropic for practical purposes, provided that the Poisson intensity is not too large. The results of formal hypothesis testing (e.g., testing differences between radii at different angles) are also consistent with a hypothesis of isotropy. Of course this isotropic



(a) (b)

Figure 8: *These two images attempt to show the time of labelling for each pixel in the domain: the brighter the pixel is, the sooner it was labelled. (a): one realization of the presented growth process (with equal λ_i's) with one single initial pixel located in the centre of the image. (b): average of two hundred such realizations. From these images the isotropic property of the growth process is apparent. Size of each realization: 256 $\times$ 256.*

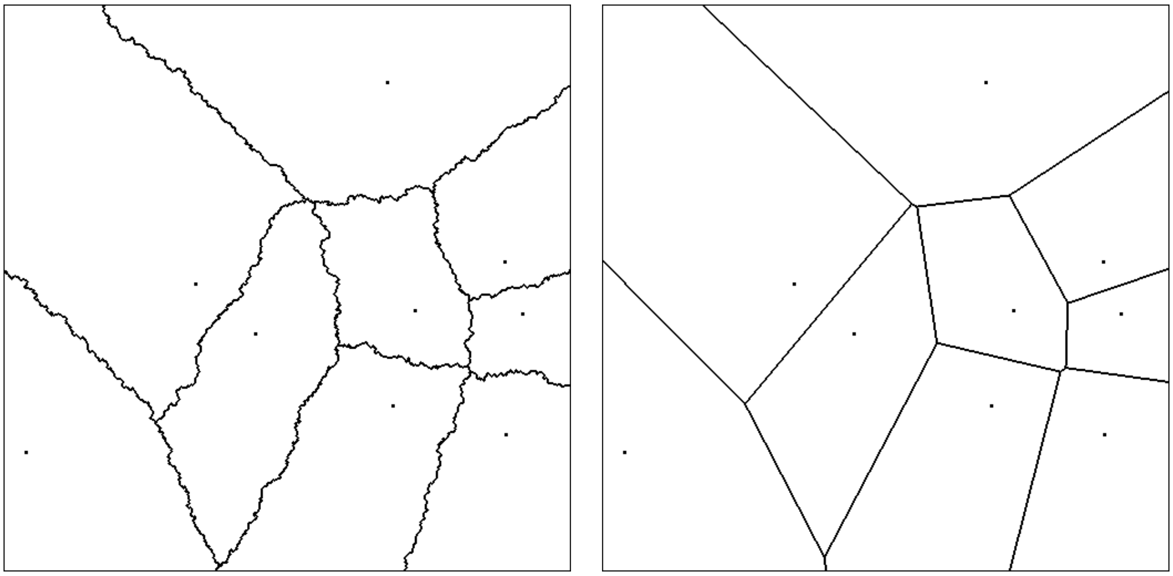
property of the tessellation pattern no longer holds if we apply the last vertical scaling step.

Fig. 8 shows a small scale simulation demonstrating the isotropic property of the growth process. Batchelor & Henry (1991) have performed a more extensive simulation for the Eden model (2^{25} pixels involved), which also demonstrates isotropy. Various researchers have attempted to obtain the asymptotic shape of different growth processes (e.g., Bramson & Griffeath 1981 and Durrett & Liggett 1981) as the grain size increases, but the results are so far limited.

4.4 Asymptotic to the Voronoi Tessellation?

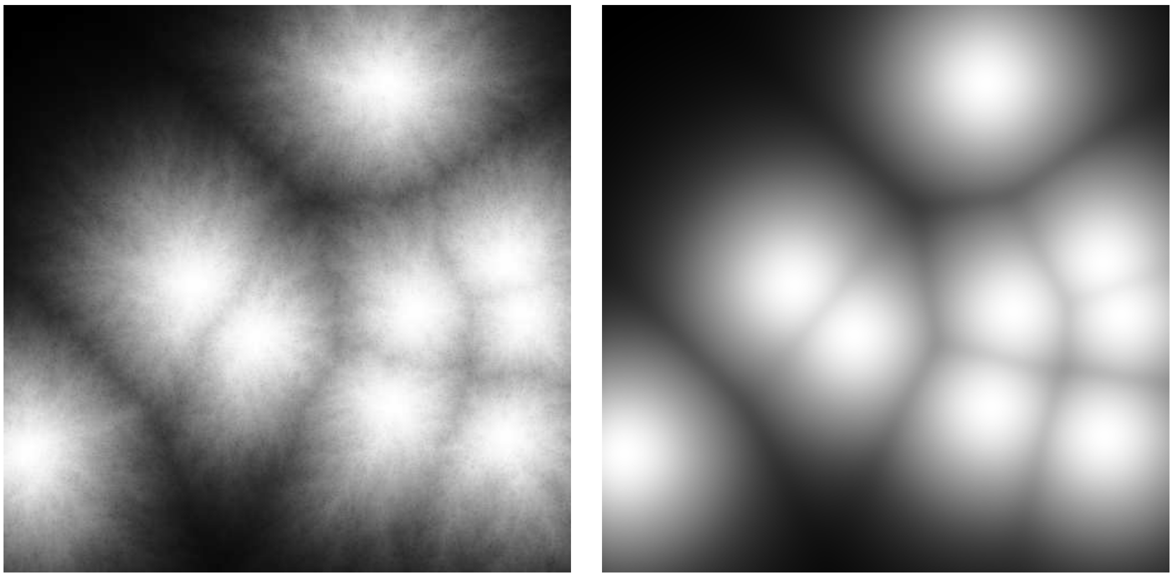
If the stochastic growth process is in fact isotropic, a sensible question to ask is: does the proposed tessellation converge to the Voronoi tessellation (with the same set of initial seeds) in some asymptotic sense?

There are two common spatial asymptotics discussed in the literature (e.g., see Stein 1995 and references therein): (i) increasing–domain asymptotic, in which the number of pixels in the image goes to infinity; and (ii) fixed–domain asymptotic, also called in–fill asymptotic, in which the pixel size shrinks to zero. However, we have not investigated these two asymptotics. We have, instead, briefly investigated a simpler third property: the behaviour of the "mean" of the boundaries of n independent realizations (all with the same seed points). Fig. 9, which shows the mean of 300 such realizations, suggests that the mean boundary is a *discrete* space version of the Voronoi tessellation. We do not have any theoretical results supporting our claim, only empirical results such as that shown in Fig. 9.



