Random phase fields and Gaussian fields for image sharpness assessment and fast texture synthesis

Arthur Leclaire

Université Paris Descartes
Laboratoire MAP5 UMR CNRS 8145
École doctorale 386 : Sciences Mathématiques de Paris Centre

Thèse

présentée par

Arthur LECLAIRE

pour obtenir le grade de

Docteur de l'Université Paris Descartes
Spécialité : Mathématiques appliquées

Champs à Phase Aléatoire et Champs Gaussiens pour la Mesure de Netteté d'Images et la Synthèse Rapide de Textures

Random Phase Fields and Gaussian Fields for Image Sharpness Assessment and Fast Texture Synthesis

Soutenue le 26 juin 2015 devant le jury composé de

Andrés Almansa	CNRS, Telecom Paristech	Rapporteur
Pierre Chainais	École Centrale de Lille	Examinateur
Agnès Desolneux	CNRS, ENS Cachan	Examinatrice
Bruno Galerne	Université Paris Descartes	Examinateur
Lionel Moisan	Université Paris Descartes	Directeur de thèse
Gabriel Peyré	CNRS, Université Paris-Dauphine	Rapporteur
Frédéric Richard	Université Aix-Marseille	Président du jury
Carola-Bibiane Schönlieb	University of Cambridge (UK)	Examinatrice

Résumé

Dans cette thèse, on étudie la structuration des phases de la transformée de Fourier d'images naturelles, ce qui, du point de vue applicatif, débouche sur plusieurs mesures de netteté ainsi que sur des algorithmes rapides pour la synthèse de texture par l'exemple.

Le Chapitre 2 présente dans un cadre unifié plusieurs modèles de champs aléatoires, notamment les champs *spot noise* et champs gaussiens, en prêtant une attention particulière aux représentations fréquentielles de ces champs aléatoires.

Le Chapitre 3 détaille l'utilisation des champs à phase aléatoire à la synthèse de textures peu structurées (microtextures). On montre qu'une microtexture peut être résumée en une image de petite taille s'intégrant à un algorithme de synthèse très rapide et flexible via le modèle *spot noise*. Aussi on propose un algorithme de désocclusion de zones texturales uniformes basé sur la simulation gaussienne conditionnelle.

Le Chapitre 4 présente trois mesures de cohérence globale des phases de la transformée de Fourier. Après une étude théorique et pratique établissant leur lien avec la netteté d'image, on propose un algorithme de déflouage aveugle basé sur l'optimisation stochastique de ces indices.

Enfin, dans le Chapitre 5, après une discussion sur l'analyse et la synthèse directe de l'information de phase, on propose deux modèles de textures à phases cohérentes qui permettent la synthèse de textures plus structurées tout en conservant quelques garanties mathématiques simples.

Abstract

This thesis deals with the Fourier phase structure of natural images, and addresses no-reference sharpness assessment and fast texture synthesis by example.

In Chapter 2, we present several models of random fields in a unified framework, like the spot noise model and the Gaussian model, with particular attention to the spectral representation of these random fields.

In Chapter 3, random phase models are used to perform by-example synthesis of microtextures (textures with no salient features). We show that a microtexture can be summarized by a small image that can be used for fast and flexible synthesis based on the spot noise model. Besides, we address microtexture inpainting through the use of Gaussian conditional simulation.

In Chapter 4, we present three measures of the global Fourier phase coherence. Their link with the image sharpness is established based on a theoretical and practical study. We then derive a stochastic optimization scheme for these indices, which leads to a blind deblurring algorithm.

Finally, in Chapter 5, after discussing the possibility of direct phase analysis or synthesis, we propose two non random phase texture models which allow for synthesis of more structured textures and still have simple mathematical guarantees.

Remerciements

Quelle chance ai-je eue d'avoir été aussi bien entouré !

Mes plus chaleureux remerciements vont tout d'abord à Lionel Moisan et Bruno Galerne qui ont encadré ma thèse. Dès le départ, j'ai été heureux de pouvoir explorer un très beau sujet de thèse, écrit pratiquement sur mesure, au carrefour de l'analyse de Fourier, des probabilités, et du traitement d'images. La richesse de votre discussion et les passionnantes questions que vous m'avez soumises m'ont beaucoup aidé à faire vivre ce sujet. Lionel, grâce à un exceptionnel recul scientifique, tu as su m'accompagner dans les deux thématiques de cette thèse que sont la modélisation de textures et la mesure de netteté, tout en me laissant une liberté intellectuelle et matérielle sans laquelle je n'aurais pu m'épanouir en tant qu'apprenti enseignant-chercheur. Bruno, c'est au cours d'un de tes exposés que ma curiosité a été éveillée sur la synthèse de textures, et il est clair que ma thèse a beaucoup profité de ton expertise dans ce domaine; ta disponibilité et ton optimisme (à toute épreuve !) m'ont permis de retrouver la motivation dans les moments difficiles. Au-delà de votre apport scientifique, j'ai été profondément touché, dans le travail et dans la vie, par votre ouverture d'esprit et votre humanité.

J'adresse aussi mes plus sincères remerciements à tous les membres de mon jury de thèse, à qui je porte une grande admiration scientifique. Je remercie particulièrement Andrés Almansa et Gabriel Peyré qui ont immédiatement accepté le rôle de rapporteur, et qui ont efficacement rempli cette tâche difficile. Leurs travaux en inpainting, restauration et modélisation de textures ont inspiré certaines parties de cette thèse. Je remercie aussi les examinateurs Pierre Chainais, Agnès Desolneux, Frédéric Richard, Carola-Bibiane Schönlieb pour les questions et suggestions pertinentes qu'ils m'ont adressées avant, et pendant la soutenance. Au-delà des échos de vos travaux que vous pourrez trouver dans cette thèse (avec la super-résolution de textures, la poursuite du texton idéal, la modélisation gaussienne de textures, et l'inpainting), j'espère que vos suggestions déboucheront dans les mois suivants sur des pistes de travail communes.

Je garderai un très bon souvenir des quatre années que j'ai passées à l'Université Paris Descartes. Je souhaite remercier les deux directrices successives du laboratoire MAP5, Annie Raoult et Fabienne Comte, pour l'écoute et le soin très particulier qu'elles portent aux doctorants. Je remercie aussi la directrice de l'UFR Math-Info, Christine Graffigne, de m'avoir attribué le monitorat en L1 que j'ai été ravi d'accomplir sous la bienveillance de François Patte et Florent Benaych-Georges. Les aspects matériels et techniques de la vie doctorale au MAP5 sont grandement facilités par Christophe Castellani, Marie-Hélène Gbaguidi, Michel Guillemot, Voehni Kheng, Marie-Antoinette Lagarde, Clémence Misseboukpo, et Isabelle Valéro, que je remercie pour leur efficacité et leur gentillesse. Je n'oublie pas non plus les personnels du LIPADE voisin, Patricia Bous, Arnaud Meunier, et Thierry Raedersdorff, avec qui j'ai eu le plaisir de partager quelques discussions informatiques et autres carrés de chocolat.

En septembre 2014, j'ai été très bien accueilli à l'École Normale Supérieure de Cachan, et pour cela je remercie la direction du département de mathématiques Alain Trouvé et Claudine Picaronny, la direction du CMLA Nicolas Vayatis, Florian de Vuyst, Véronique Almadovar, et les gestionnaires Carine Saint-Prix, Micheline Brunetti, et Virginie Pauchont. Je remercie chaleureusement Laurent Desvillettes et Frédéric Pascal d'avoir facilité mon intégration à cette équipe pédagogique dynamique, ainsi que mon prédécesseur Cyrille Hériveaux dont l'aide a été précieuse pour assumer cette rentrée chargée. Le quotidien de fin de thèse a été très éclairci grâce à la proximité de mes joyeux acolytes Sandrine Dallaporta et Tuong-Huy Nguyen (promis, après ma soutenance, on fait un couscous ensemble !). Je remercie aussi Nicolas Pajor dont l'aide informatique m'a été précieuse à la fois comme étudiant et comme personnel, et dont la gentillesse et la bonne humeur rayonnent toujours sur le premier étage du bâtiment Cournot. Enfin, je remercie les autres chercheurs de l'équipe de traitement d'images pour leur accueil chaleureux, notamment Gabriele Facciolo, Enric Meinhardt-Llopis et Jean-Michel Morel.

J'ai été beaucoup enrichi, scientifiquement et personnellement, par toutes les rencontres que j'ai pu faire dans le milieu de la recherche. En participant au projet MATAIM, j'ai eu le privilège de côtoyer Anne Estrade et Hermine Biermé (dont la culture en champs aléatoires a irrigué mon travail) ainsi que Yann Gousseau (dont les travaux en modélisation de textures ont inspiré cette thèse à plusieurs titres). Au cours des réunions MISS, j'ai aussi pu faire connaissance avec les chercheurs du CNES Gwendoline Blanchet, Christophe Latry et Bernard Rougé dont l'expérience en imagerie satellitaire m'a apporté un point de vue différent sur mes travaux en netteté d'images. J'ai aussi eu la chance de participer à plusieurs colloques et conférences, dans lesquels j'ai pu notamment rencontrer Jean-François Aujol (que je remercie pour ses encouragements bienveillants), Cécile Louchet (qui en grande soeur de thèse a su soulager mes angoisses), et Mila Nikolova (qui a reçu nos travaux en netteté et déflouage de façon très enthousiaste). Je remercie aussi Gaël Mahé du LIPADE de m'avoir apporté sa vision de traiteur de signaux, et de continuer l'exploration de la cohérence de phase du signal audio. Et enfin, même si je ne peux tous les citer, ma plus sincère gratitude va aux enseignants-chercheurs que j'ai côtoyés au MAP5, Flora Alarcon, Avner Bar-Hen, Olivier Bouaziz, Sylvain Durand, Jean-Claude Fort, Servane Gey, Joan Glaunès, Georges Koepfler, Raphaël Lachièze-Rey, Nicolas Meunier, Nicole Rigal et Adeline Samson, avec une pensée particulière pour Antoine Chambaz (qui outre les carrés de chocolat, partage aussi son amour de la musique), Jérôme Dedecker (incollable sur les limites du théorème central-limite), Julie Delon et Edoardo Provenzi (qui m'ont tous deux aidé à vérifier l'équation 3 bavards ensemble $= -1$ heure de travail + beaucoup d'amitié).

Au cours de ce doctorat, j'ai été maintes fois contaminé par la bonne humeur virulente des Thésards et Jeunes Docteurs, dont certains à la nuit tombée se transforment en Très Joyeux Drôles d'oiseaux. Qu'ils soient à Descartes, à Télécom, à Cachan, ou ailleurs, je remercie Alasdair, Alkéos, Angelina, Anne-Claire & Christophe, Ardo, Aurélie, Axel, Baptiste, Carlo, Cecilia, Charlotte L., Claire, Diarra, Fabien, Irène, Julie, Julien, Kevin, Lara, Laureen, Laurent, Léon, Maël, Marc, Mariella, Mario, Maxime, Maud, Nicola, Nicolas C., Nicolas M., Oriel, Rachel, Re-

becca, Ronan, Samuel, Samy, Sira, The-Minh, Thibaud, et Thomas. J'ai tissé des liens privilégiés avec ceux issus de la même filiation scientifique, et en particulier ma cousine/nièce Fanny (vive les pauses macarons au soleil !), mon cousin Guillaume (qui, malgré les apparences, n'est pas tombé dans la potion magique étant petit), ma petite soeur Anne-Sophie (dont les histoires très détaillées et le rire aigu me rappellent l'insouciance de la jeunesse), et mon presque frère jumeau Rémy (sans qui je n'aurais pu imaginer le plaisir d'un saucisson balnéaire sur fond de Joe Dassin). Je remercie aussi Pierre & Anaïs, imbattables pour le remontage de moral d'urgence à force de sorties piscine, de pauses ensoleillées, et d'apéritifs improvisés; j'espère pouvoir un jour vous rendre la pareille. Le groupe des doctorants de Paris Descartes trouve sa cohésion au travers de son fameux GTTJD, géré de mains de maîtres par Charlotte, Christèle, Gwennaëlle et Jean (et son alter-ego fictionnel José Bigre, à qui je souhaite bon courage pour la recherche de poste), qui outre ses nombreux exposés très intéressants, propose des karaokés des plus enflammés. Et je n'oublie pas une très chère ex-organisatrice : Gaëlle, tu as vogué très vite vers de glorieux horizons (quelle idée de finir sa thèse en trois ans ?!?); nos pauses chocolat me manquent beaucoup, mais mon admiration et mon amitié perdurent.

Je dois aussi beaucoup à tous les amis qui m'ont soutenu. Ceux du lycée : Alexis, Cindy, Florian, Grégoire, Jérémy, Kayvan, Sarah; sans oublier Sébastien & Marjorie (et la petite Cléa !) qui, malgré la distance, sont toujours là quand il faut. Un grand merci aussi aux amis rencontrés dans la suite de mes études. D'abord à Jehanne & Élie, Marie & Nicolas T., Nicolas B., et Sary & Marine, pour leur amitié si solide et leur grande générosité. Puis le festif G4 : Camille, Émilie & Romain, Kevin, Séverine et ses membres d'honneur Laure & Nicolas, Nina & Jérémie, Stéphane & Ombline, Timon & Tiphaine (et les petits Élise et Benoît !), Pierre, et Vong & Vonguette; la bonne humeur propagée à cet étage nous poursuit inlassablement ! J'ai partagé avec tous ceux-ci une certaine passion pour les sciences, nourrie par d'excellents professeurs, motivés et motivants. Pour ne citer que ceux qui ont été véritablement déterminants dans ma formation à Nancy, je remercie du fond du coeur Marie-Laure Duhaut, Daniel Noircler, Pierrette Visot, et à l'université, Anne de Roton et le regretté Patrick Sargos, qui m'ont transmis le goût de l'analyse de Fourier et ont fait vivre mes ambitions.

Enfin, ma réussite scolaire n'aurait pas été possible sans l'appui inconditionnel de ma famille. Mes parents Dominique, Jean-Pierre et Marie-Odile ont su me convaincre très rapidement de l'importance de suivre des études épanouissantes, et m'ont donné tous les moyens de m'accomplir; qu'aurais-je pu espérer de plus ? Je tiens à remercier aussi mes oncles, tantes, cousins et cousines, et notamment Francis qui a éveillé ma curiosité mathématique au plus jeune âge, et André dont j'apprécie la discussion à chaque déjeuner dominical ! Un grand merci aussi à ma belle-famille et en particulier Ouiza et Nabyl qui m'ont permis de passer mon premier nouvel an sous le soleil méditerranéen. Et bien sûr, mille mercis à Sonia, dont l'amour adoucit ma vie chaque jour. Je termine par une pensée émue pour mes grands-parents qui m'ont transmis une inextinguible soif de connaissances, et dont l'affection et la fierté m'ont poussé beaucoup plus loin que ce que je pouvais imaginer au départ.

Contents

Notations

$|E|$ is the cardinality of the set E.
If $E, F \subset \mathbb{R}^d$, $E - F = \{\ x - y\ ;\ x \in E, y \in F\ \}$.
$\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$.
If $z \in \mathbb{C}^*$, then $\arg(z) \in \mathbb{T}$ denotes the argument of z.
A^T is the transpose matrix of A.
$\bar{A}$ is the complex conjugate of A.
A^* is the adjoint of A ($A^* = \bar{A}^T$).
$A^{1/2}$ is the square root of the symmetrical non-negative matrix A
If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^k$, $\mathbf{x}.\mathbf{y} = \mathbf{x}^T\mathbf{y}$
$\|v\|_p$ is the ℓ^p norm of a complex vector v
$\|v\|$ is the euclidian norm of the complex vector v ($\|v\| = \|v\|_2$)
For $p \in [1, \infty[$, $\ell^p(\mathbb{Z}^2, \mathbb{R}^{m\times n})$ is the Banach space of functions $f : \mathbb{Z}^2 \to \mathbb{R}^{m\times n}$ such that

$$\|f\|_p^p = \sum_{\mathbf{x}\in\mathbb{Z}^2} \|f(\mathbf{x})\|^p < \infty\ .$$

If $p = \infty$, $\ell^\infty(\mathbb{Z}^2, \mathbb{R}^{m\times n})$ is the Banach space of functions $f : \mathbb{Z}^2 \to \mathbb{R}^{m\times n}$ such that

$$\|f\|_\infty = \sup_{\mathbf{x}\in\mathbb{Z}^2} \|f(\mathbf{x})\| < \infty\ .$$

$\tilde{f}(\mathbf{x}) = f(-\mathbf{x})$.
$\mathcal{H}_d$ is a set of (complex) Hermitian matrices of size $d \times d$ ($A^* = A$).
$\mathcal{H}_d^{\geqslant 0}$ is the set of non-negative Hermitian matrices of size $d \times d$.
$\mathcal{U}(E)$ is the uniform probability distribution on E.
$\mathcal{P}(\lambda)$ is the Poisson distribution with intensity λ,

$$\mathcal{P}(\lambda) = e^{-\lambda} \sum_{n\geqslant 0} \frac{\lambda^n}{n!}\,\delta_n\ .$$

$\mathbb{E}(X)$ is the expectation of the matrix-valued random variable X.
$\mathrm{Var}(X)$ is the variance of the real random variable X.
If X, Y are two complex random vectors of same dimension,

$$\mathrm{Cov}(X, Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))^*)$$

and by extension, $\mathrm{Cov}(X) = \mathrm{Cov}(X, X)$.
We shall denote the Nyquist frequencies corresponding to a discrete rectangle of size $M \times N$ by

$$\boldsymbol{\eta}_x = (-M/2, 0) \quad \boldsymbol{\eta}_y = (0, -N/2) \quad \boldsymbol{\eta}_{xy} = (-M/2, -N/2)\ .$$

Chapter 1

Introduction

Contents

This thesis would certainly not have been written, had I any detectable drawing talent. Being unable to reproduce what you see in your everyday life, however beautiful it may be, encourages to develop a certain fascination for automatic image synthesis.

If we admit that images are expected to reflect the reality in the most faithful manner, several issues are at stake: the design of objects contours in the right proportions, the precise reproduction of the human or animal morphology, the respect of geometrical constraints like perspective. But an image realism also highly depends on the credibility of the textural content with which objects are covered, as one can see in Fig. 1.1. In a certain sense, textures may not be our first concern when representing the spatial organization of a scene, but they must seem realistic, especially if one decides to focus on them. Finding an automatic way to produce a (possibly large) image that resembles a (possibly small) given texture sample is a problem known as by-example texture synthesis. This thesis was motivated in great part by texture modelling and texture synthesis, and is mostly meant to study the role of Fourier phase (that is, the argument of the Fourier transform) in texture modelling and image quality assessment.

1.1 Stochastic Texture Models

The ambiguous concept of texture

One main difficulty of texture modelling is that the set of texture images is not precisely defined. According to the lexical definition [Simpson & Weiner 1989], this word has several meanings all related to the composition of a piece of material (fabric, rock, food, etc). The meaning that is relevant here is inherited from fine arts: a texture is the graphical representation of an object surface. The consideration of one single object implicitly assumes some kind of homogeneity.

