MICROTEXTURE INPAINTING THROUGH GAUSSIAN CONDITIONAL SIMULATION

Bruno Galerne*, Arthur Leclaire†, Lionel Moisan*

* Université Paris Descartes, MAP5, CNRS UMR 8145, France. † CMLA, ENS Cachan, France.

ABSTRACT

Image inpainting consists in filling missing regions of an image by inferring from the surrounding content. In the case of texture images, inpainting can be formulated in terms of conditional simulation of a stochastic texture model. Many texture synthesis methods thus have been adapted to texture inpainting, but these methods do not offer theoretical guarantees since the conditional sampling is in general only approximate. Here we show that in the case of Gaussian textures, inpainting can be addressed with perfect conditional simulation relying on kriging estimation. We thus obtain a microtexture inpainting algorithm that is able to fill holes of any shape and size in an efficient manner while respecting exactly a stochastic model.

Index Terms—Inpainting, Gaussian texture, texture synthesis, conditional simulation, kriging

1. INTRODUCTION

Since the seminal papers [1, 2, 3], random phase fields have proven to be a powerful tool for microtexture modeling. In particular, given a microtexture, a Gaussian model can be easily derived and sampled [4] and latter approximated by a spot noise model to allow for fast by-example synthesis [5]. In addition, the theoretical benefits of the Gaussian model have been exploited to address texture analysis [6] and texture mixing [7]. In this paper, we take advantage of the Gaussian model in order to address microtexture inpainting.

The inpainting problem consists in recovering hidden parts of an image based on the surrounding content. Many methods have been proposed for inpainting, falling into three categories. The first are deterministic methods which use a variational or PDE-based framework to fill the missing regions by connecting the edges or level lines in a way that respects the Gestaltist’s good continuation principle [8, 9, 10, 11, 12, 13, 14, 15]. The common drawback of these methods is their inability to handle texture precisely.

The second are stochastic methods. Inpainting is by essence an ill-posed problem. However, if one has a pertinent stochastic model \( p(u) \) for the whole image \( u \), a perfect inpainting solution can be obtained by conditional simulation, that is simulating an image \( u \) with distribution \( p(u) \) given the known values of \( u \). This stochastic point of view is not really helpful in general since one cannot sample the distribution of realistic images. However, this approach is of particular interest when dealing with texture images since they are often well represented by tractable stochastic models. The first works in the direction of stochastic inpainting suggest to fill the masked zone by matching specific statistics, like co-occurrence matrices [16] or histograms of filter responses [17]. Truly conditional methods later appeared based on Markov random fields, which are inherently adapted to conditional sampling [18]. Using patch comparison to approximate local conditional sampling, the authors of [19, 20] were led to a celebrated non-parametric texture synthesis algorithm which can also be used for hole filling. However, these synthesis algorithms may degenerate after a spatial boundary (growing garbage effect), and besides are quite dependent on the filling order, thus not ensuring the prolongation of large-scale features.

Taking profit of the first attempts in deterministic and stochastic inpainting, hybrid methods have emerged, which consist in filling both geometric and textural contents in a coherent manner. After the proposition of [21] to extend non-parametric sampling to general image synthesis, several authors improved this work by varying the pixel filling order [22, 23, 24]. Other methods suggest to separate geometry and texture and to inpaint them either with two different techniques [25] or to minimize a unified functional [26]. The last work paves the way for inpainting with adapted dictionaries [27, 28, 29, 30]. More recently, several authors gave a unified view of example-based image inpainting in a variational framework [31, 32].

These methods are able to produce a satisfying result on a large class of images with a reasonable computational time, but in general do not provide any theoretical guarantee in the synthesized content. In contrast, we demonstrate in this paper that if one restricts to a Gaussian random model, microtexture inpainting can be achieved by perfect conditional sampling. Our algorithm relies on kriging estimation [33] and fills holes with two independent components, namely the kriging component which extends long scale structures from the hole boundary, and the innovation component which adds an independently generated textural content containing small scale details. The proposed algorithm can inpaint holes of any shape and size, provided that a faithful Gaussian texture model can be estimated within the unmasked image area. For
large holes, this algorithm does not suffer from the *growing garbage* effect encountered in [19], and besides, edge-like structures in the textures are naturally restored in a way that respects the covariance.