(a) proposed tessellation

(b) Voronoi tessellation



(c) one realization of growth time

(d) averages of three hundred realizations

Figure 9: *These images aim to answer the question: is the proposed tessellation asymptotic to the Voronoi tessellation? (a): one realization of the proposed tessellation together with initial Poisson seeds (black dots). (b): the corresponding Voronoi tessellation. (c): time of labelling for (a) (see caption of Fig. 8 for an explanation). (d): averages of three hundred realizations of time of labelling.*

5 Possible Variants

Variants of the proposed tessellation model are of course possible. For example, we can start the growth process with non–Poisson initial seeds, or we can modify the growth process such that the growth in the horizontal (or the vertical) direction is faster than the growth in the other direction (this issue will be further discussed in Section 6). Below we discuss two simple but interesting variants: a Johnson–Mehl type modification and a three–dimensional (3D) extension.

5.1 Johnson–Mehl Type Modification

A simple modification to the proposed tessellation is to allow the birth of new seeds at any time step during the growth process. This kind of modification is often called a Johnson–Mehl type modification (e.g., see Stoyan et al. 1987, Chapter 10). It often produces more interesting and complex patterns than the original model. For example, it is easy to find a smaller grain lying totally inside a bigger grain. See Fig. 10 for a realization.

5.2 Three Dimensional Extension

The formulation of the proposed model permits simple and straightforward higher dimensional extensions. Fig. 11 displays different cross sections of a realization of such a 3D extension. One noteworthy observation is that, if a (3D) grain has a U–shape, then in some cross sections it may appear as two distinct grains.