Figure 1.1: **The importance of texture for image realism**. One can see two versions of the same virtual character (copyright Nintendo), the first one (left) designed in 1998, and the second one (right) designed in 2008. One main difference between these two images is the rendering of texture, which makes the new image much more realistic.

Several precised definitions have been given in the image processing literature. In [Haralick 1979], a texture is seen as "organized area phenomena", whereas it is defined in [Cross & Jain 1983] as a "stochastic, possibly periodic, two-dimensional image field". The authors of [Chellappa & Kashyap 1985] suggest to oppose "deterministic" and "stochastic" textures: a deterministic texture is characterized by a set of "primitives" (i.e. micro-objects) whose arrangement obeys a (possibly random) placement rule, whereas a stochastic texture is only described by statistical characteristics. One can argue that these abstract concepts correspond more to an early attempt of texture modelling than to a clear delimitation of the set of texture images. Still, all these authors expressed the importance of randomness in texture modelling, which agrees with the lexical definition because an object formation can be governed by deterministic and stochastic rules.

More recently, the authors of [Wei *et al.* 2009] adopted a very large definition of texture (which is shared by many researchers in the graphics community) as "an image containing repeating patterns" [with] "a certain amount of randomness". As illustrated in Fig. 1.2, this definition is larger than the lexical one. But its success is undeniable, maybe because it encompasses many (if not all) examples of textures shown in research articles about texture synthesis, but also because this definition still allows to think the concept of texture by negation of geometrical content (a point of view that was precised for example in [Meyer 2001] and [Aujol *et al.* 2005]). And this negative definition complies well with a two-step image synthesis scheme that first positions graphical objects in the image and later covers them with repeated patterns.

A related question is to ask which mathematical objects will be involved in

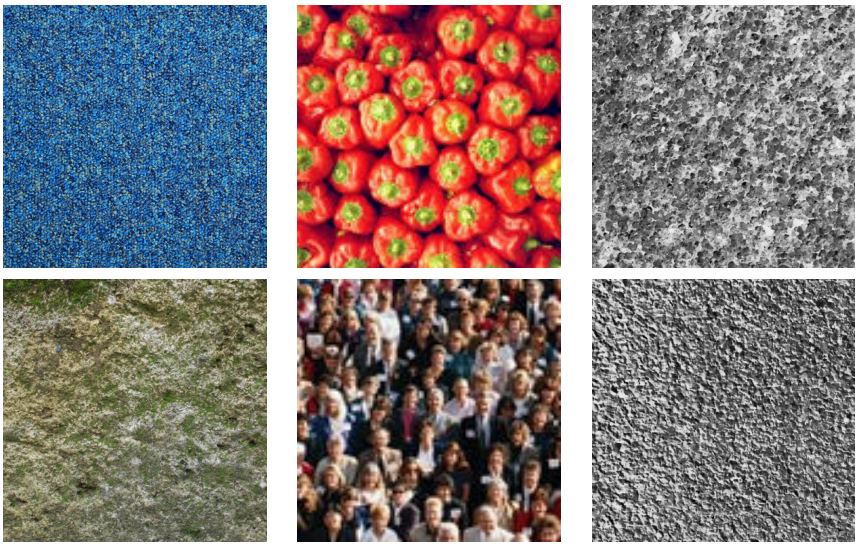


Figure 1.2: **Examples of natural textures.** Even if the term "texture" usually refers to the graphical representation of an object surface (first column), it can more generally refer to images having repeated patterns (second and third column), as suggested by [Wei *et al.* 2009]. If one can distinguish such repeating patterns, then the period of repetition is sometimes referred to as the texture scale (large scale in the second column, and fine scale in the third column).

the modelling of textures. In this thesis, we will only work with discrete-sampled models, and thus, as suggested by [Cross & Jain 1983], textures will be modelled by random fields F defined on a subdomain Ω of $\mathbb{Z}^2$ (which will often be a finite subdomain of $\mathbb{Z}^2$, or the whole plane $\mathbb{Z}^2$) and with values in $\mathbb{R}^d$ (taking $d = 1$ for gray-level images, and $d = 3$ for color images). In other words, for all $\mathbf{x} \in D$, $F(\mathbf{x})$ is a $\mathbb{R}^d$-valued random variable, and thus, ommiting to write the sample ω of the probability space, we will often describe F as a random function

$$F : D \longrightarrow \mathbb{R}^d \ .$$

A non-trivial problem is then to impose the texture homogeneity through a formal property of the random field F. A common way to do this is to impose stationarity of the random field F with respect to the translations of D. Because of some degenerate cases[1], stationarity is not sufficient to ensure the spatial homogeneity of the realizations. But conversely, (so far) we have not found any reason that prevent us to consider a homogeneous texture image as the realization of some stationary random process (in particular if all pieces of the texture have been obtained by the same physical process). So imposing stationarity is still a relevant constraint, which can have some practical consequences, as we will see later.

[1] think of a randomly positioned segment in a finite domain

From constrained random models to a highly complex reality

In probability theory, the introduction of a random process is generally motivated by a physical or a practical problem. To mention a few classical examples, Brownian motion [Doob 1990] can represent the position of a moving microscopic particle in a viscous fluid, discrete martingales [Doob 1990] can model the gain evolution in several gambling games, and Ising model [Brémaud 1999] can describe the organization of magnetic dipoles in a ferromagnetic material. In these examples, the physical behavior is translated into simple mathematical rules (respectively, prescribed covariance function, null incremental conditional expectation, local conditional distribution) which completely describe the corresponding random process and allow for extensive mathematical study.

The realizations of these random models can already be considered as interesting creations in the sense that they satisfy some criteria which make them very peculiar. For example, almost surely, a realization of the Brownian motion is continuous and nowhere differentiable. Of course, other constructive and non-constructive methods were formerly known to show the existence of such counter-intuitive behavior. But one can still be intrigued by this example which illustrates the remarkable creative power of random methods (in particular if one remembers that a Brownian motion can be seen as a limit of properly renormalized "heads or tails" sequences). Stochastic methods can also generate visual curios since many fractal objects can be obtained with self-similar random fields [Mandelbrot & Van Ness 1968], [Fournier *et al.* 1982].

Nevertheless, the constraints making these models interesting (from a mathematical point of view) also make them very specific and therefore not sufficient (in terms of visual similarity) to account for the large diversity of natural textures. That is why, in texture synthesis by example, in some sense we follow the inverse approach (maybe closer to the statistical paradigm): we choose a very large model with as less constraints as possible, adjust it based on an exemplar image, and then draw a realization of this random model to produce a new texture sample. The main question is thus: can we find a class of random functions able to model a wide variety of texture images for which the model analysis and synthesis is easy, and which still allows for some kind of mathematical analysis?

The statistics of texture perception

In order to precise the model, we now have to discuss the important statistics that rule the human perception of textures. The statistics of a random field F are usually sorted by order, a statistic of order p being any quantity of the form

$$\mathbb{E}\left[\varphi\Big(F(\mathbf{x}_1),\dots,F(\mathbf{x}_p)\Big)\right] \ ,$$

where φ is a measurable function such that the expectation exists and where $\mathbf{x}_1,\dots,\mathbf{x}_p$ are p points in D. For example, knowing all the first-order statistics is equivalent to knowing the marginal distribution, i.e. the probability distribution of one value $F(\mathbf{x})$ of the random field (notice that this distribution does not depend

on $\mathbf{x}$ since F is stationary). Among the statistics of order p, we single out the moments of order p, which are obtained by taking a p-linear function φ in the above expectation. In particular, the moments of order 1 are given by $\mathbb{E}[F(\mathbf{x})] \in \mathbb{R}^d$ and the moments of order 2 are given by

$$\mathbb{E}[F(\mathbf{x})F(\mathbf{y})^T] \in \mathbb{R}^{d\times d} \ .$$

We will see in Chapter 2 that, if the first-order moments are prescribed, then the moments of order 2 are as informative as the energy spectrum (which in a periodic discrete setting is represented by the expected square modulus of the discrete Fourier transform). Sometimes, one does not consider the raw values of the random field but rather the values of a linear transform (for example the Fourier transform, or a wavelet transform); one can then define the p-th order distribution after this linear transformation as above.

The first-order distribution of a texture is perceptually important because it defines the color distribution. Also, the second-order moments are important because they encode the frequency content of the texture. But are these statistics sufficient? Can we state a global rule for texture perception? These problems were dealt with by Julesz in a famous series of articles in which he exhibited several examples and counter-examples illustrating the complexity of the two previous questions. In the article [Julesz 1962], he formulated the conjecture that two textures with identical second-order statistics are pre-attentively indistinguishable, meaning that they cannot be discriminated without scrutiny. Indeed, for a wide class of textures, these first-order statistics are sufficient for discrimination. However, he showed in [Julesz 1981] the existence of counterexamples, and two of them are reproduced in Fig. 1.3. More precisely, he exhibited

1. pre-attentively indistinguishable textures with same second-order statistics but different higher-order statistics,

2. pre-attentively distinguishable textures with same second-order statistics but exhibiting different local conspicuous features based on closure, connectivity, granularity.

With the latter example, he thus refuted his own first conjecture. The former example illustrates the difficulty of the human visual system (HVS) to compute high-order statistics. The two textures of this former example can actually be attentively discriminated by paying particular attention to local details. Beyond the refutation of his first conjecture, the experiments of this article thus led him to two important conclusions. The first one is the clear separation between two texture perception systems: a pre-attentive one that uses globally summed statistics of low order, and an attentive one that takes profit of extended local analysis of features. The second conclusion is that the perception of texture is mostly based on the computation of first-order statistics on the textures *or* on local features which are called textons. For example, the HVS is sensitive to the density of elongated blobs, of terminations, of crosses, etc. Julesz thus introduced the word texton referring to any textural unit that allows for pre-attentive texture discrimination. He then

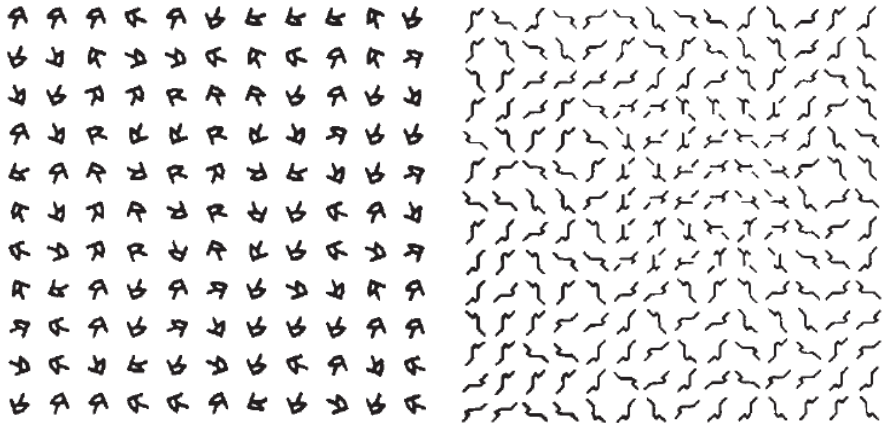


Figure 1.3: **Counterexamples to the first Julesz conjecture**. These texture images are borrowed from [Julesz 1981, Fig. 2 and Fig. 5]. The left image is made of two pre-attentively indistinguishable textures with same second-order statistics but different higher-order statistics. The right image is made of two pre-attentively distinguishable textures with same second-order statistics; the pre-attentive discrimination is based on the density of terminators (i.e. end of lines).

formulated his modified conjecture: the pre-attentive system for texture perception cannot compute global statistics of order > 2.

So must the Fourier modulus be thrown out?

However, the fact that texture perception is focused on first-order distribution of textons must not prevent us from exploring texture models based on second-order statistics. Indeed, on the one hand, since Julesz did not provide an exhaustive and operative description of textons, it is not straightforward to derive a clear texture model that precisely respects the density of textons. On the other hand, considering the complexity of texture modelling, it is legitimate to exploit every available mathematical tool that could build a texture model whose success would be measured not only by the adequation with the theory of texture perception, but also according to the multiple applicative issues it may solve (should it be texture classification, texture analysis, fast texture synthesis, etc).

In particular, in this thesis, we will thoroughly study texture models based on first and second-order moments. As it is well-known (and sometimes referred to as Wiener-Khintchine theorem [Yaglom 1987]), the energy spectrum is the Fourier transform of the autocorrelation, so that prescribing the second-order moments is equivalent to impose the expected squared modulus of the Fourier transform. Even if this link between the autocorrelation and the frequency content is already a convincing argument to use second-order moments, one can also argue that the autocorrelation can be used to analyze periodic patterns and to single out relevant texture scales (see [Haralick 1979] and references therein).

1.2 Random Phase Texture Models

Fourier modulus rescued by the success of random phase texture models

Another way to justify the use of second-order moments in texture modelling is to observe the power of frequency-based synthesis algorithms. The richness of frequency-based models was demonstrated for image synthesis by the authors of [Perlin 1985], [Lewis 1984], [Lewis 1989] and [Van Wijk 1991]. In particular, in the first one, Perlin introduces the (slightly overloaded) concept of "noise" as a procedure to compute and display the values of a synthetic texture at prescribed spatial positions. The authors of these four articles notice that a wide variety of textures can be obtained by adjusting the distribution of the energy spectrum. Section 5 of [Lewis 1984] details two methods to set up the output texture: one can directly operate in Fourier domain (spectrum painting); or one can design a spatial kernel h that is convolved by a (possibly white) noise called excitation function (sparse convolution). The author explains in [Lewis 1989] that in the case of sparse convolution, the spectrum of the output texture is the spectrum of the excitation function multiplied by $|\hat{h}|^2$ where $\hat{h}$ is the Fourier transform of h. Thus, spectrum painting and sparse convolution are two ways to prescribe the output texture spectrum. The author of [Van Wijk 1991] proposes the spot noise model which amounts to convolve a spot h with a Poisson point process. He also discusses the interpretation in spectral domain by claiming that a sample of spot noise texture can be obtained via a multiplication by a random phase shift in Fourier domain, with a slight confusion due to the ambiguity of the term "white noise". But the major contribution of van Wijk is to exhibit many textures that can be obtained by simple geometric variations of the spot h.

Several years later, these frequency-based models were rigorously studied in [Galerne *et al.* 2011b]. In this article, a clear distinction is made between the random phase noise (RPN) and the asymptotic discrete spot noise (ADSN). The RPN model is a random field with prescribed Fourier modulus, and maximally random Fourier phase coefficients. The ADSN model is obtained as the limit for larger and larger intensity λ of properly renormalized spot noise functions

$$\mathbf{x} \longmapsto \sum_{i \geqslant 1} h(\mathbf{x} - \mathbf{X}_i) \ ,$$

where $(\mathbf{X}_i)$ is a point process of intensity λ. Based on the classical central-limit theorem, the authors of [Galerne *et al.* 2011b] show that the ADSN is a Gaussian random field. They also clear up the confusion between the RPN and ADSN models by showing that they differ in Fourier domain by a Rayleigh noise on the Fourier modulus. Besides, they detail the application of RPN and ADSN models to by-example texture synthesis, drawing a particular attention to the handling of color distribution. They show that for natural exemplar textures, the realizations of both models are perceptually similar. Since both these models have maximally random Fourier phase (in the entropy sense), the textures that are reproduced by RPN or ADSN in a visually satisfying manner are called random phase textures. They are

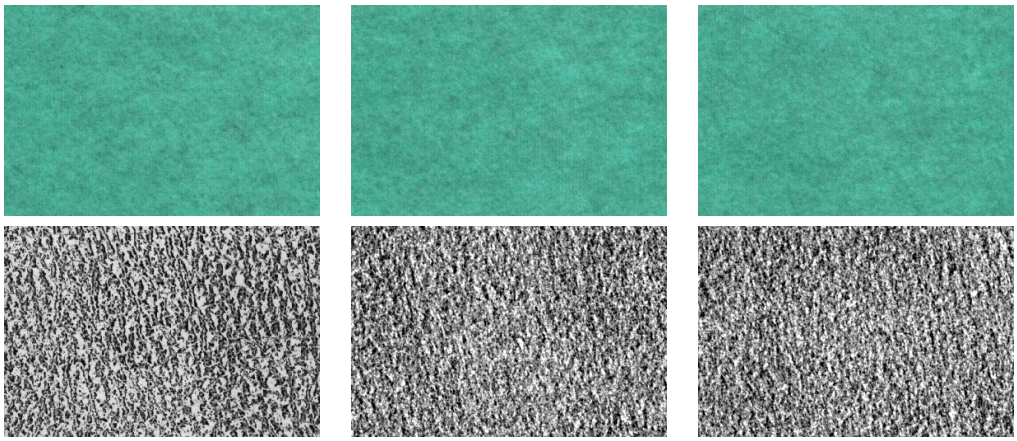


Figure 1.4: **Examples of RPN and ADSN synthesis**. For each row, from left to right, one can see a natural texture, and the synthesis results obtained with RPN and ADSN models. In the first case, the original texture is well-reproduced by RPN or ADSN algorithm and thus one would say that this original texture is a random phase texture. Notice that random phase textures do not have salient features nor large-scale geometric elements.

also sometimes called microtextures because, as one can see in Fig 1.4, their perceptual characteristics are concentrated on fine non-salient details. By opposition, textures with salient geometric details are called macrotextures. Actually, the perceptual unit for microtextures is exactly the Fourier modulus which led the authors of [Desolneux *et al.* 2012] to a practical definition of texton for Gaussian textures. Inspired by the work of [Galerne *et al.* 2011b], the richness of the Gaussian texture model can be explained *a posteriori* by the universality that it inherits from the central-limit theorem.

In Chapter 2, we give a more detailed presentation of the spot noise model and other random phase models. Using the general framework of random fields on $\mathbb{Z}^2$ (with a special attention to spectral representations) allows for comparison with the autoregressive [Chellappa & Kashyap 1985] and moving-average models [Cadzow *et al.* 1993]. The notion of texton for Gaussian texture [Desolneux *et al.* 2012], [Xia *et al.* 2014] is introduced but mainly used to explain the simulation of Gaussian random fields in simple terms; we will not extensively discuss the link between this definition of texton and the perceptual one given by Julesz. In this chapter, we will also recall the definition of the optimal transport distance between texture models [Xia *et al.* 2014] and extend it to the case of texture models on $\mathbb{Z}^2$.