Texture inpainting was already addressed in [34, 35] with kriging estimation, but, to the best of our knowledge, not with Gaussian conditional sampling. Let us mention however that local Gaussian conditional models were used for structured texture synthesis in [36].

### 2. GAUSSIAN CONDITIONAL SIMULATION

In this section, following [33], we rely on kriging estimation to explain the perfect conditional sampling of Gaussian random vectors that is at the heart of our microtexture inpainting algorithm.

Let $\Omega$ be a finite set. For a set $A \subset \Omega$ and a function $f : \Omega \to \mathbb{R}$ we denote by $|A|$ the cardinality of the finite set $A$, and $f|_A$ the restriction to $A$ of the function $f$. Let $(F(x))_{x \in \Omega}$ be a Gaussian vector with mean zero (for the sake of simplicity) and covariance function $\Gamma(x, y) = \text{Cov}(F(x), F(y)) = \mathbb{E}(F(x)F(y)), x, y \in \Omega$.

Given prespecified values $\varphi : C \to \mathbb{R}$ on a set $C \subset \Omega$ of conditioning points, conditional Gaussian simulation consists in sampling the conditional distribution of $F$ given that $F|_C = \varphi$.

One defines the **kriging estimator** at $x \in \Omega$ as the conditional expectation $F^*(x) = \mathbb{E}(F(x) | F(c), c \in C)$. This means that $F^*(x)$ is the best least-square estimation of $F(x)$ that can be obtained as a measurable function of $(F(c))_{c \in C}$. In particular for all $x \in C$, $F^*(x) = F(x)$.

A standard result of probability theory [37] ensures that in the Gaussian case $F^*(x)$ is the orthogonal projection of $F(x)$ on the subspace of linear combinations of $(F(c))_{c \in C}$ (for the $L^2$-distance between square-integrable random variables). Hence, there exist coefficients $(\lambda_c(x))_{c \in C}$ such that

$$F^*(x) = \sum_{c \in C} \lambda_c(x) F(c).$$

These deterministic numbers $(\lambda_c(x))_{c \in C}$ are called the **kriging coefficients**. The following theorem is the core result for conditional Gaussian simulation.

**Theorem 2.1** ([33]). $F^*$ and $F - F^*$ are independent. Consequently, if $G$ is independent of $F$ and has the same distribution, then $H = F^* + (G - G^*)$ has the same distribution as $F$ and satisfies $H|_C = F|_C$.

**Proof.** Due to the orthogonal projection, $(F^*, F - F^*)$ is a Gaussian vector whose components $F^*$ and $F - F^*$ are uncorrelated and thus independent.

A conditional sample of $F$ given $F|_C = \varphi$ can thus be obtained with $\varphi^* + G - G^*$ where the **kriging component** $\varphi^*$ is defined by $\varphi^*(x) = \sum_{c \in C} \lambda_c(x) \varphi(c), x \in \Omega$ and where the independent component $G - G^*$ is called the **innovation component**. To do so one needs to compute the kriging coefficients $(\lambda_c(x))_{c \in C}$. Since $F^*(x) - F(x)$ is orthogonal to $F(c)$, we obtain that $\mathbb{E}(F^*(x)F(c)) = \mathbb{E}(F(x)F(c)) = \Gamma(x, c)$, and using (1) to substitute $F^*(x)$ shows that $(\lambda_c(x))_{c \in C}$ is a solution of the following $|C| \times |C|$ linear system

$$\forall c \in C, \sum_{d \in C} \lambda_d(g) \Gamma(d, c) = \Gamma(x, c).$$

In the following, we will assume that the matrix $\Gamma|_{C \times C}$ is invertible (which turned out to be true in all our inpainting experiments) so that this system admits a unique solution. If we denote by $\Lambda = (\lambda_c(x))_{x \in \Omega, c \in C}$ the $|\Omega| \times |C|$ matrix of all kriging coefficients, the previous system rewrites $\Lambda = \Gamma_{|C \times C}^{-1}$, where $\Gamma_{|C \times C}$ and $\Gamma_{|C \times C}$ are the restrictions of the covariance matrix $\Gamma$. Finally, writing the Gaussian vectors in column, sampling $F$ given $F|_C = \varphi$ amounts to sample

$$\Lambda \varphi + G - \Lambda G|_C, \quad \text{where} \quad \Lambda = \Gamma_{|C \times C}^{-1}$$

and $G$ has the same distribution as $F$ and is independent of $F$. The required number of operations is $O(|C|^3 + |\Omega||C|^2)$ in addition to the simulation cost for $G$.