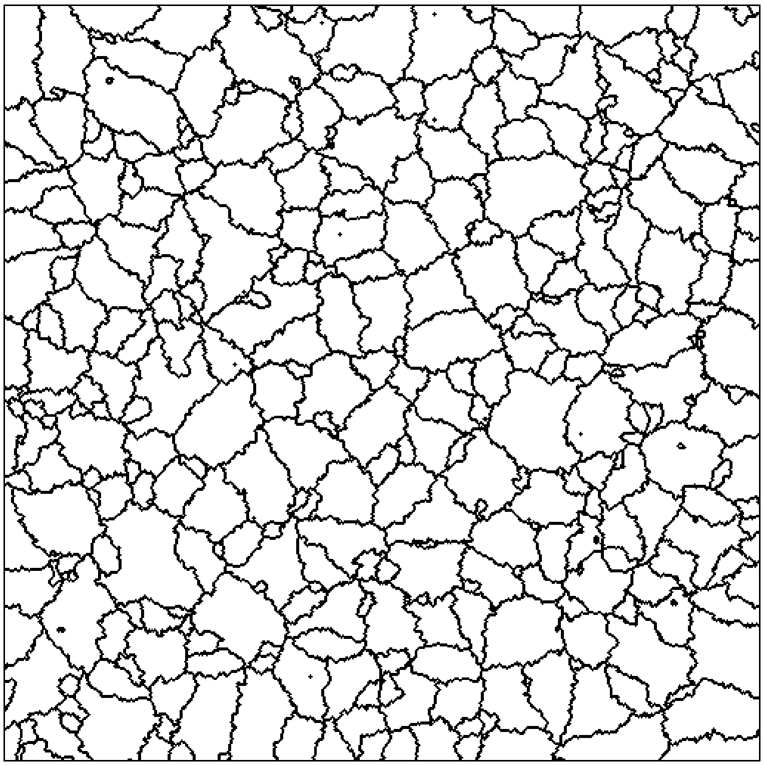


Figure 10: *A realization of the Johnson–Mehl type modification.*

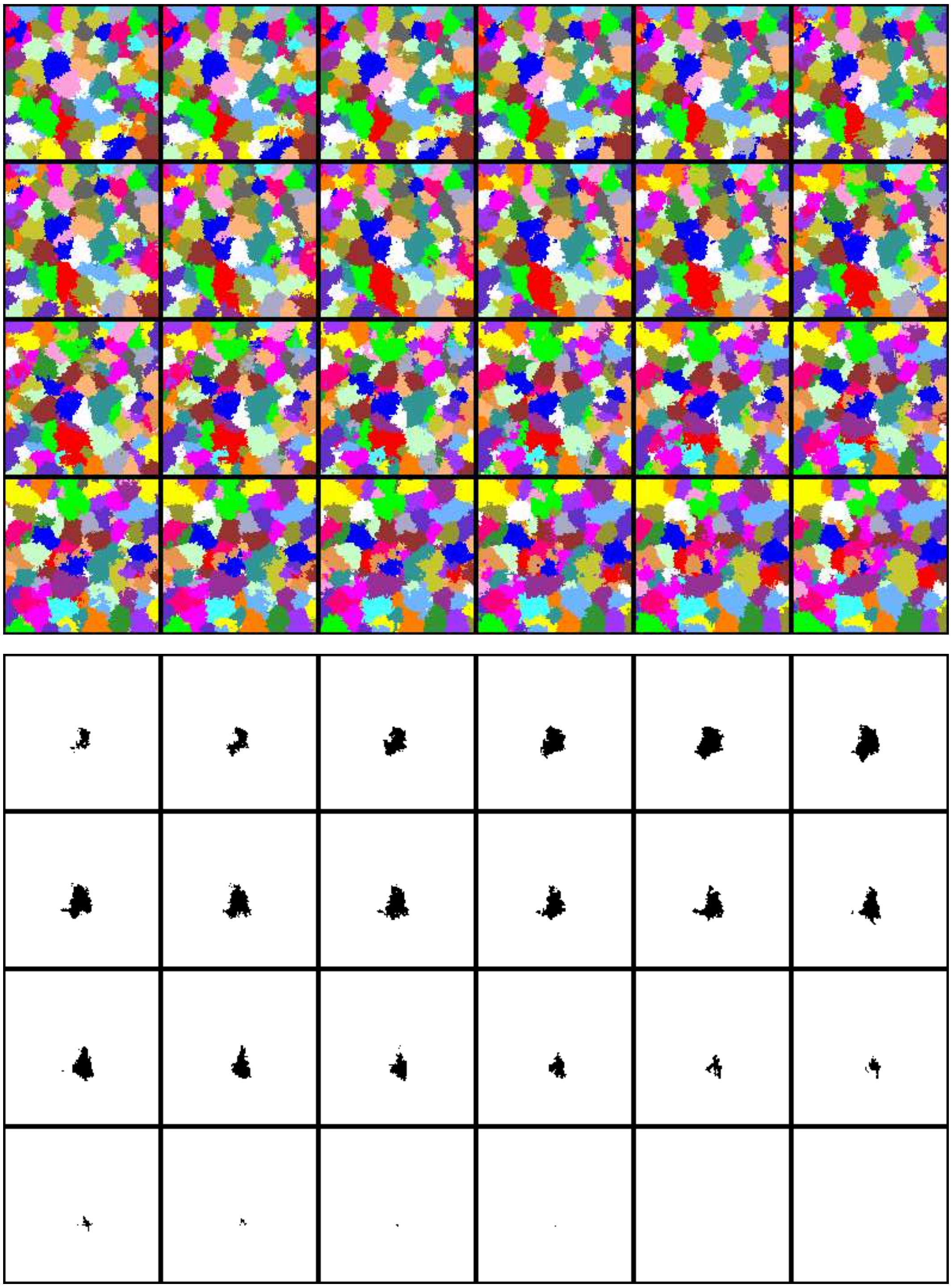


Figure 11: *Three dimensional extension. Top: 24 successive cross sections of one 128 $\times$ 128 $\times$ 128 realization. Bottom: one single highlighted grain.*

6 Shape Modelling of Aluminium Grain Images

This section demonstrates how the proposed tessellation model can be applied to simulate the shape of the aluminium grains. The semi–automatically segmented aluminium grain image displayed in Fig. 2 will be taken as the target for our simulation.

6.1 Choice of Model Parameters

Recall that there are 4 independent model parameters (see Section 4.1): λ, I, r_1 and r_2. Given a segmented image such as the one displayed in Fig. 2, the intensity of the Poisson seeds λ can be estimated by

$$\frac{\text{number of segmented grains}}{\text{area of the image}}.$$

(This estimator for λ is biased as it suffers from edge effects, but since this is not an important issue here, hence is ignored.)

Our empirical experience suggests that the effects of I, r_1 and r_2 are somewhat confounded. This makes estimation of I, r_1 and r_2 difficult. So we arbitrarily choose

$$I = \text{“random boundary formation rule”} \qquad \text{and} \qquad r_1 = r_2 = 1$$

to simulate the shape pattern.

6.2 Choice of Scaling Factor

Since the widths of the aluminium grains tend to be longer than the heights, a further step of rescaling with respect to the vertical direction is necessary. Two simple statistics, ρ_{MEAN} and ρ_{MAX}, have been used to estimate the scaling factor of the image displayed in Fig. 2. They are defined as:

$$\rho_{\text{MEAN}} = \frac{1}{k}\sum_{j=1}^{k}\frac{\text{width}_{\text{MEAN},j}}{\text{height}_{\text{MEAN},j}} \qquad \text{and} \qquad \rho_{\text{MAX}} = \frac{1}{k}\sum_{j=1}^{k}\frac{\text{width}_{\text{MAX},j}}{\text{height}_{\text{MAX},j}},$$

where

$$\begin{aligned}
\text{width}_{\text{MEAN},j} &= \text{(area of the } j\text{th segmented grain)/(number of rows it occupies)},\\
\text{height}_{\text{MEAN},j} &= \text{(area of the } j\text{th segmented grain)/(number of columns it occupies)},\\
\text{width}_{\text{MAX},j} &= \text{maximum number of pixels the } j\text{th segmented grain occupies in one row},\\
\text{height}_{\text{MAX},j} &= \text{maximum number of pixels the } j\text{th segmented grain occupies in one column},
\end{aligned}$$

and k is the total number of segmented grains in the image. For the segmented image in Fig. 2, both ρ_{MEAN} and ρ_{MAX} produced estimates remarkably close to 2, and so we took the scaling factor to be 2. See Fig. 19(b) for such a scaled tessellation.

Other than rescaling, another way to generate elongated grains is to modify the growth process in such a way that the growth rate in the horizontal direction is faster than the vertical direction. However we did not pursue this approach. The reasons are that it is unclear how the relative growth rate can be estimated from an image and that the real aluminium grains were in fact flattened (rescaled) after their formation.

7 Colour Modelling

The primary concern of this section is the modelling of the *colour* characteristics of the aluminium grain images. Once a scaled tessellation is obtained (e.g., by using the tessellation model described before), the next stage for simulating an aluminium grain image is to assign a colour value to each of the simulated "empty" regions. We first note that, for a given real aluminium grain image, it is not always possible to tell whether a region with all its pixels having similar colour values is in fact one true grain, or is composed of two or more neighbouring grains with similar colours. In these situations, we treat such a region as a single grain, and we will ignore the spatial relationship amongst the colour values when assigning simulated colours to the empty regions. We shall concentrate on the modelling of the distribution of the colour values.

7.1 A Quadratic Model for Colour Values

In digital colour images, a particular colour is stored as three values: the red, green and blue (RGB) components. Therefore in our modelling, colours are treated as realizations of trivariate random variables. Given a real aluminium grain image, how can one extract data in a format that can be worked with (i.e., a series of trivariate colour values)? One way to achieve this is to first segment the image, and then for each segmented grain, compute the mean RGB values. These computed mean values, denoted by (r_j, g_j, b_j), $j = 1, \ldots, n$, with n as the total number of RGB values, can be taken as the "workable data". The image in Fig. 2 has 164 segmented grains, and the corresponding (r_j, g_j, b_j)'s are plotted pairwise in Fig. 12. These extracted RGB values will be taken as the observed data that we will work with. Fig. 12 seems to suggest that these observed RGB values are related by a quadratic relationship in three–dimensional (3D) colour space.