We explain in Chapter 3 how random phase models can be used to perform fast by-example microtexture synthesis. After discussing the estimation procedure for random phase models, we provide several examples of random phase texture synthesis in order to highlight the limits of RPN and ADSN synthesis. The main contribution of this chapter is the synthesis-oriented texton (SOT) which realizes

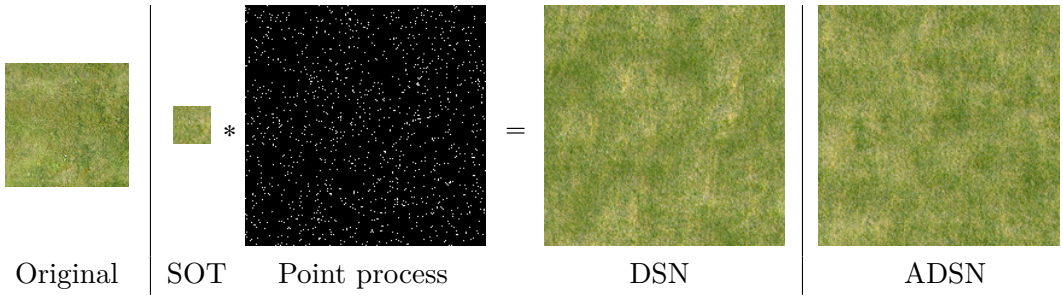


Figure 1.5: **Spot noise synthesis with the synthesis-oriented texton.** The original texture shown on the left can be successfully synthesized via the convolution of the synthesis-oriented texton with a sparse point process. The resulting discrete spot noise is a satisfying approximation of the asymptotic discrete spot noise. The DSN and ADSN associated with the SOT are shown on the right.

a compact summary of a microtexture that can be used as the kernel for spot noise synthesis. We show that many microtextures can be successfully synthesized with a small kernel convolved with a sparse Poisson point process, as illustrated in Fig. 1.5. The SOT thus allows fast and flexible synthesis by direct sampling of the spot noise process. In this chapter, we also propose a solution to the inpainting problem in the case of Gaussian textures by making use of Gaussian conditional simulation [Lantuéjoul 2002].

1.3 The Importance of Phase in Images

What do we miss if we leave out the Fourier phase?

Even if the random phase texture models have a certain success in texture synthesis, their richness is still limited. In particular, a part of Julesz's textons are not directly encoded in the energy spectrum, so that more evolved texture models must bear some kind of structure in the Fourier phase.

Actually, with a naive approach of Fourier transforms, it may seem tempting to consider Fourier modulus more important than Fourier phase for the three following wrong reasons:

1. many basic scholar examples of Fourier transforms are computed with odd or even functions and are thus real;

2. discrete Fourier transforms are often represented through the modulus component and only occasionally accompanied with the phase component;

3. most theorems linking the signal regularity with the decreasing speed at infinity of the Fourier transform only involve its modulus.

But, since a translation of the signal is expressed in Fourier domain by the addition of a linear function to the phase, it is clear that the phase component will play a role in the spatial organization of the signal. As concerns regularity, as illustrated in Fig. 1.6, shuffling the phases of a piecewise smooth function with

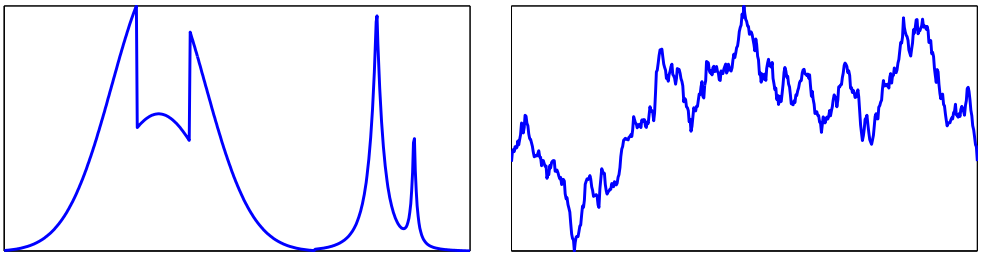


Figure 1.6: **Phase randomization of a bounded variation function.** The signal shown on the right has been obtained by randomizing the phase information of the bounded variation signal shown on the left. Both these signals thus have the same Fourier modulus, but the second one looks much more irregular.

discontinuities leads to a new signal having large oscillations everywhere; and still, this does not modify the Fourier modulus! This means that the regularity expressed by the decreasing speed of Fourier coefficients does not reflect the regularity in the sense of bounded variation functions.

Beyond the context of texture modelling, several works confirmed the importance of Fourier phase in image processing and perception. The most famous one is certainly the paper [Oppenheim & Lim 1981] whose authors bring several arguments justifying the importance of phase, and in particular the one stating that the geometry of an image persists if its phase component is being imposed a completely different modulus information (a random one or one taken from another image). Also, several perceptual studies (for example [Morrone & Burr 1988], [Field *et al.* 2000] and [Hansen & Hess 2006]) confirm the HVS sensitivity to local phase shifts (i.e. phase values of localized Fourier transforms). To precise the results of [Oppenheim & Lim 1981], the authors of [Gegenfurtner *et al.* 2003] measure in a more quantitative manner the persistence of image perception when the phase information is more and more corrupted. Inspired by the work of Morrone and Burr, the author of [Kovesi 2000] (and of the technical report [Kovesi 1999]) suggests to address edge detection by exploiting a concept of local phase coherence. Also, the authors of [Peters & Itti 2008] show that Fourier phase can be used to produce saliency maps, or in other words, to predict the regions of an image that will receive more visual attention from an observer.

Global phase coherence and image quality assessment

Therefore, by showing that the image geometry is mainly encoded in the Fourier phase, these works suggest that the precision of the image geometry may be linked to the coherence of the phase information. And indeed, there has been several attempts to measure the image quality based on phase coherence. Before quoting the corresponding articles, let us briefly discuss the stakes of image quality assessment. We only propose here a very modest evocation of these difficult questions and refer the reader to [Chandler 2013] for a largely more exhaustive review.

Of course, measuring the intrinsic quality of an image is a quite ill-posed problem, since many observers may have different subjective methods and criteria to precise why they are satisfied or not by a particular image. The problem is a bit simpler if one has to assess the quality of an image by comparing it to a supposedly ideal version; in this case, we would speak of full-reference image quality assessment. The problem of full-reference automatic evaluation is already very wide and has led to a large research effort. It is now commonly admitted that mean-square error (MSE) and peak-signal-to-noise ratio (PSNR) are not sufficient to compare image processing algorithms. Thus, new measures have been proposed for full-reference evaluation, like the structural similarity (SSIM) in [Wang *et al.* 2004] and the visual signal-to-noise ratio (VSNR) in [Chandler & Hemami 2007], even though the PSNR values are still widely trusted, in particular when it comes to decide if a denoising or deblurring algorithm is worth publishing or not.

However, in many applications, the ideal version of the image is not available; we then speak of no-reference image quality assessment. In articles about full-reference or no-reference quality assessment, a largely shared goal is to obtain quality indices that are correlated to perceptual scores measured on a group of subjects confronted to a certain image database; this is a way to cope with the subjective nature of image quality. Many questions can be asked, from the most to the least difficult:

- From these two images, which one do you prefer?
- From these two corrupted versions of the same image, which one do you prefer?
- From these two images with different levels of the same artifact, which one do you prefer?

Several interesting no-reference quality measures have been proposed in the literature (see [Chandler 2013] for a detailed list). A standard approach is to evaluate the overall quality by examining the geometric details on which human observers may pay attention. For example, the authors of [Marziliano *et al.* 2004] propose a direct analysis of image edges in order to detect blur and ringing artifacts (with the intent to compare the quality of JPEG2000-compressed images). On the other hand, the authors of [Zhu & Milanfar 2010] notice that blur or noise tend to destroy the anisotropy of certain image patches (like edge patches), and thus they suggest to measure image quality by exploiting the singular values of a matrix containing local gradient values. Of course, if the goal is to correlate with perceptual quality measurements, it also seems relevant to adjust certain parameters of the method based on the mean opinion of human observers; for example, the method of [Ferzli & Karam 2009] compares the edges width to a threshold value called "Just Noticeable Blur" (JNB). The JNB computation is based on a perceptual experiment and represents the value under which blur becomes unnoticeable to a human observer.

Returning to the importance of phase in image analysis, the authors of [Wang & Simoncelli 2004] make the connection between the local phase coherence defined by Kovesi, and the perception of blur. In this article, they translate

the idea of Kovesi using the language of complex wavelet transforms, which allows them to propose a natural two-dimensional analogue of phase coherence based on coarse-to-fine predictions of the phase values. This link between local phase coherence and perceived blur was later exploited in [Hassen *et al.* 2010] to define a no-reference image quality index which is able to penalize a remarkably large class of artifacts (different kinds of blurs, noises, compression artifacts).

Meanwhile, another line of articles proposed to define a notion of global Fourier phase coherence, and used it to address sharpness evaluation. More precisely, the authors of [Blanchet *et al.* 2008] propose to define a notion of global phase coherence (GPC) by measuring how much the regularity of an image is affected by the randomization of the phase information. Given an image u, they define

$$\mathrm{GPC}(u) = -\log_{10} \mathbb{P}(\mathrm{TV}(U) \leqslant \mathrm{TV}(u)) \ ,$$

where U is the RPN associated to u, and where $\mathrm{TV}(u)$ is the discrete periodic total variation (TV) of u. Since we have seen that the phase randomization turns a BV signal into a much more oscillating one, the probability of the righ-hand side is very small, thus motivating the logarithmic scale. Even if a computation of GPC relies on an expensive Monte Carlo simulation, the authors of [Blanchet *et al.* 2008] are able to give practical evidence that the GPC reflected some kind of image quality. This approach is made viable by the heuristic (but non-trivial) remark that the corruption of the phase information is more destructive (with respect to the image geometry) in an image that has low levels of noise and blur. In [Blanchet & Moisan 2012], a variant of GPC is proposed by replacing the RPN field U by its Gaussian analogue (that is, ADSN). The major advantage of this new index, called sharpness index (SI), is that it can be expressed with a closed-form formula (modulo a reasonable Gaussian approximation of $\mathrm{TV}(U)$), and thus can be easily computed (using only four fast Fourier transforms). This considerable simplification made SI appropriate for further applications, as illustrated by the parametric deblurring experiment presented in [Blanchet & Moisan 2012].

In Chapter 4 of this thesis, we will pursue the work initiated by the authors of [Blanchet *et al.* 2008] and [Blanchet & Moisan 2012] on global phase coherence and no-reference image quality assessment. Accordingly, we will mainly consider the concept of image quality[2] by opposition to certain well-identified artifacts that corrupt a particular image. The choice of artifacts that will be considered is driven by our applicative purposes which are mostly focused on image restoration. We will take particular care to blur, noise, ringing and aliasing, which are the four major stakes of image denoising and deblurring, whereas only marginal attention will be paid to compression artifacts. In this chapter, we will first recall the definition of GPC, SI, and present the index S as a further simplification of SI. We will give many analytical and probabilistic properties of these indices, thus illustrating the great advantage to dispose of relatively simple formulae (as opposed to most image quality indices). After that, we provide a more practical study which first proposes many experiments relating phase coherence indices with perceived sharpness, and

[2] or sharpness, which will be considered as synonym.

next shows how the indices SI and S can be used in a blind deblurring algorithm based on a simple stochastic optimization framework.

1.4 Non Random Phase Texture Models

How to account for the phase in a texture model?

Going back to texture modelling and inspired by these works which clarify the role of phase in image perception, it seems that more expressive texture models could be obtained by considering not only the first and second-order moments but also the information contained in the Fourier phase. In particular, one would expect from a phase-dependent synthesis algorithm to be able to reproduce the edge-like structures (which is impossible with RPN and ADSN synthesis algorithms). Such models could also have a significant impact in texture analysis because of the importance of linear structures in medical images: fiber structures are visible all over bone radiographs used for the diagnosis of osteoporosis [Benhamou *et al.* 1994], and spiculated lesions (identified by convergence points of segments) play an important role in the diagnosis of breast cancer [Sampat & Bovik 2003].

Unfortunately, as we observed in the first years of this thesis, the direct processing of the global Fourier phase is a difficult problem. The first section of Chapter 5 summarizes our work in this direction. We show in particular that the phase constraints induced by stationarity are not rigid enough to inspire direct by-example synthesis of the phase component.

How to account for edges in a texture model?

Putting aside the constraints of Fourier representation, one can wonder: which texture models are truly sensitive to edge structures?

Of course, it is possible to design texture models whose corresponding synthesis algorithms naturally produce sharp edges. Let us mention for example the dead leaves model introduced in [Matheron 1968] and later studied in [Bordenave *et al.* 2006], which consists in throwing on a given domain several shapes with different colors that will progressively occlude each other. The occlusion principle used in this model naturally produces sharp edges. If one relaxes this principle with a transparency factor, one gets another texture model called transparent dead leaves [Galerne & Gousseau 2012] whose realizations also exhibit sharp edges. Sharp edges can also be directly produced by random tessellations [Lantuéjoul 2002]. These models have seldom been used to address by-example texture synthesis: still, the method of [Gousseau 2002] indeed emulates dead leaves (by throwing shapes defined as filled connected components of level sets) and thus manages to reproduce sharp edges, but fails to grasp the complex inter-dependences of geometric structures that exist in structured textures.

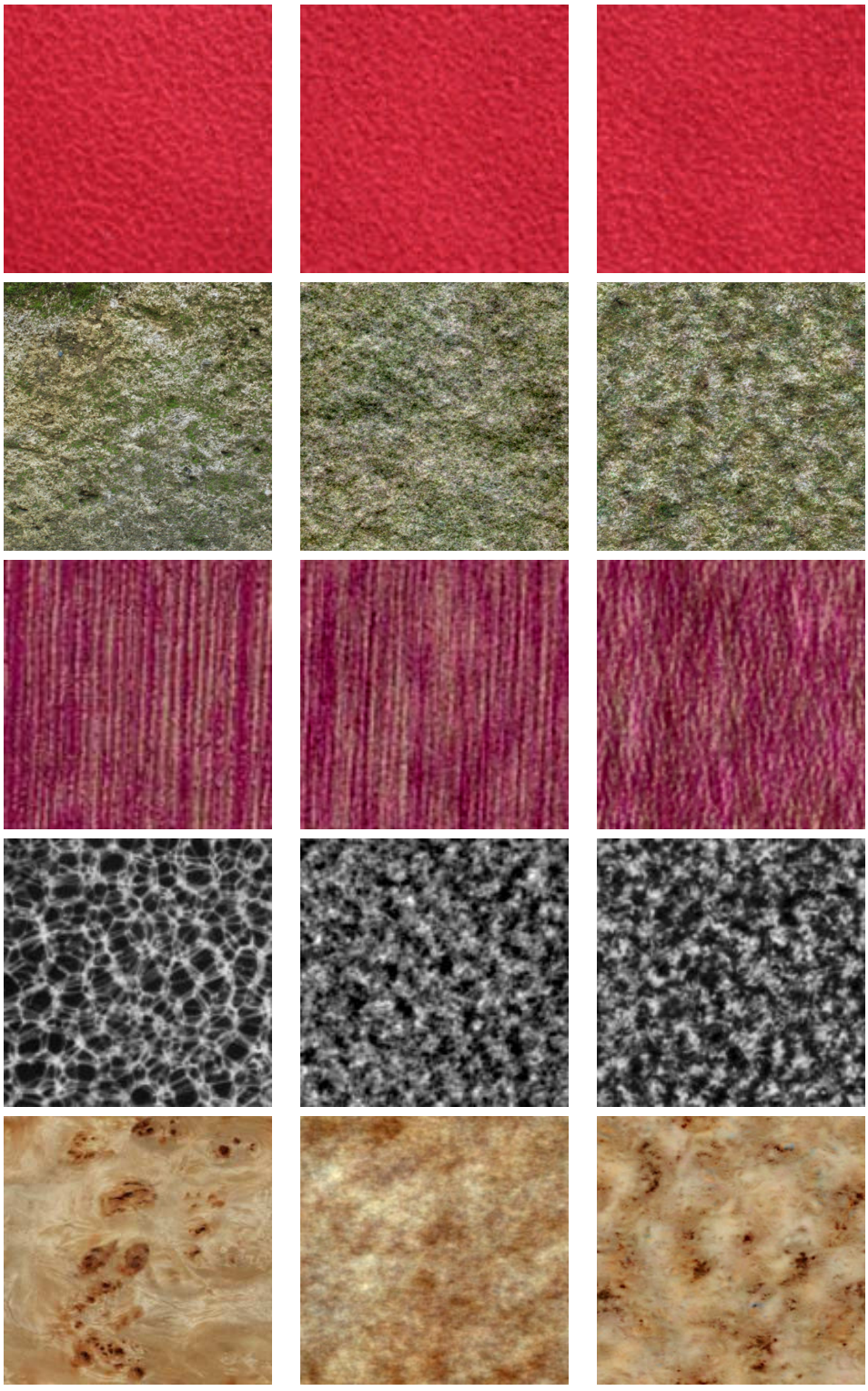
Actually, since edge structures can be seen as specific salient features, we can formulate another question whose conceptual aspect will be more inspiring: does

there exist a texture synthesis model/algorithm that precisely respect the local conspicuous features and thus complies well with Julesz's texton theory?

This question has been partially answered through the use of filter banks and wavelet transforms. Indeed, since the atoms of certain wavelet transforms have the aspect of local salient features, it seems convincing to address by-example texture synthesis via a method based on the first-order distribution of the wavelet coefficients. The problem has been addressed in those terms by the authors of [Heeger & Bergen 1995] who proposed a synthesis algorithm that matches the histograms of sub-bands in a steerable pyramid. This work is justified by the perceptual studies [Bergen & Adelson 1988] and [Malik & Perona 1990] which, in some sense, takeover Julesz's texton theory. In fact, the later article [Malik *et al.* 1999] tends to bridge the gap between the two approaches: its authors propose an operative definition of texton, as frequently-occurring local filter responses (obtained by K-means clustering performed over filter responses at many pixels). As concerns texture synthesis, the algorithm of [Heeger & Bergen 1995] is quite successful for by-example synthesis of microtextures, but for textures with more structured local features the results are in general not convincing. As one can see in Fig. 1.7, Heeger-Bergen algorithm is less precise than RPN or ADSN models in terms of frequency content. But in a few number of cases, it (surprisingly!) allows to reproduce non random phase features.

The main drawback of Heeger-Bergen algorithm is that it only enforces the marginal distribution of each filter response, without respecting the correlations between the filter responses. And, as the Fourier phase, these correlations are crucial to produce edges or conspicuous features. A very satisfying answer to that problem was brought in [Portilla & Simoncelli 2000]. Indeed, in this article, Portilla and Simoncelli propose to define a statistical texture model parametrized by several well-chosen second-order statistics computed in a complex wavelet transform. We refer to the original article for the detailed list of the chosen statistics but let us mention that among these, they capture local phase measurements through the complex correlation between a wavelet coefficient and its parent coefficient at the adjacent coarser scale. This certainly agrees with the importance of local phase shifts in image perception. Also, to address synthesis, this statistical model is sampled using an iterative procedure that alternately adjust all the chosen statistical constraints.