### 3. MICROTEXTURE INPAINTING

In this section, we use the Gaussian conditional sampling scheme describe in the previous section to perform microtexture inpainting. The input is a masked texture $u : \Omega \to \mathbb{R}$ the values of which are unknown on the mask domain $M \subset \Omega$ and we want to sample these values based on the unmasked values in $M^c$. This amounts to adopt a Gaussian texture model $U$ on $\Omega$, and to sample it on $M$ given the surrounding values. This raises two main difficulties: 1) One must estimate the Gaussian model for $U$ from the masked image $u$; 2) The computation of the kriging components must be practical, without a cubic complexity in $|M^c|^3$.

#### 3.1. Gaussian Texture Model Estimation

The Gaussian model that we use for conditional simulation is given by an asymptotic discrete spot noise (ADSN), which is the convolution of a finitely-supported function $h : \mathbb{Z}^2 \to \mathbb{R}$ with a normalized Gaussian white noise on $\mathbb{Z}^2$. This ADSN has zero mean and covariance $\Gamma(x, y) = h \ast \hat{h}(y - x)$, where $\hat{h}(x) = h(-x)$, which only depends on $x - y$ since the ADSN is stationary. In order to estimate an ADSN from the masked exemplar, we consider a subimage $v$ of $u$ defined on a subdomain $\omega \subset M^c$. Following [4, 5], we compute the empirical mean $\bar{v}$ and the normalized spot $t_v = \frac{1}{|v|}(v - \bar{v})$ that we extend by zero-padding to $\mathbb{Z}^2$. The ADSN is then simply given by the convolution $t_v \ast W$ where $W$ is a normalized Gaussian white noise on $\mathbb{Z}^2$. We must choose $\omega$ large enough to capture a representative piece of the desired texture.
3.2. Textural Inpainting Algorithm

Ideally one would like to consider the conditioning set $C = M^c$, but because of the $O(|C|^3)$ complexity, we restrict $C$ to a band of width $w$ pixels along the outside boundary of $M$. Let us mention that in the case where $U$ has a Markov property with a neighborhood of width $w$, this restriction of the conditioning set would still give a perfect conditional sampling given all the values in $M^c$. We then sample $U$ on $M$ given $U|_C = u|_C$ with the algorithm explained in Section 2. However, we exploit the stationarity of our Gaussian model to compute efficiently the kriging component. Indeed, thanks to stationarity, a matrix-vector multiplication with the covariance matrix $\Gamma$ is a convolution by the image $c_v = t_v * \tilde{t}_v$. Hence to compute the kriging component $\phi^* = \Lambda\phi = \Gamma_{B_{|C^c}C}(\Gamma^{-1}_{|C^c}C\phi)$ we first compute $\psi = \Gamma^{-1}_{|C^c}C\phi$ that we extend by zero-padding to get a function $\Psi$ on $\Omega$, and then compute $\phi^* = c_v * \Psi$. The whole process is summarized in Algorithm 1.

**Algorithm 1: Microtexture inpainting**

Inputs: texture $u$ known outside the mask $M \subset \Omega$, conditioning points $C \subset M^c$
- From a restriction $v$ of $u$ to a subdomain $\omega \subset M^c$, compute the mean $\bar{v}$, $t_v = \frac{1}{|\omega|} (v - \bar{v}) \in \mathbb{R}^{|\Omega|}$ (extended by 0 outside $\omega$), and $c_v = t_v * \tilde{t}_v$
- Store the matrix $\Gamma_{|C^c}C(c, d) = c_v(e - d)$, $c, d \in \Omega$
- Draw an ADSN sample $U = t_u * W$
- Compute $\psi_1 = \Gamma^{-1}_{|C^c}C(u|_C - \bar{v})$ and $\psi_2 = \Gamma^{-1}_{|C^c}C U|_C$
- Extend $\psi_1$ and $\psi_2$ by zero-padding to get $\Psi_1$ and $\Psi_2$
- Compute the kriging components $(u - \bar{v})^* = c_v * \Psi_1$ and $U^* = c_v * \Psi_2$
- Output: $\bar{v} + (u - \bar{v})^* + U - U^*$