Let (R_j, G_j, B_j) be the "true" but unobservable colour of the ith segmented grain. We propose using the following simple but effective model for the observed RGB values (r_j, g_j, b_j)'s:

$$\begin{aligned}
r_j &= R_j &&\sim N(\mu_R, \sigma_R^2), \\
b_j &= B_j + \epsilon_j &&= \alpha_0 + \alpha_1 R_j + \alpha_2 R_j^2 + \epsilon_j \qquad (2) \\
g_j &= G_j + \delta_j &&= \beta_0 + \beta_1 R_j + \beta_2 B_j + \beta_3 R_j^2 + \beta_4 B_j^2 + \beta_5 R_j B_j + \delta_j
\end{aligned}$$

where the α's and β's are model parameters, and R_j, δ_j and ϵ_j are independent Gaussian

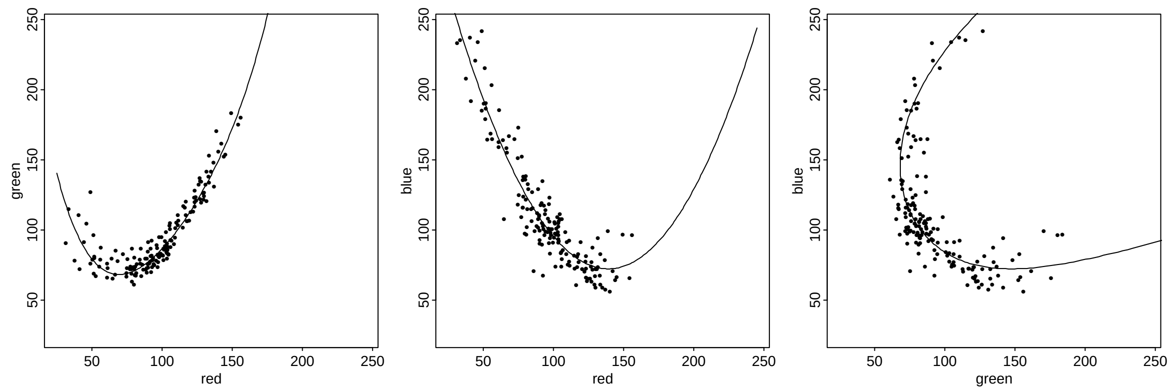


Figure 12: *Pairwise plots of raw RGB values, with fitted model (2) in solid lines.*

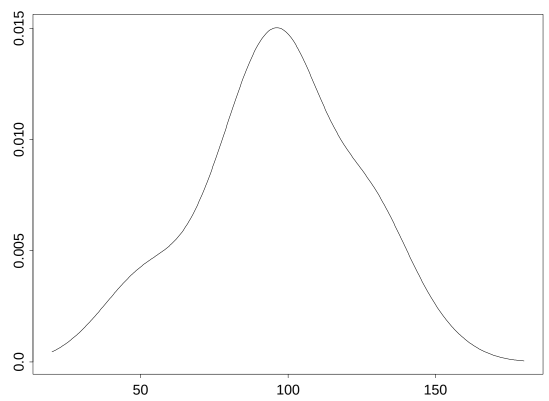


Figure 13: *Nonparametric density estimate of the r_j's.*

random variables: $\delta_j, \sim N(0, \sigma_\delta^2)$, $\epsilon_j \sim N(0, \sigma_\epsilon^2)$ (by $X \sim N(\mu, \sigma^2)$ we mean X is a Gaussian random variable with mean μ and variance σ^2). Notice that in the above model both g_j and b_j are assumed to be observed with errors (δ_j and ϵ_j), while r_j is error–free, and is assumed to be Gaussian. Fig. 13 is a nonparametric density estimate (e.g., see Silverman 1986) of r_j, which empirically justifies the assumed Gaussianity of the r_j's. A more sophisticated model in which r_j, g_j and b_j are all assumed to be error–contaminated can be found in Lee (1997).

We understand that the above quadratic model (2) does not bear any physical explanation

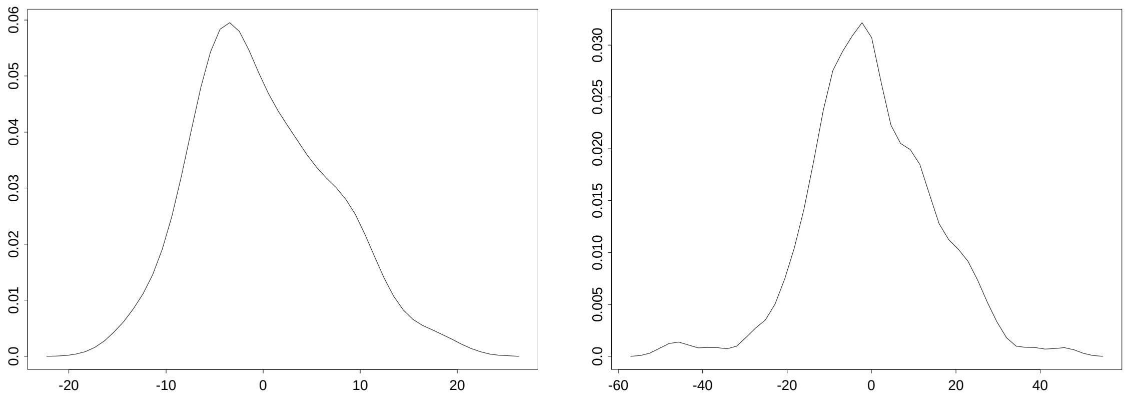


Figure 14: *Nonparametric density estimates of residuals of g_j (left) and b_j (right).*

to the generation of the observed "true" colour values. However, because of its simplicity and efficiency, we recommend its use here.

We computed the maximum likelihood estimates (which are also the least–square estimates in this situation) of the parameters. They are listed in Table 1. The corresponding fitted 3D quadratic curve is plotted in Fig. 12. Also, nonparametric density estimates of the residuals of the fitted g_j and b_j are plotted in Fig. 14; these residuals are reasonably zero–meaned Gaussians.

Parameters	μ_R	σ_R^2	α_0	α_1	α_2	σ_ϵ^2
Estimates	96.40	26.63^2	367.8	-4.253	0.01530	14.43^2

Parameters	β_0	β_1	β_2	β_3	β_4	β_5	${\sigma_\delta}^2$
Estimates	292.4	-3.177	-1.956	0.01667	0.005277	0.008784	6.882^2

Table 1: *Listing of parameter estimates.*

7.2 Simulation of Artificial Aluminium Grain Colours

To simulate an artificial aluminium grain colour value, we first simulate an r_j from $N(\hat{\mu}_R, \hat{\sigma}_R^2)$, where $\hat{\mu}_R = 96.40$ and $\hat{\sigma}_R^2 = 26.63^2$, and then simulate g_j and b_j using model (2) with parameter estimates given in Table 1. Fig. 15 displays pairwise plots of such a set of simulated colour values, and these simulated colour values can be assigned to the empty regions of a tessellation to form a artificial (noise–free) colour aluminium grains image. However, one additional step is required after each empty region is assigned with a colour: if the Euclidean distance between the simulated colour values of two neighbouring regions is less than a pre–set threshold, these two neighbouring regions are merged together. Because of this additional merging step, the intensity of the Poisson seeds for simulating a tessellation should be increased. Fig. 19(c) provides an example of a simulated colour image.