Among many conceptual advances, Portilla and Simoncelli demonstrate the richness of this parametric texture model by showing an amazing variety of textures (microtextures and also macrotextures) that are well-reproduced by their synthesis algorithm. Besides, their work also show that a small set of parameters (710 in their experiments but that could be even less) suffices to describe a very wide class of texture images. However, one of the weakness of Portilla-Simoncelli algorithm, which is already mentioned in the final section of their article, is that the output synthesis is not properly defined as a stationary random field (partly because they use a decimated wavelet transform, but also because the model is only partially determined by the iterative synthesis algorithm whose convergence is not proved). Also, since this iterative procedure is very slow, it is far from meeting the



Original RPN Heeger-Bergen

Figure 1.7: **Comparison between RPN and Heeger-Bergen synthesis algorithms**. In each row, from left to right, one can see an original texture, and the synthesis results of RPN algorithm and Heeger-Bergen algorithm. These results have been obtained with the online demos [Galerne *et al.* 2011a] and [Briand *et al.* 2014]. These algorithms perform well for microtexture synthesis, but as one can see in the third row, RPN is actually more precise. However, for macrotexture synthesis, both algorithms fail in general. Nevertheless, Heeger-Bergen algorithm sometimes reproduces non random phase features on very peculiar exemplars (last row).

requirements of real-time image synthesis.

Texture modelling with Markov random fields

Another weakness of Portilla-Simoncelli algorithm (mentioned in the "Discussion" section of [Portilla & Simoncelli 2000]) is that it sets once and for all, among the second-order distribution of filter responses, a subset of interesting statistical constraints whose choice is essentially motivated by empirical observations. Therefore, the synthesis algorithm still has a few failure cases and thus one would mitigate the first sentence of their abstract: their statistical model is just nearly universal (which was still, of course, an incredible breakthrough in texture modelling).

A much less specific approach was proposed two years before by the authors of [Zhu *et al.* 1998]. In this article, Zhu et al. completely identify the concept of texture with the one of random field. Subsequently, they explain that if one has access to several exemplars of the same texture, then texture analysis boils down to inference on a probability distribution, and that, after this inference step, one can perform texture synthesis by sampling according to that probability distribution. Of course, apart from the Gaussian case where such an analysis/synthesis pipeline is easily followed, these two steps are very hard to solve in a general setting. In their FRAME algorithm, Zhu et al. suggest to infer a Gibbs distribution by selecting automatically a few filters (among a large filter bank) that will capture the important features of the considered texture.

The FRAME algorithm thus automatically recognizes the features that will play a role in the perception of a specific texture. In terms of textons, this analysis step is close to the one of [Malik *et al.* 1999] because in some sense, both articles go against the search of a universal notion of texton. Indeed, since the concept of texture is quite vague, we do not know *a priori* what kind of local details will drive our perception, and thus it is natural to define the textons with respect to a specific class of textures. In the example of Gaussian textures, this justifies again the definition of texton that was given in [Desolneux *et al.* 2012]. This idea has also been pursued in [Zhu *et al.* 2005] whose authors propose a more generative view of textons.

One of the advantage of FRAME is that it is formulated in the framework of Gibbs distributions, which are shown to be equivalent to Markov random fields by the Hammersley-Clifford theorem [Besag 1974], [Winkler 2006]. Let us recall that a Markov random field (MRF) is a probability distribution P on a set of images $\mathbb{R}^\Omega$ such that (with obvious notations for the conditional distributions) for every pixel $\mathbf{x}$,

$$P(\ u(\mathbf{x}) \mid u(\mathbf{y}),\ \mathbf{y} \neq \mathbf{x}\ ) = P(\ u(\mathbf{x}) \mid u(\mathbf{y}),\ \mathbf{y} \in \mathcal{N}_{\mathbf{x}}\ , \mathbf{y} \neq \mathbf{x}\ )\ ,$$

where $\mathcal{N}_{\mathbf{x}}$ is a neighborhood of $\mathbf{x}$. The right-hand side of this equation is sometimes called local specification of the MRF.

Therefore, Markov random fields form a clearly defined probabilistic model which is very suited to texture analysis because it is completely determined by the local specification. Even if MRF models are generally presented only in the case of

a finite image domain[3] Ω, relying on a precise random model is still an appreciable theoretical asset (which does not have [Portilla & Simoncelli 2000] for example). MRF have already been used for texture analysis and synthesis long before the article of Zhu et al., for example in [Cross & Jain 1983], [Chellappa & Chatterjee 1985] and [Geman & Graffigne 1986]. But these three articles use simple local specifications parametrized by a few number of coefficients which can be estimated on an exemplar; hence the richness of these models is very limited.

Nonetheless, anyone who performs texture synthesis based on a MRF model is confronted to the fact that in general, the sampling of MRF involves heavy numerical simulations. Apart from very specific cases, the simulation of MRF relies on Gibbs sampling, which is a very costly algorithm especially with large size neighborhoods $\mathcal{N}_{\mathbf{x}}$, and whose convergence speed may depend on the model (because it is already true for the "simple" case of Ising model). This practical problem is thus the main limitation of FRAME synthesis algorithm.

The incredible success of non-parametric texture synthesis by example

Fortunately, a tremendous simplification was brought quasi-simultaneously by the authors of [Efros & Leung 1999] and [Wei & Levoy 2000]. They showed that by-example texture synthesis could be performed by progressive filling of the synthesis domain with a kind of copy-paste technique: in order to synthesize a new pixel $\mathbf{x}$, one only has to look for exemplar pixels whose neighborhoods resemble the already synthesized pixels in the neighborhood of $\mathbf{x}$. Then, one can randomly sample from these exemplar values as suggested in [Efros & Leung 1999] or simply copy the most resembling exemplar value as in [Wei & Levoy 2000]. Apart from slight modifications involved in the final algorithms (for example in [Wei & Levoy 2000], the multiresolution scheme, and the tree-structured vector quantization which greatly fastens the execution), the simple "copy-paste" principle was established. In terms of MRF, it means that one has access to a partial realization of the random field, sampling according to the local conditional distribution can be approximated by this simple copy-paste operation. This tremendous simplification certainly explains the success of their method over the one proposed in [Paget & Longstaff 1998] which consists in a precise non-parametric estimation and sampling of the local conditional distributions of the MRF. The algorithms of [Efros & Leung 1999] and [Wei & Levoy 2000] exhibit good results even on structured textures. As one can try with the online demo [Aguerrebere *et al.* 2013], these algorithms may be less precise for microtexture synthesis than random phase algorithms, but they perform well on many more structured textures.

The works of Efros-Leung and Wei-Levoy have inspired many further texture synthesis articles and methods; here we will mention only a few ones that we judge apropos with regards to the problematics of this thesis; the interested reader may refer to [Wei *et al.* 2009] for a more exhaustive (though already outdated) review. Several authors remarked that in Efros-Leung algorithm, only a few pixels have a

[3] Several technical difficulties arise when considering Markov random fields defined over the infinite lattice $\mathbb{Z}^2$, see for example [Prum 1997].

truly random choice when sampled, so that the copy-paste principle could be accelerated by processing patches instead of pixels. Thus, the method of [Liang *et al.* 2001] accelerates the one of [Efros & Leung 1999] by using conditional sampling of patches instead of pixels. The authors of [Efros & Freeman 2001] also exploit such a patch-based sampling method but obtain much better results since they cautiously handle the boundary between adjacent overlapping blocks: a minimum error boundary cut is found with a dynamic programming method. The authors of [Kwatra *et al.* 2003] bring a considerable algorithmic improvement to Efros-Freeman' method by using graphcut optimization to stitch the texture blocks. Let us also mention the work [Ashikhmin 2001] which modifies Wei-Levoy' algorithm by encouraging verbatim copy; this modification is quite questionable in terms of random simulation but undeniably fastens the method by avoiding exhaustive search in the exemplar texture. The methods of [Wei & Levoy 2000] and [Ashikhmin 2001] have inspired the concept of image analogies [Hertzmann *et al.* 2001] which has numerous applications, although it brings only minor improvements as concerns texture synthesis. The patch similarity in [Efros & Leung 1999] also inspired the "non-local means" denoising algorithm [Buades *et al.* 2005]. Finally, the unpublished work [Wei & Levoy 2002] achieves order-independence in the pixels filling; we would like to highlight here that this feature is conceptually attractive because it guarantees some kind of stationarity in the output random field.

Texture Optimization

The order-independence is inherently respected by synthesis algorithms based on texture optimization [Kwatra *et al.* 2005] (and its surface analogue [Han *et al.* 2006]). Given an exemplar texture $u : \Omega_u \longrightarrow \mathbb{R}^d$, Kwatra et al. suggested to synthesize a texture $v : \Omega_v \longrightarrow \mathbb{R}^d$ by finding a local minimum of the "texture energy"

$$E(v) = \sum_{\mathbf{y} \in \Omega_v} \min_{\mathbf{x} \in \Omega_u} \|p_u(\mathbf{x}) - p_v(\mathbf{y})\|^2 \ ,$$

where $p_u(\mathbf{x})$ denotes the patch of u that is centered on pixel $\mathbf{x}$. The authors propose to minimize this energy, starting from a random initialization, by iterating simultaneous projections of all the pixel values to the exemplar values with the closest resembling neighborhood. Here too, the final algorithm is a bit more complicated (multiresolution framework, reweighted least-square for optimization, and K-means acceleration for patch search in the exemplar) but this method is an inherently stationary improvement of [Wei & Levoy 2000] and produces both impressive and stable results on highly structured textures (the stability of the algorithms will be discussed below). Texture optimization can also address constrained synthesis [Kwatra *et al.* 2005], [Ramanarayanan & Bala 2007], [Kim *et al.* 2012]. Finally, let us mention that similar functionals have been used to address video inpainting problems [Wexler *et al.* 2004], [Newson *et al.* 2014] and also inspired the PATCHMATCH algorithm for image editing [Barnes *et al.* 2009].

Maybe one last drawback of texture optimization algorithms is that they do not meet with the requirements of computer graphics. For this kind of applications,

the algorithm is often asked to be procedural (meaning that the output texture can be evaluated at any spatial position (in a continuous setting)), parallel (meaning that the output pixel values can be computed separately), and must have low memory storage (because of constraints involved in GPU execution). Nowadays, many algorithms allow for real-time texture synthesis, but they are often limited to a certain class of textures. Since the seminal paper [Perlin 1985], several procedural noises have been proposed [Cook & DeRose 2005], [Goldberg *et al.* 2008] (see [Lagae *et al.* 2010a] for a detailed review), with recent progress in procedural noise by example [Lagae *et al.* 2009], [Lagae *et al.* 2010b], [Galerne *et al.* 2012], [Gilet *et al.* 2014]; but all these methods essentially concentrate on random phase textures. Closer to the aforementioned texture optimization models, lies the synthesis algorithm of [Lefebvre & Hoppe 2005] which is parallel but not procedural, and was followed by several variants like [Dong *et al.* 2008] or [Han *et al.* 2008]. In this article, Lefebvre and Hoppe propose to generate, in a multiscale framework, more and more detailed versions of the texture, by performing, at each scale, a jittering step (that introduces some randomness) and a correction step (inspired by [Ashikhmin 2001]). A main strength of this method is that it has been designed to allow for parallel evaluation and thus benefits from fast GPU execution Besides, since it combines many advantages of previous methods, it is able to handle a very large class of textures. If we could here express some personal point of view, we would say that this method is, up to date, the most impressive synthesis algorithm with regards to its execution speed and the diversity of well-reproduced textures.

Towards texture models that combine the flexibility of the spot noise model and the richness of patch-based models

Despite all these efforts directed towards texture synthesis, we believe that there is still room for improvement in the research of texture models having certain mathematical guarantees. In particular, such guarantees could solve the stability problems encountered with methods inspired by the "copy-paste" principle. For example, Efros-Leung algorithm is not stable for two reasons. On the one hand, two output textures computed on the same exemplar may look different. And on the other hand, on certain textures, the result is corrupted by the famous "growing-garbage" effect (which is already mentioned by [Efros & Leung 1999] and illustrated in Fig. 1.8). These stability problems can certainly be explained by the weakness of mathematical guarantees provided by Efros-Leung algorithm. In fact, the authors of [Levina & Bickel 2006] showed that Efros-Leung algorithm is a consistent resampling method provided that the size of the exemplar tends to infinity. Such a result is of course very welcomed but unfortunately not sufficient to ensure practical stability of the synthesis algorithm[4]. In contrast, these stability problems do not happen anymore with truly stationary schemes. For example, with natural texture samples, several results of RPN or ADSN synthesis have similar appearance; also,

[4]It would certainly be interesting to pursue the work of [Levina & Bickel 2006] by proving an (approximate) consistency result based only on a finite observation, even if it requires very strong assumptions of the random field.

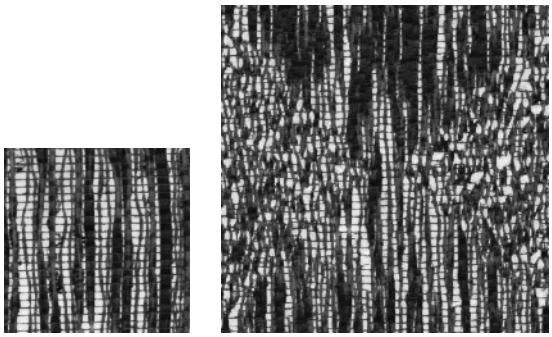


Figure 1.8: **Growing-garbage effect**. The synthesized texture on the right has been obtained with Efros-Leung algorithm (using the online demo [Aguerrebere *et al.* 2013]) applied to the original texture shown on the left. In the output image, one can clearly distinguish a spatial boundary beyond which the synthesis fails (we say that the algorithm is "growing garbage").

to the best of our knowledge, no stability problems have been encountered in the stationary method of [Kwatra *et al.* 2005].

Hence the need of clear random texture models that are simple enough to allow for mathematical analysis (and thus for theoretical guarantees and/or fast sampling), and whose richness comes close to the one of recent patch-based models. Such a model would certainly rely on the analysis of patch distributions. Considering the patch distribution of a texture is, to some extent, a way to reconcile the approaches based on filter banks, and the approaches based on Markov random fields (thereby avoiding the difficulties attached to conditional distributions). This point of view is justified by the findings of [Varma & Zisserman 2003] whose authors showed that better results could be obtained in texture classification by direct analysis of the patch distribution, compared to the results obtained via a filter bank (even if the filters support is larger than the patch domain). Therefore, this suggests that textons could be directly defined through the analysis of the patch distribution.

One successful way to analyze a patch distribution is to make use of visual dictionaries. Since the concept of dictionary originated in the need of adaptive sparse representations of images, such approaches in texture synthesis could be considered as the natural continuation of filter-based models and in particular [Zhu *et al.* 2005]. The author of [Peyré 2009] successfully combines texture optimization [Kwatra *et al.* 2005] and a dictionary learning technique [Mairal *et al.* 2008]; the corresponding synthesis algorithm is able to synthesize structured textures in a slightly over-regularized manner. Later, this work about sparse texture synthesis was pursued in [Tartavel *et al.* 2014]. The authors of this article suggest to minimize a functional made of different terms. One term constrains the dictionary atoms to be used with the same frequencies than in the exemplar texture (which makes the output textures seem less regular than in [Peyré 2009]). The other terms constrain the color distribution, and also the global Fourier modulus (which, again, is the only extracted information in RPN

and ADSN synthesis). The output randomness is obtained through stochastic initialization of the minimization process. Unfortunately the synthesized texture obtained by the algorithm of [Tartavel *et al.* 2014] is still quite difficult to describe as a stationary random field. Indeed, further research would be required to understand more thoroughly the distribution of images obtained by applying a a global optimization procedure to a random field. If the functional is considered to be a Gibbs potential, this method seems more as a way to extract local maxima *a posteriori* than a method to sample according to the Gibbs distribution.

In Chapter 5 of this thesis, we have followed a different path: we tried to find simple random fields (inspired by the aforementioned synthesis algorithms) whose simulation is straightforward and that are still able to reproduce a wide class of textures. Inspired by the aforementioned works in texture synthesis, we tried to design new texture models that combine the benefits of both spot noise models and patch-based models.

In Section 5.2 we define localized variants of the ADSN model which allow to resynthesize on the same domain non-stationary microtextures. The corresponding simulation scheme is still very efficient because such a local ADSN can be approximated by a low-intensity discrete spot noise (as for the ADSN associated to the synthesis-oriented texton in Section 3.3). Besides, we can fully describe the distribution of the random field since it is Gaussian with easily computed first-order and second-order moments. This algorithm is able to resynthesize non-stationary microtextures (like the ones encountered in radiographic images) because it locally respects the texture autocorrelation. Besides, experimental results show that this synthesis commutes with the addition of a sufficiently smooth component; this method can thus be understood as approximate resynthesis of a Gaussian texture conditioned on a low-frequency component. The corresponding synthesis results illustrate the potential of Gaussian models with relaxed stationarity constraint.

Finally, in Section 5.3 we propose to define bi-level texture models which consist of a coarse-scale random phase field on which fine scale details are added through a local function. Such a model agrees with the paradigm [Galerne *et al.* 2011b] that a texture can be considered as random phase if it is seen from sufficiently far away. We present one instance of bi-level models by using a patch-based local function whose construction is inspired by [Kwatra *et al.* 2005]. It leads to a by-example synthesis algorithm whose results are comparable to the ones of [Kwatra *et al.* 2005] but with a simpler synthesis algorithm and with additional mathematical guarantees. One main conceptual asset of this model is that it can be seen as a local function of a stationary Gaussian random field. In particular, the stationarity and the long-range independence obtained with this texture model are good guarantees for stability, which is illustrated by the fact that it can be used to synthesize structured textures on very large domains.

1.5 Detailed Outline of the Thesis

We here give a detailed overview of the following chapters, with the main contributions written in bold characters.

Chapter 2

In Chapter 2, we propose a **unified framework for stationary texture modelling** in which we define, study, and **compare several well-known texture models**, like the spot noise model, the random phase noise, moving averages, autoregressions, and Gaussian Markov random fields (GMRF). For the sake of readability, we distinguish between texture models that are defined over a periodic finite domain $\Theta = \mathbb{Z}/M\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$ called circular models, and texture models that are defined over the whole plane $\mathbb{Z}^2$. In the first case, the spectral representation is given by the DFT, whereas it is obtained with Herglotz theorem in the second case.

After recalling the construction and well-known properties of the spot noise model, we discuss several questions related to ADSN random fields. In particular, in the circular gray-level case, we recall that any stationary Gaussian random field can be obtained as an ADSN, and in the circular color case, we propose a **canonical decomposition of a $\mathbb{R}^d$-valued Gaussian stationary random field** into a sum of d independent ADSN fields. The notion of texton is recalled for circular Gaussian models [Desolneux *et al.* 2012] and also introduced for Gaussian models on $\mathbb{Z}^2$ (but mainly used for simulation purpose)

Finally, we recall the definition of the optimal transport distance between circular random fields and recall (with a simplified proof) the result of [Xia *et al.* 2014] that expresses the distance between two circular ADSN fields. We also extend this definition and this result to the infinite lattice by defining an **optimal transport distance between stationary random fields on $\mathbb{Z}^2$**.

Chapter 3

In Chapter 3, we explain how the random phase models introduced in Chapter 2 can be used to perform fast by-example synthesis of microtextures.