Considering the complexity, the linear systems $\Gamma^{-1}_{|C^c}C\phi$ can be solved with standard numerical techniques in $O(|C|^3)$. Then, the kriging components can be obtained with the DFT in $O(|\Omega| \log |\Omega|)$, which is the complexity to generate the ADSN sample $U$. Thus, the number of conditioning points must stay low for the conditional simulation to be feasible.

**Extension to Color Images** Algorithm 1 can be straightforwardly adapted to color textures. Indeed, the simulation of the ADSN model does not change, provided that we convolve the three channels by the same Gaussian white noise [4]. Besides, the kriging framework can be adapted by taking the conditional expectation with respect to the three channels of conditional values. In the end, it amounts to compute a larger matrix $\Gamma$ containing all the cross-correlations between the three channels. Then, the linear systems $\Gamma^{-1}_{|C^c}C\phi$ are of size $3|C| \times 3|C|$, and the convolution involved in $\Gamma_{|C^c}C\phi$ is now a convolution by a matrix-valued function.

4. RESULTS

First, in Fig. 1 we validate the methodology based on conditional simulation by inpainting a masked texture with a previously estimated Gaussian model. One can see that, in contrast to PDE-based methods, this algorithm is able to fill holes in textures of any size and shape. In particular, the kriging component extends edge-like structures while local details are added with the innovation component, in a way that globally preserves the texture covariance.

Let us also discuss on this example the influence of the parameter $w$ which is the width of the conditioning points. Intuitively, this value must not be too small in order to get a sufficient summary of the surrounding context based on the conditioning set $C$, but also not too large to keep a reasonable computational time (see Subsection 3.2). In our experiments we chose the value $w = 3$. We observed in other experiments (not shown here) that this is the minimal value ensuring a good line preservation. As already said, with a Gaussian Markov model, restricting the conditioning points to the mask border would not affect the conditional distributions. Let us mention however that the ADSN models that we use here do not satisfy the Markov property. But the visual impact of this conditioning restriction seems visually harmless.

Now, let us present microtexture inpainting with a Gaussian model estimated on the masked exemplar. In Fig. 2, we inpaint a circular hole in the left part of the texture while learning the Gaussian model on the right part. One can observe that the algorithm produces a satisfying result on the two first examples because the hole cannot be perceived in...
the inpainted texture and it may be difficult to distinguish the original and the output. The result is however slightly less satisfying on the third example because the corresponding Gaussian model exhibits interference patterns which are not visible in the original.

In Fig. 3, we compare our method to the one of [12] based on the total variation (TV), and to the example-based method of [23]. As expected, textures are completely lost with the TV-based model. In contrast, the result of [23] is more precise, but the border of the inpainted domain is still visible. It is also the case with our method, which produces a more textured pattern.

Our Gaussian inpainting algorithm is also compared to Efros-Leung algorithm in Fig. 4. As one can see, the frequency content of the estimated Gaussian model is precisely reproduced in the filled zone. In contrast, Efros-Leung algorithm fails to reproduce the long-range correlations which exist in this microtexture. In Fig. 4 we also show the kriging component $\tilde{v}+(u-\tilde{v})^+$ involved in the conditional simulation; one can see that this component contains long-range diagonal correlations that are characteristic of this wood texture.

To conclude, this paper shows that Gaussian conditional sampling is an interesting algorithm to fill large holes in microtextures provided that the Gaussian model has been properly estimated. For small to medium holes, the proposed algorithm is very efficient (the results of Figures 2, 3 were obtained in about 1 sec.). Its main limitation is that the Gaussian model is only valid for a limited class of texture. Still, extending Gaussian conditional method to non-stationary Gaussian random fields could be of potential interest to inpaint non-texture images.