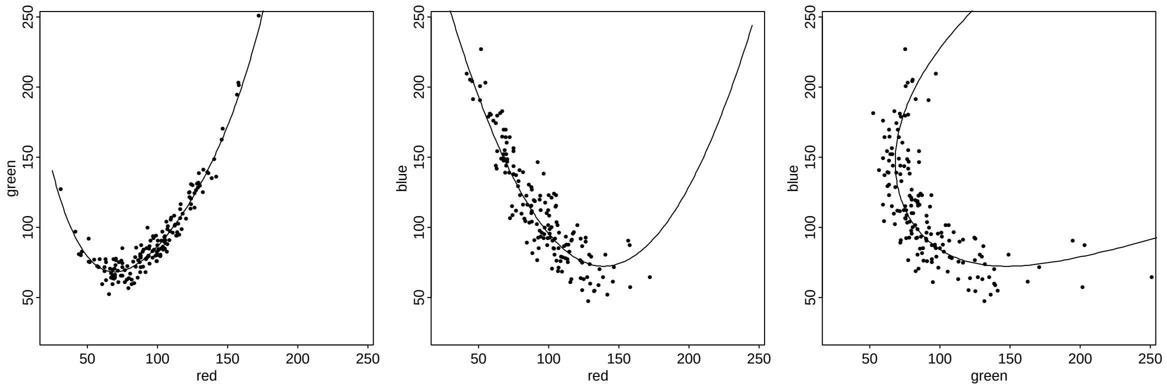


Figure 15: *Pairwise plots of simulated colour values. Also plotted in solid lines is the fitted model (2). Compare these plots with Fig. 12.*

8 Noise Modelling

This section considers the modelling of the *noise* characteristics of the aluminium grain images. A careful examination of the aluminium grain images suggests that the noise is *spatially correlated*; Fig. 16 displays a cross–section of the empirical autocovariance function of the "extracted noise" (see below) of the red band, which supports the claim that the noise is spatially correlated. It will be convenient to model this noise using a stationary Gaussian random field (SGRF), and hence it has been decided to use SGRFs to model the noise (e.g., see Chellappa & Kashyap 1985). Therefore our aim is: given an observed noise image which can be well modelled by an SGRF T, we would like to simulate other noise images (i.e., other SGRFs) which would *look like* and *possess similar statistical properties* to T.

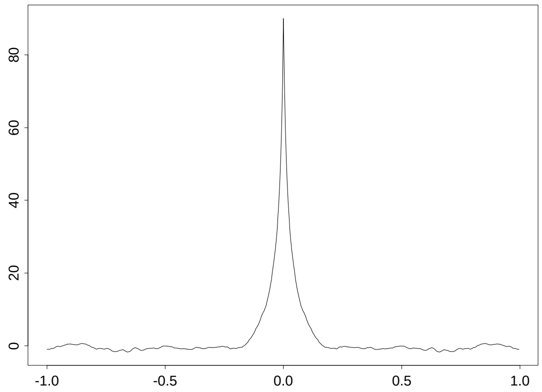


Figure 16: *Cross–section of the empirical autocovariance of extracted noise.*

We shall use the technique developed in Lee & Berman (1997) to achieve our aim. Specifically, Lee & Berman (1997) proposed an automatic spectrum estimation procedure, which is designed to produce an estimated spectrum of an observed SGRF T in such a way that any subsequent SGRFs simulated from this estimated spectrum would look like and possess similar statistical properties to T.

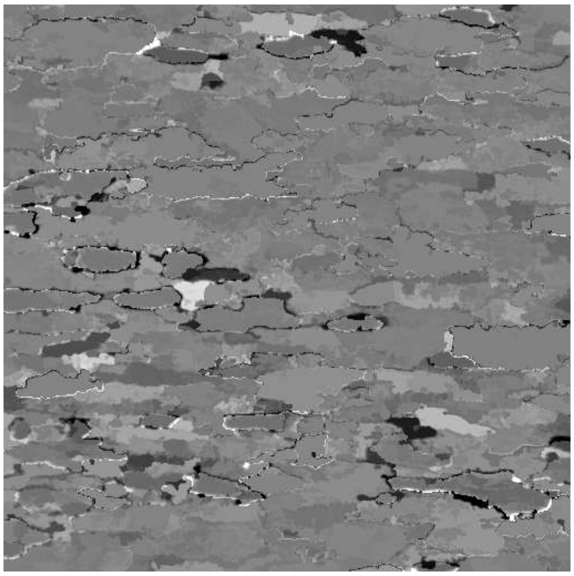
Before we proceed, we first mention one legitimate suspicion is that, the noise is not only spatially correlated, but is also correlated *across* the colour bands (e.g., the noise of the red band is correlated with the noise of the green band). However for simplicity, this across–band correlation will be ignored.

8.1 Noise Modelling of Aluminium Grain Images

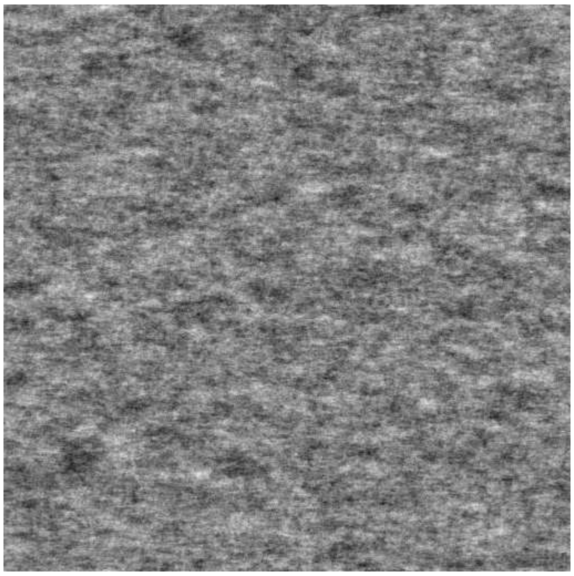
As for the colour modelling, given a real aluminium grain image, the first step of our noise modelling is to extract data in a workable format. We can apply the same trick: given a real image, we segment it and compute the mean colour values for each segmented grain. Then the difference between this mean–computed segmented image and the real observed image can be taken as the noise. See Fig. 17(a) for the "extracted noise" of the red band of the image in Fig. 2.