In Section 3.1, we show how to estimate a Gaussian model from an exemplar texture $u : \Omega \longrightarrow \mathbb{R}^d$. In particular, we **compare periodic and non-periodic estimators of the covariance function**. Next, in Section 3.2 we present several examples of random phase texture synthesis, with the underlying desire to precisely highlight the limits of RPN and ADSN synthesis.

In Section 3.3 we present the main contribution of this chapter: we explain how to derive from an exemplar texture a so-called **synthesis-oriented texton** (SOT). This texton is designed to hold in a prescribed compact support, and to realize a good approximation of the ADSN model associated to the exemplar. Besides, the SOT spreads over all the available support, and thus the corresponding DSN realizes a good approximation of the Gaussian limit even for a reasonably low intensity. We will thus show that **many microtextures can be convincingly synthesized with a DSN model associated to a very small SOT** (31×31) **and**

with less than 50 operations per pixel. We can thus take profit of the numerous advantages of direct spot noise synthesis, in particular fast parallel local evaluations. Motivated by a peculiar example of SOT, we also show that **the raw optimal transport distance is not sufficient to measure precisely the perceptual proximity between two textures** because of the equal contribution of all frequencies of the Fourier domain. We thus propose an **optimal transport distance including frequency weights** which reflects more accurately our texture perception and which is used to **get a more robust SOT**.

Finally, in Section 3.4, we apply random phase texture models to the problem of textural inpainting. Since textural inpainting can be clearly formulated as conditional simulation, **we obtain a satisfying textural inpainting algorithm for microtextures** by adapting the **Gaussian conditional simulation** scheme based on kriging estimation [Lantuéjoul 2002].

Chapter 4

Chapter 4 is devoted to a thorough presentation of the global phase coherence indices GPC, SI and S. In Section 4.3, we give basic properties of these indices and also **analyze the regularity of SI and** S. We also show that **the phase coherence indices of random phase fields are expected to be low** (we actually compute the exact distribution of the GPC of a RPN). A related question is to study the TV of a random phase field. In Appendix 4.A, we provide an **explicit approximation of the expected TV in the RPN model** (which can be plugged in the numerical computation of GPC). Also, in Appendix 4.B, in a well-chosen asymptotic framework, we justify the **Gaussian approximation of the TV distribution of certain ADSN models** (which underlies the operative definition of SI).

In Section 4.4, we provide a practical study which relates these indices to the image sharpness. We provide many experiments that illustrate their behavior on natural images, and in particular, with the same methodology than [Blanchet *et al.* 2008] and [Blanchet & Moisan 2012] we confirm that **these indices are sensitive to blur, noise and ringing artifacts**. However, we illustrate in the end of this section that these indices have no reason to reflect perfectly the perceptual notion of sharpness (in particular when facing phase-coherent artifacts). Nevertheless, the index S can still be considered as a good prior for deblurring an image u. In Section 4.5, we propose to find a restoration kernel k that maximizes the sharpness $S(k * u)$ of the restored image. Restricting to the class of kernels having radial and unimodal DFT, then we can use a simple stochastic scheme to optimize this functional. **This method can thus treat isotropic blur by applying a well-chosen deconvolution filter** and compares well to other purely non-linear methods while keeping the benefits of linear filtering.

Chapter 5

In Section 5.1, we illustrate the difficulty of direct phase analysis/synthesis. We first show that **the stationary assumption imposes to consider the phase modulo the linear functions (called ramp functions)**. We will see that the bispectrum phase provides such a representation but does not seem appropriate for by-example texture synthesis. To end this section, we illustrate that the phase coherence indices of Chapter 4 are not sufficient to assess precisely the plausibility of an exemplar texture in the random phase model.

The rest of this chapter is devoted to texture models which combine the benefits of Gaussian models and patch-based sampling. In Section 5.2 we define a local spot noise model which allows to resynthesize non-stationary microtextures. The corresponding Gaussian limit can still be approximated by a low-intensity local spot noise, which allows for fast simulation. Several examples of local spot noise show that **the Gaussian model gets much richer with a relaxed stationarity constraint**.

Finally, in Section 5.3 we propose to define bi-level models as local functions of coarse-scale stationary Gaussian random fields. We formulate **simple mathematical properties of such bi-level models**, like stationarity or long-range independence, which can be understood as **strong guarantees of stability and innovation**. Using a patch-based local function inspired by [Kwatra *et al.* 2005], we show that such **a bi-level model can synthesize many macrotextures on very large domains**.

1.6 Publications

The content of Chapter 4 has been recently published as is in the journal article

> **No-reference image quality assessment and blind deblurring with sharpness metrics exploiting Fourier phase information** (Arthur Leclaire, Lionel Moisan), *Journal of Mathematical Imaging and Vision*, 2015.

A preliminar version of this work was presented in the conference paper

> **Blind Deblurring Using a Simplified Sharpness Index** (Arthur Leclaire, Lionel Moisan), *proceedings of the fourth International Conference on Scale Space and Variational Methods in Computer Vision*, Lecture Notes in Computer Science, vol. 7893, pp. 86-97, 2013.

The technical issue of border effects in the computation of SI was discussed in

> **Une Variante non Périodique du Sharpness Index** (Arthur Leclaire, Lionel Moisan), *Actes du GRETSI*, 2013.

Finally, the synthesis-oriented texton was presented in the conference paper

> **A Texton for Fast and Flexible Gaussian Texture Synthesis** (Bruno Galerne, Arthur Leclaire, Lionel Moisan), *proceedings of the European Signal Processing Conference*, 2014.

Chapter 2

Random Phase Fields

Contents

The objective of this chapter is to introduce the random phase fields that will be used as texture models in Chapter 3 and for phase coherence evaluation in Chapter 4. In particular, a wide part of this chapter is devoted to the study of discrete Gaussian random fields.

The Gaussian random fields that were first proposed for texture modelling consist of 2-D generalizations of the autoregressive and moving average fields obtained with a Gaussian excitation function. Such models were used for texture synthesis in [Chellappa & Kashyap 1985] and [Cadzow *et al.* 1993], for texture classification in [Chellappa & Chatterjee 1985], and also for texture segmentation in [Chellappa 1985]. Notice that the term "autoregressive" is not adapted to the two-dimensional case for which there is no natural definition of the past and future of the process; this is why the corresponding texture model is called "noncausal autoregressive". The spectral density of these Gaussian processes is a trigonometric polynomial (often with low degree) in the moving average case, and a trigonometric rationale function in the autoregressive case.

The moving average fields are a particular case of filtered white noises. In the computer graphics community, such models were proposed for texture synthesis

in [Lewis 1984], [Lewis 1989] and [Van Wijk 1991]. In particular, the spot noise model of [Van Wijk 1991] consists of the superposition of randomly-shifted copies of a kernel h positioned according to a Poisson point process. In other words, the spot noise is the convolution of a kernel h with a Poisson white noise. In [Van Wijk 1991], the author notices that the Fourier transform of the spot noise is obtained by multiplying the Fourier coefficients of h by a scale factor and a phase shift. He clearly states the randomness of the phase shifts, but not the one of the scale factor, leading to a confusion between the spot noise model and any other filtered white noise model. However the experiments shown in his article clearly demonstrates the richness of the spot noise model. In contrast to the moving average and noncausal autoregressive models, the spectral density of the spot noise model is not restricted to a parametric class.

The mathematical analysis of the spot noise model was presented in [Galerne *et al.* 2011b]. In this article, is is shown that when the intensity of the Poisson point process tends to infinity, the discrete spot noise converges to the so-called asymptotic discrete spot noise (ADSN) which is the convolution of the kernel with a Gaussian white noise. The authors also make a clear distinction between the ADSN and the random-phase noise (RPN) which is a process with fixed Fourier modulus and with uniform random Fourier phase. However, they show that both ADSN and RPN can model a wide class of microtextures composed of texture images for which the phase does not convey any useful information.

The stationary Gaussian model (and in particular ADSN), besides from its interest in texture synthesis demonstrated in [Galerne *et al.* 2011b], leads to fruitful mathematical developments. For example it was used in [Grosjean & Moisan 2009] to study the detectability of spots on textured background using the *a-contrario* methodology. Also, an expression of the optimal transport distance between two ADSN random fields was given in in [Xia *et al.* 2014] and [Desolneux *et al.* 2015] (with an application to texture mixing in the former article).

The Gaussian random fields are characterized by their moments of order 1 and 2. Even if this property is very convenient on a theoretical point of view (for example when doing conditional simulation as in Section 3.4), it is also their main limitation as texture models. Indeed, as shown in [Julesz 1981], the pre-attentive perception of texture by the human visual system is not reduced to the statistics of order 1 and 2 but also depends on local conspicuous features. More precisely, since their Fourier phase is uniform, the ADSN and RPN fields cannot model textures with geometrical salient elements (like sharp edges), according to the observations of [Oppenheim & Lim 1981]. This property restricts their performance in by-example texture synthesis, as will be seen in Chapter 3.

In this chapter, we present in a unified framework several random phase fields that will be later used as texture models. In particular we study the RPN, ADSN, and general Gaussian models, with particular attention on the spectral representation of the random fields. We explain the available simulation scheme for such Gaussian models, and we take this opportunity to define the textons associated to a Gaussian model. We also define and compute the optimal transport distance between ADSN fields. For the sake of clarity, we prefer to distinguish between the

models defined on a circular finite domain, and those defined on the infinite lattice $\mathbb{Z}^2$. Indeed, the analysis and simulation of circular random fields is easy since the discrete Fourier basis is an eigenvector basis of the corresponding covariance operators. In contrast, the spectral representation of stationary random fields on $\mathbb{Z}^2$ may not be sampled on a discrete grid in Fourier domain. Since we are interested in color texture models, we consider random phase fields with values in $\mathbb{R}^d$ where d denotes the number of channels (3 for an RGB image). The main contributions of this chapter are the decomposition of a multi-channel Gaussian field as a sum of ADSN fields (Subsection 2.1.4), a simpler proof of the expression giving the optimal transport distance between circular ADSN fields (Subsection 2.1.6), and its extension to the case of ADSN fields defined over the whole plane $\mathbb{Z}^2$ (Subsection 2.2.7).

2.1 Circular Stationary Random Fields

Let $\Theta = \mathbb{Z}/M\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$ be a periodic rectangular discrete domain of size $M \times N$.

2.1.1 General Definitions

In this subsection, we recall general definitions and properties of circular stationary random fields and their spectral representation. Even if these general results are well-known (see for example [Doob 1990], [Galerne *et al.* 2011b], [Xia *et al.* 2014]), they are recalled with their proofs, for the sake of completeness.

Random Fields and Covariance Functions

Definition 2.1.1. A random field on Θ is a random process F on Θ with values in $\mathbb{R}^d$, meaning that for all $\mathbf{x} \in \Theta$, $F(\mathbf{x})$ is a random variable in $\mathbb{R}^d$. Omitting to write the probability sample ω, we will often denote F as a random function $F : \Theta \longrightarrow \mathbb{R}^d$. It is said to be of order p if for all $\mathbf{x} \in \Theta$, $\mathbb{E}(\|F(\mathbf{x})\|^p) < \infty$. If F is of order one, its expectation (or mean field) is the function $m : \Theta \longrightarrow \mathbb{R}^d$ defined by

$$\forall \mathbf{x} \in \Theta, \quad m(\mathbf{x}) = \mathbb{E}(F(\mathbf{x})) \; .$$

If F is of order two, its spatial covariance is the function $\Gamma_F : \Theta \times \Theta \longrightarrow \mathbb{R}^{d \times d}$ defined by

$$\forall \mathbf{x}, \mathbf{y} \in \Theta, \quad \Gamma_F(\mathbf{x}, \mathbf{y}) = \mathrm{Cov}(F(\mathbf{x}), F(\mathbf{y})) = \mathbb{E}((F(\mathbf{x}) - m(\mathbf{x}))(F(\mathbf{y}) - m(\mathbf{y}))^T) \; .$$

The random field F is said to be Gaussian if any linear combination of its values is Gaussian. The distribution of a Gaussian random field F with mean m and covariance Γ will be denoted by $\mathcal{N}(m, \Gamma)$.

Notice that a covariance Γ can also be seen as the covariance matrix of a random vector in $\mathbb{R}^{\Theta \times d}$, and thus inherits the property of such covariance matrices. But since we are particularly interested in the spectral analysis, it is worth writing these

properties with a clear distinction between the spatial index $\mathbf{x} \in \Theta$ and the channel index $j \in \{1, \dots, d\}$. The covariance satisfies the symmetry constraint

$$\Gamma_F(\mathbf{y}, \mathbf{x}) = \Gamma_F(\mathbf{x}, \mathbf{y})^T$$

and the non-negativity constraint

$$\forall f \in (\mathbb{C}^d)^\Theta, \quad \sum_{\mathbf{x},\mathbf{y} \in \Theta} f(\mathbf{x})^T \Gamma_F(\mathbf{x}, \mathbf{y}) \overline{f(\mathbf{y})} \geqslant 0 \ ,$$

where $\overline{f(\mathbf{y})}$ refers to the complex conjugate of $f(\mathbf{x})$. The covariance is associated to the non-negative Hermitian form on $(\mathbb{C}^d)^\Omega$

$$\begin{array}{ccc} (\mathbb{C}^d)^\Theta \times (\mathbb{C}^d)^\Theta & \longrightarrow & \mathbb{C} \\ (f, g) & \longmapsto & \displaystyle\sum_{\mathbf{x} \in \Theta, \mathbf{y} \in \Theta} f(\mathbf{x})^T \Gamma_F(\mathbf{x}, \mathbf{y}) \overline{g(\mathbf{y})} \end{array} \ .$$

Conversely, given any function Γ which satisfies these two properties, one can build a random field on Θ with expectation 0 and covariance Γ.

Definition 2.1.2. A random field F on Θ is said to be circular stationary (or simply circular) if for every $n \geqslant 1$, $\mathbf{x}_1, \dots, \mathbf{x}_n \in \Theta$, $\mathbf{v} \in \Theta$, $(F(\mathbf{x}_1), \dots, F(\mathbf{x}_n))$ has the same distribution as $(F(\mathbf{x}_1 + \mathbf{v}), \dots, F(\mathbf{x}_n + \mathbf{v}))$.

If F is a second-order stationary random field on Θ, its expected value

$$m = \mathbb{E}(F(\mathbf{x})) \in \mathbb{R}^d$$

does not depend on the location $\mathbf{x}$, and the covariance between two samples

$$\mathrm{Cov}(F(\mathbf{x}), F(\mathbf{y})) = \mathbb{E}((F(\mathbf{x}) - m)(F(\mathbf{y}) - m)^T) \in \mathbb{R}^{d \times d}$$

only depends on the shift $\mathbf{y} - \mathbf{x}$.

Definition 2.1.3. The covariance function of a second-order circular random field F on Θ is the function $C_F : \Theta \longrightarrow \mathbb{R}^{d \times d}$ such that

$$\forall \mathbf{x}, \mathbf{y} \in \Theta, \quad \mathrm{Cov}(F(\mathbf{x}), F(\mathbf{y})) = C_F(\mathbf{x} - \mathbf{y}) \ .$$

By abuse of notation, we still denote by $\mathcal{N}(m, C)$ the distribution of a Gaussian circular stationary random field with mean m and covariance function C.

If F is a second-order circular random field on Θ, we thus have

$$\forall \mathbf{x}, \mathbf{y} \in \Theta, \forall \mathbf{v} \in \Theta, \quad \Gamma_F(\mathbf{x} + \mathbf{v}, \mathbf{y} + \mathbf{v}) = \Gamma_F(\mathbf{x}, \mathbf{y}) \ .$$

This property is sometimes referred to as "Γ_F is circulant".

Spectral Representation

Let us recall that the discrete Fourier transform (DFT) of a (matrix-valued) function $f : \Theta \longrightarrow \mathbb{R}^{p\times q}$ is the function $\hat{f} : \Theta \longrightarrow \mathbb{C}^{p\times q}$ defined by

$$\forall \boldsymbol{\xi} \in \Theta, \quad \hat{f}(\boldsymbol{\xi}) = \sum_{\mathbf{x}\in\Theta} f(\mathbf{x}) e^{-i\langle \boldsymbol{\xi}, \mathbf{x}\rangle} , \tag{2.1}$$

where $\langle \boldsymbol{\xi}, \mathbf{x}\rangle = 2\pi(\frac{x_1\xi_1}{M} + \frac{x_2\xi_2}{N})$ for $\boldsymbol{\xi} = (\xi_1, \xi_2) \in \Theta$ and $\mathbf{x} = (x_1, x_2) \in \Theta$. Since f is real valued, we have $\hat{f}(-\boldsymbol{\xi}) = \overline{\hat{f}(\boldsymbol{\xi})}$. Notice that, if f, g are two matrix-valued functions on Θ such that the product fg is well defined, then we can define their (circular) convolution by

$$\forall \mathbf{x} \in \Theta, \quad f * g(\mathbf{x}) = \sum_{\mathbf{y}\in\Theta} f(\mathbf{y}) g(\mathbf{x} - \mathbf{y}) .$$

The DFT of the convolution is given by $\widehat{f * g} = \hat{f}\hat{g}$. Indeed, for all $\boldsymbol{\xi} \in \Theta$, we have

$$\begin{aligned}
\widehat{f * g}(\boldsymbol{\xi}) &= \sum_{\mathbf{x}\in\Theta} f * g(\mathbf{x}) e^{-i\langle \boldsymbol{\xi}, \mathbf{x}\rangle} \\
&= \sum_{\mathbf{x},\mathbf{y}\in\Theta} f(\mathbf{y}) g(\mathbf{x} - \mathbf{y}) e^{-i\langle \boldsymbol{\xi}, \mathbf{x}\rangle} \\
&= \sum_{\mathbf{x},\mathbf{y}\in\Theta} f(\mathbf{y}) e^{-i\langle \boldsymbol{\xi}, \mathbf{y}\rangle} g(\mathbf{x} - \mathbf{y}) e^{-i\langle \boldsymbol{\xi}, \mathbf{x}-\mathbf{y}\rangle} \\
&= \sum_{\mathbf{x},\mathbf{z}\in\Theta} f(\mathbf{y}) e^{-i\langle \boldsymbol{\xi}, \mathbf{y}\rangle} g(\mathbf{z}) e^{-i\langle \boldsymbol{\xi}, \mathbf{z}\rangle} = \hat{f}(\boldsymbol{\xi})\hat{g}(\boldsymbol{\xi}) .
\end{aligned}$$

In particular, if we set $\tilde{f}(\mathbf{x}) = f(-\mathbf{x})$, then the (non-centered) autocorrelation of a function $f : \Theta \longrightarrow \mathbb{R}^p$ defined by

$$\forall \mathbf{v} \in \Theta, \quad f * \tilde{f}^T(\mathbf{v}) = \sum_{\mathbf{x}\in\Theta} f(\mathbf{x}) \tilde{f}(\mathbf{v} - \mathbf{x})^T = \sum_{\mathbf{x}\in\Theta} f(\mathbf{x}) f(\mathbf{x} - \mathbf{v})^T$$

satisfies

$$\widehat{f * \tilde{f}^T} = \hat{f}\widehat{\tilde{f}^T} = \hat{f}\hat{f}^* ,$$

where $\hat{f}(\boldsymbol{\xi})^* = \overline{\hat{f}(\boldsymbol{\xi})^T}$ is the transposed conjugate of $\hat{f}(\boldsymbol{\xi})$.

Let F be a second-order circular random field on Θ with mean m and covariance function C. One has for all $\mathbf{v} \in \Theta$,

$$\mathbb{E}(F * \tilde{F}^T(\mathbf{v})) = \sum_{\mathbf{x}\in\Theta} \mathbb{E}(F(\mathbf{x}) F(\mathbf{x} - \mathbf{v})^T) = |\Theta|(C(\mathbf{v}) + mm^T) .$$

Taking the DFT of both sides, we get

$$\begin{aligned}
\forall \boldsymbol{\xi} \in \Theta\backslash\{0\}, \quad \hat{C}(\boldsymbol{\xi}) &= \frac{1}{|\Theta|}\mathbb{E}\Big(\widehat{F * \tilde{F}^T}(\boldsymbol{\xi})\Big) = \frac{1}{|\Theta|}\mathbb{E}(\hat{F}(\boldsymbol{\xi})\hat{F}(\boldsymbol{\xi})^*) , \\
\hat{C}(0) &= \frac{1}{|\Theta|}\mathbb{E}(\hat{F}(0)\hat{F}(0)^*) - |\Theta| mm^T ,
\end{aligned}$$

or equivalently,

$$\frac{1}{|\Theta|}\mathbb{E}((F-m)*(\tilde{F}-m)^T(\mathbf{v}))=C(\mathbf{v})\ . \tag{2.2}$$

Besides, since $\mathbb{E}(F(\mathbf{x}))=m$ for all $\mathbf{x}$, we have

$$\begin{aligned}\forall\boldsymbol{\xi}\in\Theta\backslash\{0\},\quad \mathbb{E}(\hat{F}(\boldsymbol{\xi}))&=0\ ,\\ \mathbb{E}(\hat{F}(0))&=|\Theta|m\ ,\end{aligned}$$

so that we get the following proposition.

Proposition 2.1.1. *If F is a circular random field on Θ with values in $\mathbb{R}^d$ and with covariance function C, then*

$$\forall\boldsymbol{\xi}\in\Theta,\quad \hat{C}(\boldsymbol{\xi})=\frac{1}{|\Theta|}\mathrm{Cov}(\hat{F}(\boldsymbol{\xi}))=\frac{1}{|\Theta|}\mathbb{E}((\hat{F}(\boldsymbol{\xi})-\mathbb{E}\hat{F}(\boldsymbol{\xi}))(\hat{F}(\boldsymbol{\xi})-\mathbb{E}\hat{F}(\boldsymbol{\xi}))^*)$$

In particular, $\hat{C}(\boldsymbol{\xi})$ belongs to the set $\mathcal{H}_d^{\geqslant 0}$ of Hermitian non-negative matrices of size $d\times d$, and we have $\hat{C}(-\boldsymbol{\xi})=\overline{\hat{C}(\boldsymbol{\xi})}$. The function $\hat{C}:\Theta\longrightarrow\mathcal{H}_d^{\geqslant 0}$ is called the power spectrum of F.

Notice that for all $f,g\in(\mathbb{C}^d)^\Theta$,

$$\begin{aligned}\sum_{\mathbf{x},\mathbf{y}}f(\mathbf{x})^TC(\mathbf{x}-\mathbf{y})\overline{g(\mathbf{y})}&=\frac{1}{|\Theta|}\sum_{\mathbf{x},\mathbf{y}}\sum_{\boldsymbol{\xi}}f(\mathbf{x})^T\hat{C}(\boldsymbol{\xi})e^{i\langle\boldsymbol{\xi},\mathbf{x}-\mathbf{y}\rangle}\overline{g(\mathbf{y})}\\ &=\frac{1}{|\Theta|}\sum_{\boldsymbol{\xi}}\left(\sum_{\mathbf{x}}f(\mathbf{x})e^{i\langle\boldsymbol{\xi},\mathbf{x}\rangle}\right)^T\hat{C}(\boldsymbol{\xi})\overline{\sum_{\mathbf{y}}g(\mathbf{y})e^{i\langle\boldsymbol{\xi},\mathbf{y}\rangle}}\\ &=\frac{1}{|\Theta|}\sum_{\boldsymbol{\xi}}\hat{\tilde{f}}(\boldsymbol{\xi})^T\hat{C}(\boldsymbol{\xi})\overline{\hat{\tilde{g}}(\boldsymbol{\xi})}\ .\end{aligned} \tag{2.3}$$

This property means that the operator associated to the Hermitian form defined by the covariance is block-diagonal in the discrete Fourier basis.

Based on the spectral representation of the covariance function, we can give a convenient definition of the terms “white noise” and “colored noise”.

Definition 2.1.4. A circular random field $F:\Theta\longrightarrow\mathbb{R}^d$ with covariance function C is said to be a white noise if its power spectrum is constant (or equivalently if $C=C(0)\delta_0$). A circular colored noise is a random field $F:\Theta\longrightarrow\mathbb{R}^d$ that can be written $F=m+h*\mathcal{W}$ where $m\in\mathbb{R}^d$, where $h:\Theta\longrightarrow\mathbb{R}^{d\times d}$ is a matrix-valued function and where $\mathcal{W}$ is a $\mathbb{R}^d$-valued white noise on Θ. A circular rank-one colored noise is a random field $F:\Theta\longrightarrow\mathbb{R}^d$ that can be written $F=m+h*W$ where $m\in\mathbb{R}^d$, where $h:\Theta\longrightarrow\mathbb{R}^d$ and where W is a scalar white noise on Θ.

Notice that if $F=m+h*W$ with $m\in\mathbb{R}^d$, $h:\Theta\longrightarrow\mathbb{R}^d$ and with W a scalar white noise, then for each $\boldsymbol{\xi}\in\Theta\backslash\{0\}$, we have

$$\hat{C}(\boldsymbol{\xi})=\mathrm{Var}(W(0))\,\hat{h}(\boldsymbol{\xi})\hat{h}(\boldsymbol{\xi})^*$$

which is a matrix with rank $\leqslant 1$, which justifies the terminology.

The reader must be warned that this definition only concerns the moments of order 2, and is thus quite weak. In particular, this definition does not concern the phase information (argument of the DFT), which has a particular importance in image perception. The next paragraph will introduce some more terminology to refer to random fields that have the least possible structure in the phase information. Let us emphasize on the fact that a white noise has only uncorrelated samples. Following a remark on [Grenander & Rosenblatt 1953, p.2], we will use the expression "*pure white noise*" (due to Tukey) to refer to a random field whose samples are independent. Notice that in the Gaussian case, a white noise is always a pure white noise.

Random Fields with Uniform Phase

Let us introduce two subsets Θ_+, Θ_0 of Θ such that

$$\Theta = \Theta_+ \sqcup (-\Theta_+) \sqcup \Theta_0$$

is a partition of Θ. Setting

$$\boldsymbol{\eta}_x = (-M/2, 0) \quad \boldsymbol{\eta}_y = (0, -N/2) \quad \boldsymbol{\eta}_{xy} = (-M/2, -N/2)\ ,$$

the elements of $\Theta_0 = \Theta \cap \{0, \boldsymbol{\eta}_x, \boldsymbol{\eta}_y, \boldsymbol{\eta}_{xy}\}$ are the frequencies $\boldsymbol{\xi} \in \Theta$ such that $-\boldsymbol{\xi} = \boldsymbol{\xi}$, i.e. $-(\xi_1, \xi_2) = (\xi_1, \xi_2) \mod (M, N)$).

Let us also write $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$. If $u : \Theta \longrightarrow \mathbb{R}$ is a gray-level image, then a phase function for u is any function $\varphi : \Theta \longrightarrow \mathbb{T}$ such that $\hat{u} = |\hat{u}|e^{i\varphi}$. If $\hat{u}(\boldsymbol{\xi}) \neq 0$, the phase coefficient $\varphi(\boldsymbol{\xi}) \in \mathbb{T}$ is uniquely defined as an argument of $\hat{u}(\boldsymbol{\xi}) \in \mathbb{C}^*$ (denoted by $\arg(\hat{u}(\boldsymbol{\xi}))$), while any arbitrary value can be chosen if $\hat{u}(\boldsymbol{\xi}) = 0$.

Definition 2.1.5 ([Galerne *et al.* 2011b])**.** A uniform random phase function is a random function $\psi : \Theta \longrightarrow \mathbb{T}$ such that

- $\forall \boldsymbol{\xi} \in \Theta, \quad \psi(-\boldsymbol{\xi}) = -\psi(\boldsymbol{\xi})$,
- $\psi(\boldsymbol{\xi}) \sim \mathcal{U}(\mathbb{T})$ if $\boldsymbol{\xi} \in \Theta \backslash \Theta_0$ and $\psi(\boldsymbol{\xi}) \sim \mathcal{U}(\{0, \pi\})$ if $\boldsymbol{\xi} \in \Theta_0$,
- $(\psi(\boldsymbol{\xi}))_{\boldsymbol{\xi} \in \Theta_+ \cup \Theta_0}$ are independent.

A random phase field on Θ is a random field $F : \Theta \longrightarrow \mathbb{R}^d$ whose DFT can be written $\hat{F} = \hat{G}e^{i\psi}$ where ψ is a uniform random phase function, where G is a random field on Θ and where ψ and G are independent.

Actually, there is an abuse of terminology in Definition 2.1.5 because a random phase function may not be uniform; in this case we would speak of a structured phase function. In order to avoid ambiguity, one should use the more precise expression "random field with uniform phase" instead of "random phase field".

Proposition 2.1.2. *A random field with uniform phase is circular stationary.*

Proof. Let F such that $\widehat{F} = \widehat{G}e^{i\psi}$, where ψ is a uniform random phase function and G a random field G. Let $\mathbf{v} \in \Theta$ and let us consider

$$\tau_{\mathbf{v}} F(\mathbf{x}) = F(\mathbf{x} - \mathbf{v}) \; .$$

The DFT of $\tau_{\mathbf{v}} F$ is given by

$$\forall \boldsymbol{\xi} \in \Theta, \quad \widehat{\tau_{\mathbf{v}} F}(\boldsymbol{\xi}) = e^{-i\langle \boldsymbol{\xi}, \mathbf{v} \rangle} \widehat{F}(\boldsymbol{\xi}) = e^{-i\langle \boldsymbol{\xi}, \mathbf{v} \rangle + i\psi(\boldsymbol{\xi})} \hat{G}(\boldsymbol{\xi}) \; .$$

Since ψ is a uniform random phase function, so is $\boldsymbol{\xi} \longmapsto \psi(\boldsymbol{\xi}) - \langle \boldsymbol{\xi}, \mathbf{v} \rangle$. Thus $\widehat{\tau_{\mathbf{v}} F}$ has the same distribution as $\widehat{F}$, and therefore, $\tau_{\mathbf{v}} F$ has the same distribution as F, which proves that F is circular stationary. □

Notice that the converse of Proposition 2.1.2 is clearly not true because a random uniform translation of a fixed image is a circular stationary random field. Actually, the phase constraints that are imposed by the stationarity assumption will be later discussed in Subsection 5.1.1.

We will see in the sequel that interesting texture models are obtained as random fields with uniform phase. However, since the geometry of an image is mostly encoded in the phase of its Fourier transform [Oppenheim & Lim 1981], one cannot expect to find any sharp geometrical detail in a random field with uniform phase. This agrees with the next proposition which shows that a Gaussian white noise is a random field with uniform phase.

Proposition 2.1.3 (DFT of a Gaussian white noise)**.** *If W is a Gaussian white noise on Θ, then $\widehat{W}$ is a Gaussian vector of $\mathbb{C}^\Omega \sim (\mathbb{R}^2)^\Omega$ which satisfies*

$$\forall \boldsymbol{\xi} \in \Theta, \quad \widehat{W}(-\boldsymbol{\xi}) = \overline{\widehat{W}(\boldsymbol{\xi})} \; ,$$

and such that the random variables $(\widehat{W}(\boldsymbol{\xi}))_{\Theta_+ \cup \Theta_0}$ are independent. Besides, for each $\boldsymbol{\xi} \in \Theta \backslash \Theta_0$, $\widehat{W}(\boldsymbol{\xi})$ is a Gaussian vector of $\mathbb{C} \sim \mathbb{R}^2$ of covariance $MN \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$ and for $\boldsymbol{\xi} \in \Theta_0$, $\widehat{W}(\boldsymbol{\xi})$ is a real Gaussian random variable of variance MN. In particular, we have

$$\forall \boldsymbol{\xi} \in \Theta, \quad \mathbb{E}\Big(|\widehat{W}(\boldsymbol{\xi})|^2\Big) = MN \; .$$

In particular $\arg(\widehat{W})$ *is a uniform random phase function.*

Proof. $\widehat{W}$ is a Gaussian vector of $(\mathbb{R}^2)^\Omega$ because it is a linear transform of W, and the relation $\widehat{W}(-\boldsymbol{\xi}) = \overline{\widehat{W}(\boldsymbol{\xi})}$ is a well-known property of the DFT. Since W is a Gaussian white noise, its characteristic function is given by

$$\forall \varphi \in \mathbb{R}^\Theta, \quad \mathbb{E}\left[\exp\Big(i \sum_{\mathbf{x} \in \Theta} \varphi(\mathbf{x}) W(\mathbf{x})\Big)\right] = \exp\left(-\frac{1}{2} \sum_{\mathbf{x} \in \Theta} \varphi(\mathbf{x})^2\right) \; .$$

Thanks to Parseval's formula, we thus have for any $\varphi \in \mathbb{R}^\Theta$,

$$\mathbb{E}\left[\exp\left(i \frac{1}{MN} \sum_{\boldsymbol{\xi} \in \Theta} \hat{\varphi}(\boldsymbol{\xi}) \overline{\widehat{W}(\boldsymbol{\xi})}\right)\right] = \exp\left(-\frac{1}{2MN} \sum_{\boldsymbol{\xi} \in \Theta} |\hat{\varphi}(\boldsymbol{\xi})|^2\right) \; .$$

Then, one has

$$\sum_{\boldsymbol{\xi}\in\Theta} \hat{\varphi}(\boldsymbol{\xi})\overline{\widehat{W}(\boldsymbol{\xi})} = \sum_{\boldsymbol{\xi}\in\Theta_+} \hat{\varphi}(\boldsymbol{\xi})\overline{\widehat{W}(\boldsymbol{\xi})} + \overline{\hat{\varphi}(\boldsymbol{\xi})}\widehat{W}(\boldsymbol{\xi}) + \sum_{\boldsymbol{\xi}\in\Theta_0} \hat{\varphi}(\boldsymbol{\xi})\widehat{W}(\boldsymbol{\xi})$$
$$= \sum_{\boldsymbol{\xi}\in\Theta_+} 2\,\mathrm{Re}\left(\hat{\varphi}(\boldsymbol{\xi})\overline{\widehat{W}(\boldsymbol{\xi})}\right) + \sum_{\boldsymbol{\xi}\in\Theta_0} \hat{\varphi}(\boldsymbol{\xi})\widehat{W}(\boldsymbol{\xi})\ ,$$

and

$$\sum_{\boldsymbol{\xi}\in\Theta} |\hat{\varphi}(\boldsymbol{\xi})|^2 = \sum_{\boldsymbol{\xi}\in\Theta_+} 2|\hat{\varphi}(\boldsymbol{\xi})|^2 + \sum_{\boldsymbol{\xi}\in\Theta_0} |\hat{\varphi}(\boldsymbol{\xi})|^2\ .$$

Therefore

$$\mathbb{E}\left[\exp\left(i\frac{1}{MN}\left(\sum_{\boldsymbol{\xi}\in\Theta_+} 2\,\mathrm{Re}\left(\hat{\varphi}(\boldsymbol{\xi})\overline{\widehat{W}(\boldsymbol{\xi})}\right) + \sum_{\boldsymbol{\xi}\in\Theta_0} \hat{\varphi}(\boldsymbol{\xi})\widehat{W}(\boldsymbol{\xi})\right)\right)\right]$$
$$= \exp\left(-\frac{1}{2MN}\left(\sum_{\boldsymbol{\xi}\in\Theta_+} 2|\hat{\varphi}(\boldsymbol{\xi})|^2 + \sum_{\boldsymbol{\xi}\in\Theta_0} |\hat{\varphi}(\boldsymbol{\xi})|^2\right)\right)\ .$$

Since the mapping $\varphi \longmapsto \hat{\varphi}$ defines a map of $\mathbb{R}^{\Theta}$ onto $\mathbb{C}^{\Theta_+} \times \mathbb{R}^{\Theta_0}$, we get that for all $\psi \in \mathbb{C}^{\Theta_+} \times \mathbb{R}^{\Theta_0}$,

$$\mathbb{E}\left[\exp\left(i\sum_{\boldsymbol{\xi}\in\Theta_+} \mathrm{Re}\left(\psi(\boldsymbol{\xi})\overline{\widehat{W}(\boldsymbol{\xi})}\right) + i\sum_{\boldsymbol{\xi}\in\Theta_0} \psi(\boldsymbol{\xi})\widehat{W}(\boldsymbol{\xi})\right)\right]$$
$$= \exp\left(-\frac{1}{2}\left(\sum_{\boldsymbol{\xi}\in\Theta_+} \frac{MN}{2}|\psi(\boldsymbol{\xi})|^2 + MN\sum_{\boldsymbol{\xi}\in\Theta_0} \psi(\boldsymbol{\xi})^2\right)\right)\ .$$

The left-hand side is exactly the characteristic function of the Gaussian vector $(\widehat{W}(\boldsymbol{\xi}))_{\boldsymbol{\xi}\in\Theta_+\cup\Theta_0}$ of $(\mathbb{C})_+^{\Theta} \times \mathbb{R}^{\Theta_0}$, so that the last formula gives the desired independence property and the marginal distributions. □

Remark 2.1.1. The last proof is based on the characteristic function of $\widehat{W}$. Let us mention that a more common proof of this result (which is given in [Desolneux *et al.* 2015]) amounts to compute the correlations between all the components of the random vector

$$\left(\ \mathrm{Re}\left(\widehat{W}(\boldsymbol{\xi})\right)\ ,\ \mathrm{Im}\left(\widehat{W}(\boldsymbol{\xi})\right)\ ,\ \boldsymbol{\xi}\in\Theta\ \right)\ .$$

2.1.2 Circular Discrete Spot Noise

Here again, the definitions and results that are gathered in this subsection are quite standard [Papoulis 1971], [Van Wijk 1991], [Galerne *et al.* 2011b]. However, since we focus on discrete random fields, several properties can be simplified (and in particular, we provide an adapted proof for the Gaussian convergence of high-intensity discrete spot noise).