Observe that there are some strong, narrow and long features around the grain boundaries, and that the extracted noise may not be well modelled by an SGRF. Even though these features are present, we applied the technique developed by Lee & Berman (1997) to estimate the spectrum of the noise in Fig. 17(a), and from the estimated spectrum, an SGRF is generated; see Fig. 17(b).

When comparing visually the simulated noise in Fig. 17(b) with the extracted noise in Fig. 17(a), however, the result is unsatisfactory. We attribute this failure partially to the fact that the extracted noise cannot be well modelled by an SGRF. In Fig. 18 is the cross–section of the horizontal middle line of the extracted noise, which clearly shows that there are sharp jumps in the extracted noise. Such jumps, which are artifacts resulting from the imperfection of the semi–automatic segmentation and the blurred grain boundaries, violate the smoothness



(a) extracted noise



(b) simulated noise

Figure 17: *Extracted noise and the corresponding simulated noise.*

requirement of SGRFs, which has, in turn, caused the above failure. Nevertheless, we ignored this failure and added one such simulated SGRF (from the estimated spectrum of the extracted noise) to each of the three components of a simulated colour image to form our fully simulated aluminium grain image; see Fig. 19(d).

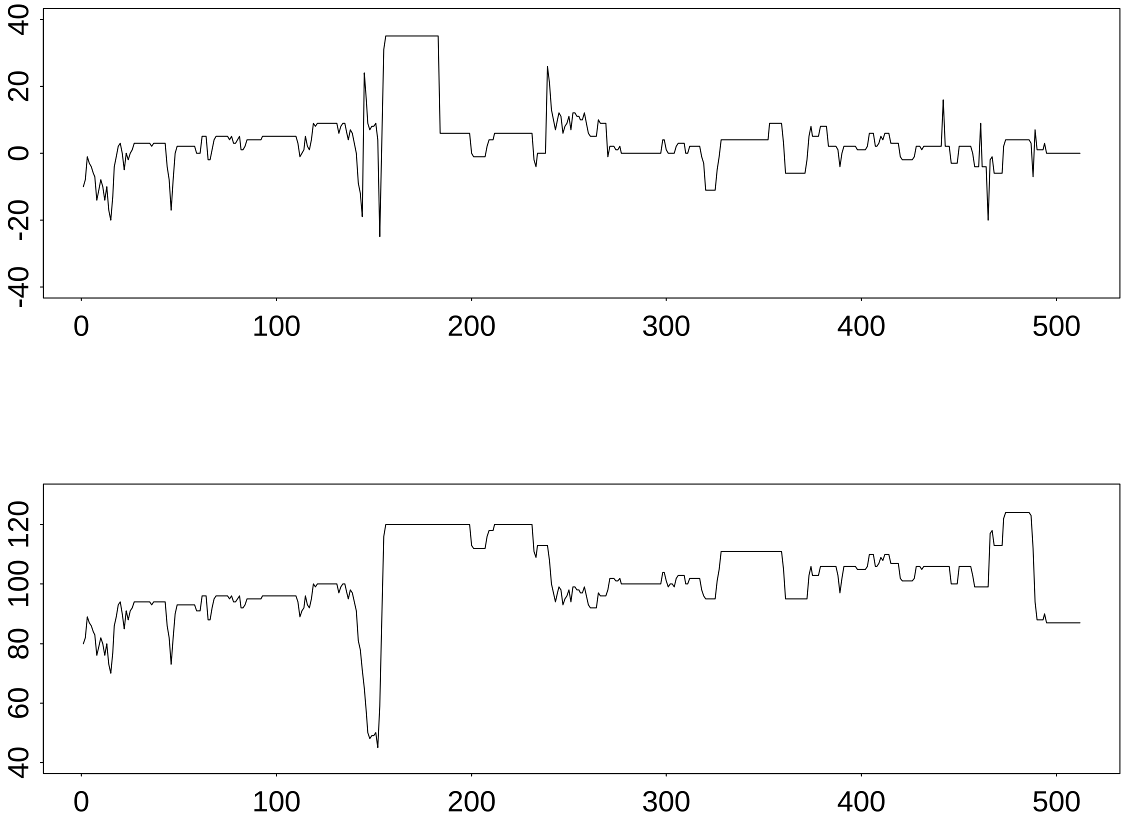
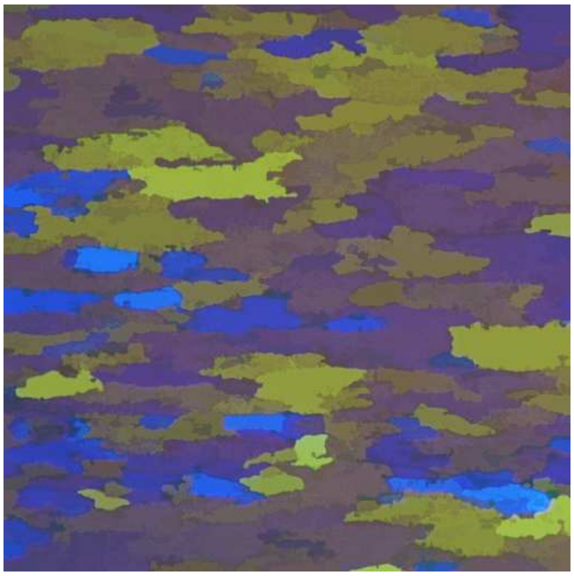
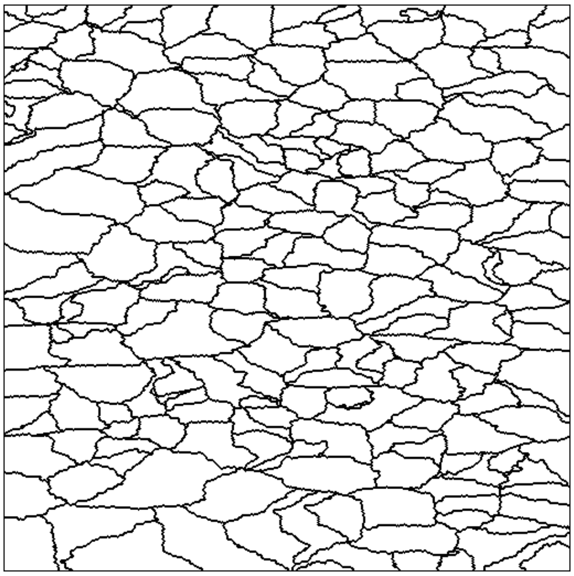