Definition, covariance

Let $h : \Theta \longrightarrow \mathbb{R}^d$ (which will be often referred to as the kernel function) and let us denote $\sum h = \sum_{\mathbf{x}\in\Theta} h(\mathbf{x})$. Let $\lambda > 0$ and let Π_λ denote a Poisson point process on Θ with intensity λ. In this case where Θ is finite, the Poisson point process Π_λ can be represented by a finite sequence $(\mathbf{X}_i)_{1\leqslant i\leqslant N_\lambda}$ of random points which are independent and uniformly distributed over Θ, and where the total number of points N_λ follows the Poisson distribution $\mathcal{P}(\lambda|\Theta|)$. For $\mathbf{x} \in \Theta$, let us introduce

$$P_\lambda(\mathbf{x}) = |\{\ i \in \{1, \dots, N_\lambda\}\ \text{ such that }\ \mathbf{X}_i = \mathbf{x}\}|$$

which follows the Poisson distribution $\mathcal{P}(\lambda)$.

Definition 2.1.6. The circular discrete spot noise (DSN) (or circular discrete Poisson spot noise) of intensity λ associated to h is the random field $F^\Theta_{h,\lambda} : \Theta \longrightarrow \mathbb{R}^d$ defined by

$$\forall \mathbf{x} \in \Theta, \quad F^\Theta_{h,\lambda}(\mathbf{x}) = \sum_{i=1}^{N_\lambda} h(\mathbf{x} - \mathbf{X}_i)\ . \tag{2.4}$$

Grouping the X_i's by their location, this rewrites

$$\forall \mathbf{x} \in \Theta, \quad F^\Theta_{h,\lambda}(\mathbf{x}) = \sum_{\mathbf{y}\in\Theta} h(\mathbf{x} - \mathbf{y})P_\lambda(\mathbf{y})\ ,$$

so that $F^\Theta_{h,\lambda} = h * P_\lambda$. Notice that P_λ is a Poisson white noise, because the $P_\lambda(\mathbf{x})$ are independent and follow the distribution $\mathcal{P}(\lambda)$. In particular, $F^\Theta_{h,\lambda}$ is a rank-one colored noise in the sense of Definition 2.1.4.

The random field $F^\Theta_{h,\lambda}$ is of second-order and circular stationary. Its expectation is given by

$$m = \mathbb{E}(F^\Theta_{h,\lambda}(\mathbf{x})) = \lambda \sum_{\mathbf{x}\in\Theta} h(\mathbf{x})\ .$$

Also, one has

$$F^\Theta_{h,\lambda} - m = h * (P_\lambda - \lambda)\ .$$

Since P_λ is a white noise, we get that the covariance function of $F^\Theta_{h,\lambda}$ is

$$C_{F^\Theta_{h,\lambda}} = \lambda h * \tilde{h}^T$$

which is the autocorrelation of h multiplied by the intensity λ.

Gaussian convergence at high intensity

Definition 2.1.7. The renormalized circular discrete spot noise of intensity λ associated to h is the random field $G^\Theta_{h,\lambda} : \Theta \longrightarrow \mathbb{R}$ defined by

$$G^\Theta_{h,\lambda} = \frac{F^\Theta_{h,\lambda} - \mathbb{E}(F^\Theta_{h,\lambda})}{\sqrt{\lambda}} = \frac{1}{\sqrt{\lambda}}\left(h * P_\lambda - \lambda \sum h\right)\ .$$

It is a circular stationary random field with mean 0 and covariance function $h * \tilde{h}^T$.

Theorem 2.1.1 ([Papoulis 1971]). *When $\lambda \to \infty$, $G^{\Theta}_{h,\lambda}$ converges in distribution to a Gaussian random field of mean 0 and covariance function $h * \tilde{h}^T$.*

Proof. If N_λ follows the Poisson distribution $\mathcal{P}(\lambda)$ with parameter λ, then we first show with Lévy's theorem that

$$\frac{N_\lambda - \lambda}{\sqrt{\lambda}} \xrightarrow[\lambda\to\infty]{(d)} \mathcal{N}(0,1) \ ,$$

meaning that $\frac{N_\lambda - \lambda}{\sqrt{\lambda}}$ converges in distribution to $\mathcal{N}(0,1)$. Indeed, computing the characteristic functions gives for all $s \in \mathbb{R}$,

$$\begin{aligned}
\mathbb{E}\left[\exp\left(\frac{N_\lambda - \lambda}{\sqrt{\lambda}}\ s\right)\right] &= \exp\left(-is\sqrt{\lambda}\right)\ \mathbb{E}\left[\exp\left(iN_\lambda \frac{s}{\sqrt{\lambda}}\right)\right] \\
&= \exp\left(-is\sqrt{\lambda}\right) \exp\left(\lambda(e^{i\frac{s}{\sqrt{\lambda}}} - 1)\right) \\
&= \exp\left(\lambda\left(e^{i\frac{s}{\sqrt{\lambda}}} - 1 - i\frac{s}{\sqrt{\lambda}}\right)\right) \\
&= \exp\left(-\frac{s^2}{2} + o(1)\right) \ ,
\end{aligned}$$

which converges to the characteristic function $e^{-s^2/2}$ of $\mathcal{N}(0,1)$ when $\lambda \to \infty$.

Subsequently, since the $P_\lambda(\mathbf{x})$ are independent with distribution $\mathcal{P}(\lambda)$, we get

$$\frac{P_\lambda - \lambda}{\sqrt{\lambda}} \xrightarrow[\lambda\to\infty]{(d)} W$$

where W is a Gaussian white noise on Θ with mean 0 variance 1. Since the convolution by h is a continuous function from $\mathbb{R}^\Theta$ to $(\mathbb{R}^d)^\Theta$, we get

$$G^{\Theta}_{h,\lambda} = h * \frac{P_\lambda - \lambda}{\sqrt{\lambda}} \xrightarrow[\lambda\to\infty]{(d)} h * W \ .$$

□

Definition 2.1.8. The circular asymptotic discrete spot noise (circular ADSN) on Θ associated to the kernel function h is $\mathcal{N}(0, h * \tilde{h}^T)$. Notice that this random field can be obtained as

$$G^{\Theta}_{h} = h * W$$

where W is a scalar Gaussian white noise on Θ with variance 1.

Notice that G^{Θ}_{h} is another example of rank-one colored noise. More precisely, the circular ADSN fields are exactly the Gaussian rank-one colored noises with zero mean.

Simulation and first examples

The simulation of circular DSN and ADSN fields is straightforward. Indeed, it amounts to convolve, respectively, a Poisson white noise or a Gaussian white noise with the kernel function h. Such white noise processes can be drawn thanks to standard sampling techniques and the convolution can be performed in the Fourier domain with complexity $\mathcal{O}(MN \log(MN))$. Therefore, this scheme, called spectral simulation, allows to draw circular DSN or ADSN fields with complexity $\mathcal{O}(MN \log(MN))$.

When the intensity λ is low, the circular DSN can be obtained by a direct summation method. It amounts to draw the number $P \sim \mathcal{P}(\lambda|\Theta|)$ of points of the Poisson process, to draw uniformly and independently P points in Θ and to perform the convolution in the spatial domain using (2.4). This method, which has a mean complexity of $\lambda|\mathrm{Supp}(h)||\Theta|$, is faster than the spectral method for a very low intensity. We will discuss it more thoroughly in Section 3.3.

One can see in Fig. 2.1 some examples of DSN associated to two different kernel functions, and with different intensities. The visual convergence of the DSN $G^{\Theta}_{h,\lambda}$ to its Gaussian limit G^{Θ}_{h} agrees with the result of Theorem 2.1.1. The aspect of the limiting random field depends on the kernel function h since the covariance function of the ADSN is $h * \tilde{h}^T$. When h is a disc kernel, the ADSN is isotropic; when h is an elongated blob, one can observe some linear correlations in the ADSN. As shown in [Van Wijk 1991], one can imagine several shapes of kernels leading to ADSN fields with very different aspects. This experiment gives a first insight in the richness of the ADSN model. However, as one can see on Fig 2.1, the textures obtained by ADSN synthesis do not have any salient features; in particular the ADSN model is not able to synthesize textures with sharp edges. The study of the Fourier transform of the ADSN field that is proposed below will help to understand this limitation.

But before, let us add some comments about the visualization of DSN processes. In this thesis, the renormalized DSN and ADSN have zero mean whereas the set of RGB values which are traditionally used in the visualization process is included in $[0, 255]^3$. Therefore, when we show examples of DSN and ADSN, a mean value m will always be added (for example, in the first example of Fig. 2.1, the mean value is gray, and in the second example, the mean value is purple). In other words, we always show $m + G^{\Theta}_{h,\lambda}$ or $m + G^{\Theta}_{h}$. In general, next to the DSN or ADSN, we show the underlying kernel h. Again the true image that is shown is $h_{visu} = m + sh$ where $s > 0$, so that the mean color m can be seen outside the support of the kernel function. The choice of s is less obvious. If a synthesis domain Θ of size $M \times N$ is fixed, we choose to set $s = \sqrt{MN}$ so that

$$\frac{1}{MN} \sum_{\mathbf{x} \in \Theta} (h_{visu}(\mathbf{x}) - m)(h_{visu}(\mathbf{x}) - m)^T$$

equals the marginal covariance $h * \tilde{h}^T(0)$ of the DSN and ADSN associated to h.

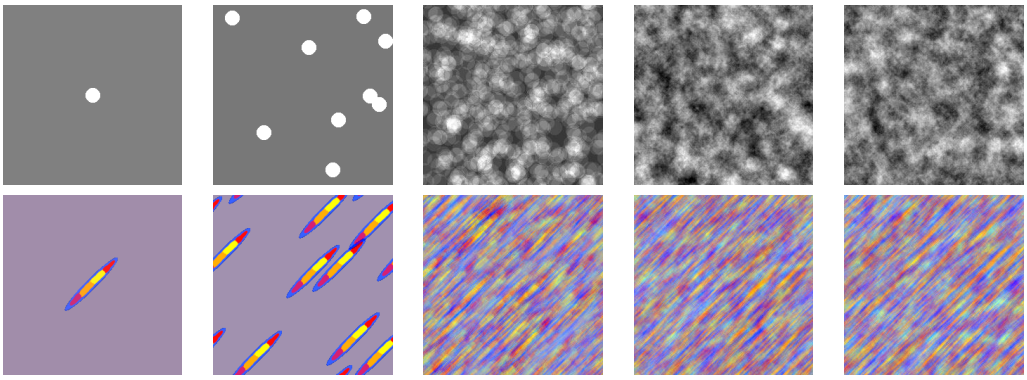


Figure 2.1: **Examples of circular discrete spot noise.** For each row, and from left to right, one can see a compactly-supported kernel function h, realizations of the renormalized circular DSN $G^{\Theta}_{h,\lambda}$ associated to h with intensities $\lambda = 10^{-4}, 10^{-2}, 1$, and a realization of the corresponding ADSN G^{Θ}_{h} . At low intensity λ, the translated copies of the kernel function are clearly visible. When the intensity λ increases, the DSN $G^{\Theta}_{h,\lambda}$ begins to look like the uniform random phase field G^{Θ}_{h} which does not present any salient feature.

Spectral representation

The discrete Fourier transform of the circular ADSN is

$$\widehat{G^{\Theta}_{h}} = \widehat{h * W} = \hat{h}\widehat{W} \, . \tag{2.5}$$

In particular, for each $\boldsymbol{\xi} \in \Theta$, $\widehat{G^{\Theta}_{h}}(\boldsymbol{\xi})$ belongs to the plane $\mathbb{C}\hat{h}(\boldsymbol{\xi})$.

Equation (2.5) together with Proposition 2.1.3 give the distribution of $\widehat{G^{\Theta}_{h}}$. In particular, the DFT coefficients of the circular ADSN G^{Θ}_{h} are independent modulo the Hermitian symmetry, and

$$\forall \boldsymbol{\xi} \in \Theta, \quad \mathbb{E}\Big(\widehat{G^{\Theta}_{h}}(\boldsymbol{\xi})\Big) = 0 \quad \text{and} \quad \mathbb{E}\Big(\widehat{G^{\Theta}_{h}}(\boldsymbol{\xi})\widehat{G^{\Theta}_{h}}(\boldsymbol{\xi})^{*}\Big) = MN\hat{h}(\boldsymbol{\xi})\hat{h}(\boldsymbol{\xi})^{*} \, .$$

Besides, G^{Θ}_{h} is a random field with uniform phase according to Definition 2.1.5. This confirms that one cannot expect to perceive any salient element in a circular ADSN field, as can be observed in Fig. 2.1.

Notice however that the phase of a Poisson white noise $\arg(\hat{P}_{\lambda})$ is not a uniform random phase field. A fortiori, the circular DSN, whose DFT is given by

$$\widehat{F^{\Theta}_{h,\lambda}}(\boldsymbol{\xi}) = \widehat{P_{\lambda}}(\boldsymbol{\xi})\hat{h}(\boldsymbol{\xi})$$

is not a random field with uniform phase. Indeed, for a DSN with low intensity, the geometrical features of the kernel function are still clearly visible; this agrees with the fact that a DSN is not a random field with uniform phase.

Binomial spot noise model

We have chosen here to study the Poisson spot noise because it allows for a straightforward generalization to $\mathbb{Z}^2$ as will be seen in Section 2.2. But, as in [Galerne *et al.* 2011b], one can also define a binomial spot noise model with a non-random total number of spots. Indeed, for $p \geqslant 1$, one can consider

$$\forall \mathbf{x} \in \Theta, \quad F_p(\mathbf{x}) = \sum_{i \geqslant 1}^{p} h(\mathbf{x} - \mathbf{X}_i) \ ,$$

where $\mathbf{X}_1, \dots, \mathbf{X}_p$ are p random independent points which follow the uniform distribution on Θ. We call it binomial spot noise because we still have $F_p = B_p * h$ where

$$B_p(\mathbf{x}) = |\{i \ \text{ such that } \mathbf{X}_i = \mathbf{x}\}|$$

follows the binomial distribution of parameters p and $\frac{1}{MN}$. Again it is clear that F_p is a circular stationary random field. Denoting $\bar{h} = \frac{1}{MN}\sum h$, we have,

$$\mathbb{E}(h(\mathbf{x} - \mathbf{X}_i)) = \frac{1}{MN}\sum h = \bar{h} \ ,$$

$$\mathrm{Cov}(h(\mathbf{x} - \mathbf{X}_i)) = \frac{1}{MN}\sum_{\mathbf{z}\in\Theta}(h(\mathbf{z}+\mathbf{t}) - \bar{h})(h(\mathbf{z}) - \bar{h})^T = \frac{1}{MN}(h - \bar{h}) * (\tilde{h} - \bar{h})^T(\mathbf{t}) \ ,$$

and since the $\mathbf{x} \longmapsto h(\mathbf{x} - \mathbf{X}_i)$ are i.i.d. random vectors of $(\mathbb{R}^d)^\Theta$, we get

$$\mathbb{E}(F_p(\mathbf{x})) = \frac{p}{MN}\sum h \ ,$$

$$C_{F_p} = \frac{p}{MN}(h - \bar{h}) * (\tilde{h} - \bar{h}) \ .$$

Notice that taking $h = \delta_0$ in the last equality shows that B_p is not a white noise.

Therefore, the classical central limit theorem shows that when $p \to \infty$,

$$\frac{F_p - \frac{p}{MN}\sum h}{\sqrt{\frac{p}{MN}}} \xrightarrow{(d)} \mathcal{N}\Big(0, (h - \bar{h}) * (\tilde{h} - \bar{h})^T\Big) \ ,$$

which means that the binomial spot noise model converges to the circular ADSN associated to $h - \bar{h}$.

Notice that the covariance of an (asymptotic) binomial spot noise always has zero sum, which is not a natural constraint. This is another reason to prefer *a priori* the Poisson spot noise.

Assuming that $\bar{h} = 0$, one could think that F_p is the conditional version of the Poisson spot noise given that the total number of spots is exactly p. Actually, there is a slight difference in the normalization which makes this fact true only if the intensity is properly set. Indeed, one can observe that the binomial spot noise F_p with p spots is the Poisson spot noise of intensity $\frac{p}{MN}$ conditioned by the fact that the total number of spots is exactly p. However, if we draw $P \sim \mathcal{P}(\lambda MN)$ and if for each $n \geqslant 1$ we draw independently a binomial spot noise F_n with n spots, then it is not true that the composition F_P is a Poisson spot noise of intensity λ. Indeed, the binomial and Poisson spot noises are not equally normalized: in the case of the Poisson spot noise, the total number of spots does not appear in the normalization.

2.1.3 Random Phase Noise

In the last paragraph we have seen that the spectral representation of the circular ADSN writes

$$\widehat{G_h^\Theta} = \hat{h}\widehat{W}$$

where $\mathbb{E}(\widehat{W}(\boldsymbol{\xi})) = 0$ and $\mathbb{E}(|\widehat{W}(\boldsymbol{\xi})|^2) = MN$. One can obtain a similar random field by convolving h by a white noise whose DFT modulus is not random leading to the following definition.

Definition 2.1.9 ([Galerne *et al.* 2011b])**.** If $h : \Theta \longrightarrow \mathbb{R}^d$ is a kernel function, the random phase noise (RPN) $R_h : \Theta \longrightarrow \mathbb{R}^d$ associated to h is defined in Fourier domain by

$$\widehat{R_h} = \sqrt{MN}\,\hat{h}\,e^{i\psi}\ ,$$

where ψ is a random phase function in the sense of Definition 2.1.5. Equivalently, thanks to the inversion formula, we have

$$\forall \mathbf{x} \in \Theta, \quad R_h(\mathbf{x}) = \frac{1}{\sqrt{MN}} \sum_{\boldsymbol{\xi}\in\Theta} \hat{h}(\boldsymbol{\xi})\, e^{i\langle \boldsymbol{\xi},\mathbf{x}\rangle + i\psi(\boldsymbol{\xi})}\ . \tag{2.6}$$

Notice that for each $\boldsymbol{\xi} \in \Theta$, $\widehat{R_h}(\boldsymbol{\xi})$ belongs to the circle

$$\{\sqrt{MN}\hat{h}(\boldsymbol{\xi})e^{i\varphi}\ ;\ \varphi \in \mathbb{T}\}\ .$$

The random phase noise is another example of rank-one colored noise. Indeed, by definition, one can write

$$R_h = h * R_{\delta_0},$$

with $\widehat{R_{\delta_0}} = \sqrt{MN}e^{i\psi}$. Since $\widehat{R_{\delta_0}}$ has a constant modulus, R_{δ_0} is a white noise. More precisely, R_h is by definition a random field with uniform phase. In particular, Proposition 2.1.2 ensures that R_h is circular stationary.