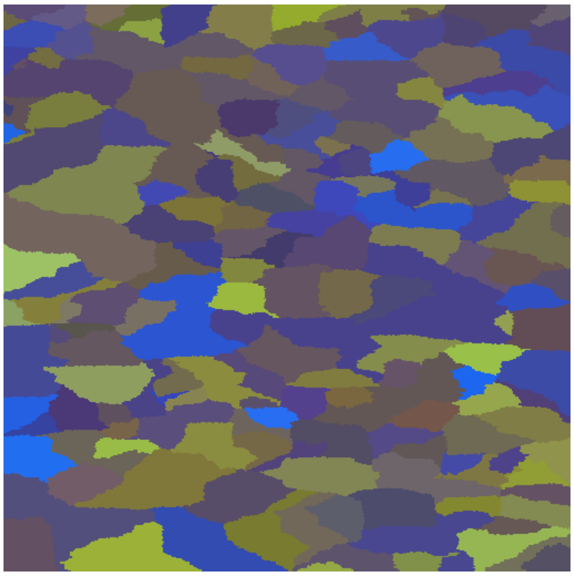
Figure 18: *Cross–sections of the extracted noise (top) and the corresponding red band of the aluminium grain image (bottom). The square feature of some parts of the plots is due to discretization effect when recording the image.*



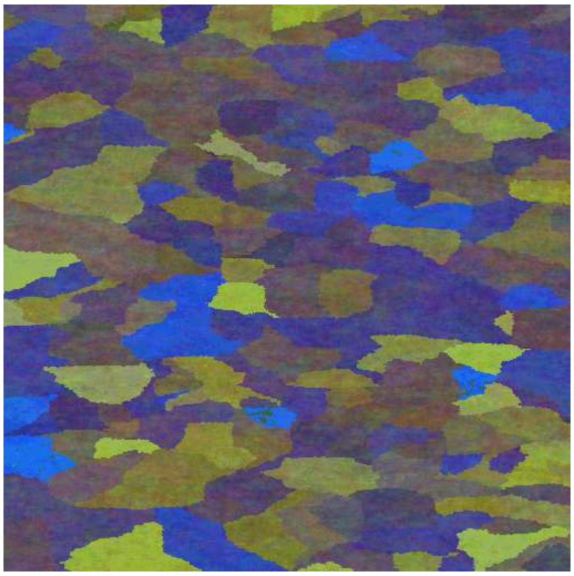
(a) real aluminium grain image



(b) shape modelling



(c) shape + colour modelling



(d) shape + colour + noise modelling

Figure 19: *Modelling of aluminium grain image.*

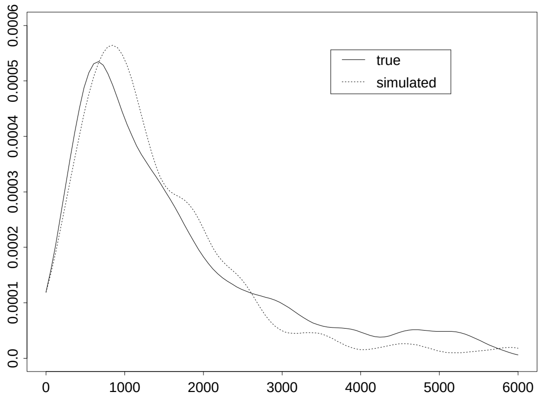
9 Comparing Real and Simulated Images

A visual comparison of a real aluminium grain image, Fig. 19(a), and a final simulated image, Fig. 19(d), suggests that our model generates tessellations which have more regular boundaries than do the real aluminium grains. This can be partially overcome by a 3D version of the tessellation model (see Section 5.2). Because the 3D simulations take much longer, we recommend using the 2D version.

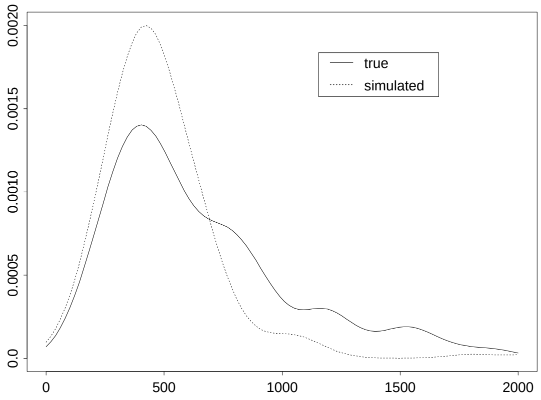
In order to compare the real and the simulated images in a quantitative fashion, we did the following. We computed nonparametric density estimates of the area and perimeter of the grains in the real and simulated images. These estimates are plotted in Fig. 20(a) and Fig. 20(b). Also, we fitted an ellipse to each individual grain. Density estimates of the ratios of the length of the fitted ellipse major axes to the length of the minor axes are plotted in Fig. 20(c), while estimates of the orientation of the major axes are plotted in Fig. 20(d). From these plots we can see that the real and simulated images are similar in global characteristics.

10 Concluding Remarks

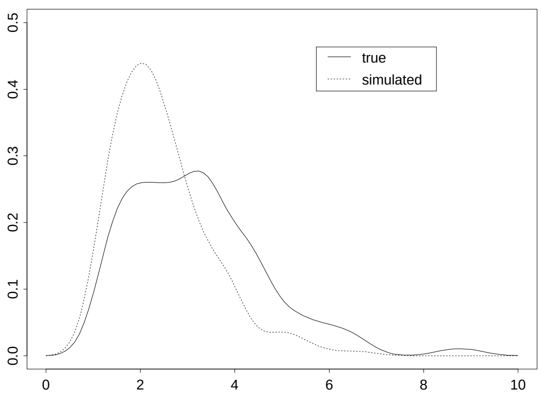
In this article we have described various models and methods for simulating artificial colour aluminium grain images. In particular, a new tessellation model is proposed. This tessellation model serves our purposes well: it is capable of generating tessellations with rough boundaries, which can be used as the basis of building synthesized aluminium grain images. In addition, it does not need a long computation time to obtain one realization: for an image of dimension 512×512 with about 220 initial seeds, it takes about 150 seconds of user time on a Sparc–10 workstation.



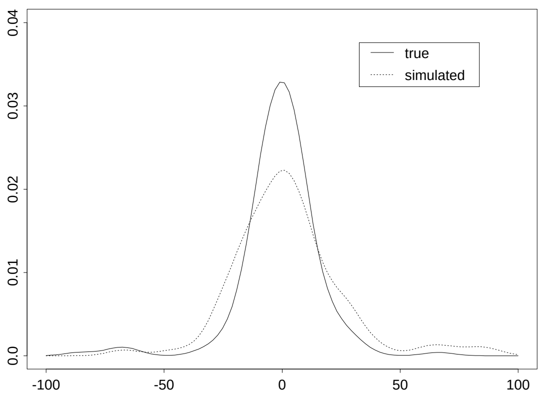
(a) area



(b) perimeter



(c) axes ratio



(d) orientation

Figure 20: *Comparing real and simulated images.*

Acknowledgement

Most of the work was completed while the author was a Ph.D. candidate at Macquarie University, Sydney, and CSIRO Mathematical and Information Sciences, Australia. He would like to express his gratitude to his supervisors, Mark Berman and Victor Solo, for their continuous encouragement, guidance and patience. He would also like to thank Richard Cowan for many discussions which bordering on coauthorship, Adrian Baddeley for making many useful comments on Lee (1997), and a referee for many constructive comments. This document was prepared using computer facilities supported in part by National Science Foundation Grant DMS 89-05292 awarded to the Department of Statistics at The University of Chicago, and by The University of Chicago Block Fund.

References

Ahuja, N. & Schachter, B. J. (1983), *Pattern Models*, John Wiley & Sons Ltd, New York.

Barnsley, M. (1988), *Fractals Everywhere*, Academic Press.

Batchelor, M. T. & Henry, B. I. (1991), 'Limits to Eden growth in two and three dimensions', *Physics Letters A* **157**, 229–236.

Brémond, R. & Jeulin, D. (1994*a*), Morphogenesis simulations with lattice gas, *in* J. Serra & P. Soille, eds, 'Mathematical morphology and its applications to image processing', Kluwer Academic Publishers, pp. 297–304.

Brémond, R. & Jeulin, D. (1994*b*), Random media and lattice gas simulations, *in* M. Armstrong & P. A. Dowd, eds, 'Geostatistical Simulations', Kluwer Academic Publishers, pp. 89–105.

Bramson, M. & Griffeath, D. (1981), 'On the Williams–Bjerknes tumour growth model I', *The Annals of Probability* **9**, 173–185.

Chellappa, R. & Kashyap, R. L. (1985), 'Texture synthesis using 2-D noncausal autoregressive models', *IEEE Transactions on Acoustics, Speech, and Signal Processing* **ASSP-33**, 194–203.

Cowan, R. & Tsang, A. K. L. (1994), 'The falling-leaves mosaic and its equilibrium properties', *Advances in Applied Probability* **26**, 54–62.

Durrett, R. (1980), 'On the growth of one dimensional contact processes', *The Annals of Probability* **8**, 890–907.

Durrett, R. & Liggett, T. M. (1981), 'The shape of the limit set in Richardson's growth model', *The Annals of Statistics* **9**, 186–193.

Eden, M. (1961), A two dimensional growth process, *in* J. Neyman, ed., 'Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability', Vol. IV, University of California Press, Berkeley, CA.

Griffeath, D. (1981), 'The basic contact processes', *Stochastic Processes and their Applications* **11**, 151–185.

Harris, T. (1974), 'Contact interactions on a lattice', *The Annals of Probability* **2**, 969–988.

Jeulin, D. (1979), Morphologie Mathématique et Propriétés Physiques des agglomérés de minerais de fer et du coke métallurgique, PhD thesis, Paris School of Mines.

Jeulin, D. (1980), 'Multi–component random models for the descrption of complex microstructures', *Mikroskopie* **37**, 130–137.

Jeulin, D. (1993), 'Random models for morphological analysis of powders', *Journal of Microscopy* **172**, 13–21.

Lee, T. C. M. (1997), Some Models and Methods in Image Segmentation, PhD thesis, Macquarie University, Sydney, Australia.

Lee, T. C. M. & Berman, M. (1997), 'Nonparametric estimation and simulation of two–dimensional Gaussian image textures', *Graphical Models and Image Processing* **59**, 434–445.

Lee, T. C. M. & Cowan, R. (1994), A stochastic tessellation of digital space, *in* J. Serra & P. Soille, eds, 'Mathematical morphology and its applications to image processing', Kluwer Academic Publishers, pp. 217–224.

Matheron, G. (1968), *Schéma Booléen Séquentiel de Partition Aléatoire*, N–83 CMM, Paris School of Mines Publication.

Meakin, P. & Sander, L. M. (1985), 'Comment on "Active zone of growing clusters: Diffusion–limited aggregation and the Eden model"', *Physical Review Letters* **54**, 2053.

Okabe, A., Boots, B. & Sugihara, K. (1992), *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*, John Wiley & Sons Ltd, Chichester.

Richardson, D. (1973), 'Random growth in a tessellation', *Mathematical Proceedings of the Cambridge Philosophical Society* **74**, 515–528.

Schwertel, J. & Stamm, H. (1997), 'Analysis and modelling of tessellations by means of image analysis methods', *Journal of Microscopy* **186**, 198–209.

Silverman, B. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London.

Stein, M. L. (1995), 'Fixed–domain asymptotics for spatial periodograms', *Journal of the American Statistical Association* **90**, 1277–1288.

Stoyan, D., Kendall, W. & Mecke, J. (1987), *Stochastic Geometry and its Applications*, Wiley: Chichester.

Thompson, S. & Rosenfeld, A. (1993), Discrete stochastic growth models for two–dimensional shapes, *in* Y.-L. O, A. Toet, D. H. Foster, H. J. Heijmans & P. Meer, eds, 'Shape in Picture — Mathematical Description of Shape in Grey–level Images', Springer Verlag, Berlin, pp. 301–318.

Williams, T. & Bjerknes, R. (1972), 'Stochastic model for abnormal clone spread through epithelial basal layer', *Nature* **236**, 19–21.

Witten, T. A. & Sander, L. M. (1981), 'Diffusion–limited aggregation, a kinetic critical phenomenon', *Physical Review Letters* **47**, 1400–1403.