The Definition 2.1.9 that was adopted here differs from the one given in [Galerne *et al.* 2011b] because of the factor $\sqrt{MN}$. In contrast, the factor $\sqrt{MN}$ appears in Theorem 2 of [Galerne *et al.* 2011b] about the simulation of the ADSN. With Definition 2.1.9, we will see that, as for the ADSN field, the first-order moments of the RPN are equal to the (non-normalized) autocorrelation of h. Besides, with the previous notation, the RPN field associated to an image $u : \Theta \longrightarrow \mathbb{R}^d$ according to [Galerne *et al.* 2011b] is R_{t_u} where

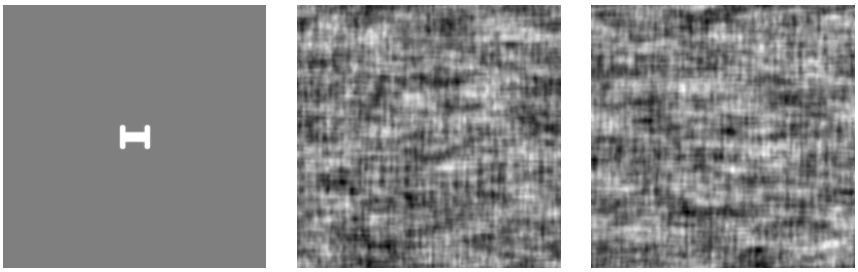
$$t_u = \frac{1}{\sqrt{MN}}(u - \bar{u}) \quad \text{and} \quad \bar{u} = \frac{1}{MN}\sum_{\mathbf{x}} u(\mathbf{x})\ .$$

This choice will also be justified again in Chapter 3.

Let us also mention that the RPN model also appears in the physics community under the name "surrogate" [Theiler *et al.* 1992], [Schreiber & Schmitz 2000], [Borgnat *et al.* 2010].

Since $\mathbb{E}(e^{i\psi(\boldsymbol{\xi})}) = 0$, the random field R_h has zero mean. Besides, we have

$$\widehat{R_h}\widehat{R_h}^* = MN\,\hat{h}(\boldsymbol{\xi})\hat{h}(\boldsymbol{\xi})^*\ ,$$



h R_h G_h^Θ

Figure 2.2: **Comparison between circular ADSN and RPN.** From left to right, one can see a kernel function h, the corresponding RPN R_h and ADSN G_h^Θ. The RPN and ADSN which are two uniform random phase fields in the sense of 2.1.5 have perceptually similar realizations.

so that, by inverse DFT, we get

$$R_h * (\widetilde{R_h})^T = MN\ h * \tilde{h}^T\ ,$$

which shows that R_h has a deterministic autocorrelation function. Recalling that

$$C_{R_h}(\mathbf{v}) = \frac{1}{MN}\mathbb{E}\left(R_h * (\widetilde{R_h})^T(\mathbf{v})\right)\ ,$$

we get the following result.

Proposition 2.1.4 ([Galerne *et al.* 2011b])**.** *The random phase noise R_h is a circular stationary random field with zero mean and covariance function $h * \tilde{h}$.*

Let us notice that for a circular stationary random field $F : \Theta \longrightarrow \mathbb{R}^d$, it is equivalent to impose that the autocorrelation $F * \tilde{F}^T$ is deterministic, and that $\widehat{F}\widehat{F}^*$ is deterministic. In particular, for $d = 1$, the autocorrelation is deterministic if and only if the DFT modulus is deterministic. Notice however that a circular stationary random field F with a deterministic autocorrelation is not necessarily a RPN. Indeed, for example with $d = 1$, this only imposes that $\widehat{F} = Me^{i\psi}$ where M is deterministic and ψ is a random phase function which is not necessarily uniform. For example, if $f : \Theta \longrightarrow \mathbb{R}$ and if $\mathbf{X}$ is a uniform point of Θ, then the translation $\tau_{\mathbf{X}} f$ of vector $\mathbf{X}$ has the same autocorrelation as f, but is not a RPN in general.

The last proposition shows that the RPN and ADSN fields associated to a kernel h have the same moments of order 1 and 2. Generally speaking, one can see that the textures that are synthesized with these models are perceptually similar (see Fig. 2.2 or the several comparisons of [Galerne *et al.* 2011b]). This empirical observation is confirmed by the fact that both RPN and ADSN fields are uniform random phase fields in the sense of Definition 2.1.5.

One can wonder if it is more interesting to work with RPN or ADSN fields. The RPN field has a simpler DFT, but its major drawback is that its spatial distribution is not explicit. In particular, we will see in Chapter 4 that the study of the total variation of uniform random phase fields is easier in the case of ADSN fields. However, thanks to Equation (2.6), one can show that the marginal distribution of

a RPN is approximately Gaussian, using central-limit theorems for non identically distributed random variables. But the approximate Gaussianity of the whole RPN process is a more difficult question.

Another strong reason to prefer ADSN fields is that they can also be defined on the discrete plane $\mathbb{Z}^2$, as we will see in Subsection 2.2.2. In contrast, as will be discussed in Subsection 2.2.3, it is not straighforward to extend the RPN model as a random field on $\mathbb{Z}^2$.

2.1.4 Circular stationary Gaussian Models and ADSN

We have seen that a circular ADSN is a circular stationary Gaussian random field. Conversely, one can wonder if every circular stationary Gaussian random field with zero mean is a circular ADSN. It is well-known to be true in the gray-level case $d = 1$ [Xia *et al.* 2014] and we provide here a short proof. However, this property is also known to be false in the color case $d > 1$. To cope with that, in this subsection we propose to decompose any $\mathbb{R}^d$-valued stationary Gaussian random field into a sum of d independent ADSN.

Proposition 2.1.5. *Let $G : \Theta \longrightarrow \mathbb{R}$ be a real-valued circular stationary random field. We suppose that G follows a Gaussian distribution with mean m. Then, there exists $h : \Theta \longrightarrow \mathbb{R}$ such that F has the same distribution as $m + G_h^\Theta$.*

Proof. Let us denote by C the covariance function of G. Since G and $m + G_h^\Theta$ are Gaussian random fields with mean m, we only have to prove that there exists h such that $C = h * \tilde{h}$. But, thanks to Proposition 2.1.1, we know that $\widehat{C} \geqslant 0$ so that we can define a kernel function $h : \Theta \longrightarrow \mathbb{R}$ by its DFT

$$\forall \boldsymbol{\xi} \in \Theta, \quad \hat{h}(\boldsymbol{\xi}) = \sqrt{\hat{C}(\boldsymbol{\xi})} \, .$$

We thus have $\widehat{C} = |\hat{h}|^2$ which is equivalent to $C = h * \tilde{h}$. □

Remark 2.1.2. The kernel built in the last proof is actually the canonical texton associated to the Gaussian model $\mathcal{N}(m, C)$, see Subsection 2.1.5.

For $d > 1$, we have seen that the circular ADSN G_h^Θ satisfies

$$\forall \boldsymbol{\xi} \in \Theta, \quad \widehat{G_h^\Theta}(\boldsymbol{\xi}) \in \mathbb{C}\hat{h}(\boldsymbol{\xi})$$

almost surely. Also,

$$\forall \boldsymbol{\xi} \in \Theta, \quad \widehat{C}(\boldsymbol{\xi}) = \hat{h}(\boldsymbol{\xi})\hat{h}(\boldsymbol{\xi})^* \, , \tag{2.7}$$

and thus $\hat{C}(\boldsymbol{\xi})$ is a matrix of rank 1 if $\hat{h}(\boldsymbol{\xi}) \neq 0$ and 0 otherwise.

Therefore, it is not difficult to construct non-ADSN circular stationary Gaussian fields. For example, let us consider a multi-channel white noise $\mathcal{W} : \Theta \longrightarrow \mathbb{R}^d$, which means that the $\mathcal{W}(\mathbf{x})$ are independent Gaussian vectors of $\mathbb{R}^d$ with covariance I_d. Equivalently, the components $W_1, \ldots, W_d$ of $\mathcal{W}$ are d independent scalar Gaussian white noises. Thus, $\widehat{W}_1, \ldots, \widehat{W}_d$ are d independent complex Gaussian white noises

in the sense of Proposition 2.1.3. Consequently, for all $\boldsymbol{\xi} \in \Theta$, the distribution of $\widehat{\mathcal{W}}(\boldsymbol{\xi})$ is not supported by a subspace of $\mathbb{C}$-dimension 1, and also

$$\widehat{C_{\mathcal{W}}}(\boldsymbol{\xi}) = \mathbb{E}(\widehat{\mathcal{W}}(\boldsymbol{\xi})\widehat{\mathcal{W}}(\boldsymbol{\xi})^*) = \begin{pmatrix} \mathbb{E}(|\widehat{W}_1(\boldsymbol{\xi})|^2) & & 0 \\ & \ddots & \\ 0 & & \mathbb{E}(|\widehat{W}_d(\boldsymbol{\xi})|^2) \end{pmatrix} = MNI_d \quad (2.8)$$

is not of rank one. This shows that $\mathcal{W}$ is not an ADSN.

However, the next proposition shows that a $\mathbb{R}^d$-valued circular stationary Gaussian random field can be written as a sum of d ADSN fields.

Theorem 2.1.2. *Let $G : \Theta \longrightarrow \mathbb{R}^d$ be a circular stationary Gaussian random field with mean m. Then there exists d kernel functions $h_1, \ldots, h_d$ such that*

$$G \sim m + G_1 + \ldots + G_d$$

where $G_1, \ldots, G_d$ are independent circular ADSN associated to $h_1, \ldots, h_d$.

Proof. Since G and $m + G_1 + \ldots + G_d$ are Gaussian with mean m, we only have to show that the covariance function C of G can be written as

$$C = h_1 * \tilde{h}_1^T + \ldots + h_d * \tilde{h}_d^T \, . \quad (2.9)$$

Let $\boldsymbol{\xi} \in \Theta$. Proposition 2.1.1 shows that $\widehat{C}(\boldsymbol{\xi})$ is a Hermitian non-negative matrix. Therefore its non-negative eigenvalues can be written $a_1^2(\boldsymbol{\xi}), \ldots, a_d^2(\boldsymbol{\xi})$, and there exists a unitary matrix $U(\boldsymbol{\xi})$ such that

$$\widehat{C}(\boldsymbol{\xi}) = U(\boldsymbol{\xi}) \begin{pmatrix} a_1^2(\boldsymbol{\xi}) & & 0 \\ & \ddots & \\ 0 & & a_d^2(\boldsymbol{\xi}) \end{pmatrix} U(\boldsymbol{\xi})^* \, .$$

If $q_j(\boldsymbol{\xi})$ denotes the j-th column of $U(\boldsymbol{\xi})$ multiplied by $a_j(\boldsymbol{\xi})$, the last equality can be written

$$\widehat{C}(\boldsymbol{\xi}) = q_1(\boldsymbol{\xi})q_1^*(\boldsymbol{\xi}) + \ldots + q_d(\boldsymbol{\xi})q_d^*(\boldsymbol{\xi}) \, .$$

Since we have $\widehat{C}(-\boldsymbol{\xi}) = \overline{\widehat{C}(\boldsymbol{\xi})}$, we can make a global choice for $(q_j(\boldsymbol{\xi}), \boldsymbol{\xi} \in \Theta)$ such that $q_j(-\boldsymbol{\xi}) = \bar{q}_j(\boldsymbol{\xi})$. Setting $\hat{h}_j = q_j$ we obtain real-valued functions h_j such that

$$\widehat{C} = \hat{h}_1\hat{h}_1^* + \ldots + \hat{h}_d\hat{h}_d^* \, .$$

Then, taking the inverse DFT, we get (2.9) as expected. □

Corollary 2.1.3. *A circular Gaussian model $\mathcal{N}(m, C)$ is an ADSN if and only if the rank of $\widehat{C}(\boldsymbol{\xi})$ is $\leqslant 1$ for all $\boldsymbol{\xi} \in \Theta$.*

Proof. The direct statement has already been proved before (cf. Equation (2.7)). For the converse, one can begin as in the proof of Theorem 2.1.2, but in this case it is possible to write

$$\widehat{C}(\boldsymbol{\xi}) = q(\boldsymbol{\xi})q(\boldsymbol{\xi})^* \, .$$

Following the end of the proof, it is thus possible to write $C = h * \tilde{h}$ for $h : \Theta \longrightarrow \mathbb{R}^d$. □

Let us remark that the decomposition of Theorem 2.1.2 is not unique. But, observing the proof, in the case where each matrix $\widehat{C}(\boldsymbol{\xi})$ has d distinct eigenvalues, one can give a canonical decomposition by requiring the eigenvalues $a_1^2, \ldots, a_d^2$ to be sorted in decreasing order

$$a_1^2 > \ldots > a_d^2 \geqslant 0$$

(of course, the $\hat{h}_j(\boldsymbol{\xi})$ are uniquely defined up to a complex factor of modulus 1). We will see another way to interpret this canonical decomposition at the end of Subsection 2.1.6.

Let us give a last remark about this decomposition. In the proof of Theorem 2.1.2, we introduced a unitary matrix $U(\boldsymbol{\xi})$ whose columns are the eigenvectors of $\widehat{C}(\boldsymbol{\xi})$. In general, the eigenvectors of $\widehat{C}(\boldsymbol{\xi})$ for different frequencies $\boldsymbol{\xi}$ will not be related (except in the case of two opposite frequencies). But, if there exists a common eigenvector basis corresponding to a unitary matrix U, then U gives also an eigenvector basis for

$$\frac{1}{MN} \sum_{\boldsymbol{\xi} \in \Theta} \widehat{C}(\boldsymbol{\xi}) = C(0)$$

which is the marginal covariance of the random field. Thus, if U is the orthogonal matrix whose columns are the eigenvectors of $C(0)$, then U is a natural candidate for the diagonalization of the matrices $\widehat{C}(\boldsymbol{\xi})$. The practical interest of the diagonalization of the matrices $\widehat{C}(\boldsymbol{\xi})$ will be questioned in Subsection 3.2.3.

2.1.5 Texton of a circular Gaussian Model

A general Gaussian random vector of mean m and covariance Γ can be sampled as $m + \Gamma^{1/2} V$ where V is a Gaussian white noise and where $\Gamma^{1/2}$ is the matrix square root of Γ (seen as a symmetrical non-negative matrix). Therefore, the simulation of a Gaussian random vector is simple as soon as the computation of the covariance square root is tractable. In the case of $\mathbb{R}^d$-valued random fields on Θ, the covariance can be seen as a matrix of size $d|\Theta| \times d|\Theta|$, which is prohibitive if only because of the storage limitation (for a gray-level image of size 1000×1000, the covariance matrix would be of size $10^6 \times 10^6$).

Hopefully, it is well-known [Wood & Chan 1994] that for circular stationary Gaussian random fields, covariance square roots can be naturally replaced by convolution operators.

Proposition 2.1.6 ([Xia *et al.* 2014])**.** *Let us consider a circular stationary Gaussian random field with mean m and covariance function C. There exists a function $t : \Theta \longrightarrow \mathbb{R}^{d \times d}$ such that $C = t * \tilde{t}^T$. Besides, if $\mathcal{W} : \Theta \longrightarrow \mathbb{R}^d$ is a multi-channel Gaussian white noise (meaning that the $\mathcal{W}(\mathbf{x})$ are independent Gaussian vectors of $\mathbb{R}^d$ with covariance I_d), then*

$$m + t * \mathcal{W} \sim \mathcal{N}(m, C) \; .$$

Such a function t is called a matrix texton *of the model $\mathcal{N}(m, C)$.*

Proof. Let us recall the decomposition (2.9) given in Theorem 2.1.2:

$$C = h_1 * \tilde{h}_1^T + \ldots + h_d * \tilde{h}_d^T \ ,$$

where $h_1, \ldots, h_d : \Theta \longrightarrow \mathbb{R}^d$. For each $\mathbf{x} \in \Theta$, let us write $t(\mathbf{x}) = (h_1(\mathbf{x}); \ldots; h_d(\mathbf{x}))$ the matrix whose columns are $h_1(\mathbf{x}), \ldots, h_d(\mathbf{x})$. Then

$$t * \tilde{t}^T = \begin{pmatrix} h_1 & \ldots & h_d \end{pmatrix} * \begin{pmatrix} \tilde{h}_1^T \\ \vdots \\ \tilde{h}_d^T \end{pmatrix} = h_1 * \tilde{h}_1^T + \ldots + h_d * \tilde{h}_d^T = C \ .$$

For the last assertion, since $\widehat{t * \mathcal{W}} = \hat{t}\widehat{W}$ and recalling Equation (2.8), we get

$$\begin{aligned} \widehat{C_{t*\mathcal{W}}}(\boldsymbol{\xi}) &= \frac{1}{MN}\mathbb{E}\Big(\widehat{t * \mathcal{W}}(\boldsymbol{\xi})\overline{\widehat{t * \mathcal{W}}(\boldsymbol{\xi})}^*\Big) \\ &= \frac{1}{MN}\hat{t}(\boldsymbol{\xi})\mathbb{E}(\widehat{\mathcal{W}}(\boldsymbol{\xi})\overline{\widehat{\mathcal{W}}(\boldsymbol{\xi})}^*)\hat{t}(\boldsymbol{\xi})^* \\ &= \frac{1}{MN}\hat{t}(\boldsymbol{\xi})(MNI_d)\hat{t}(\boldsymbol{\xi})^* = \hat{t}(\boldsymbol{\xi})\hat{t}(\boldsymbol{\xi})^* \ , \end{aligned}$$

and thus, by inverse DFT, $C_{t*\mathcal{W}} = t * \tilde{t}^T = C$ so that $m + t * \mathcal{W} \sim \mathcal{N}(m, C)$. □

Notice that in the case of a circular ADSN field with kernel function h, by definition the covariance function is given by $h * \tilde{h}^T$ where the values $h(\mathbf{x})$ are not matrices but column vectors, leading to the following definition.

Definition 2.1.10 ([Desolneux *et al.* 2012], [Xia *et al.* 2014])**.** A texton associated to a circular ADSN field with kernel function h is any function $t : \Theta \longrightarrow \mathbb{R}^d$ generating the same circular ADSN field, which means $t*\tilde{t}^T = h*\tilde{h}^T$, or equivalently, in Fourier domain, $\hat{t}\overline{\hat{t}}^* = \hat{h}\overline{\hat{h}}^*$.

Notice that if t is any texton associated to a circular ADSN field and if $\alpha \in \mathbb{R}^d$ is a constant line vector with Euclidean norm 1, then $t\alpha$ is a matrix texton associated to the same ADSN.

Among all the textons associated to a circular ADSN field, it is possible to isolate some representatives satisfying an additional constraint.

Definition 2.1.11 ([Desolneux *et al.* 2012])**.** Let us consider a circular ADSN with kernel function $h : \Theta \longrightarrow \mathbb{R}^d$.

- If $d = 1$, then the canonical texton of the ADSN field $\mathcal{N}(m, h * \tilde{h})$ is the function $t_{can} : \Theta \longrightarrow \mathbb{R}$ defined in Fourier domain by

$$\widehat{t_{can}} = |\hat{h}| \ .$$

- If $d = 3$, then the luminance texton of the ADSN field $\mathcal{N}(m, h * \tilde{h}^T)$ is the function $t_{\text{lum}} : \Theta \longrightarrow \mathbb{R}^3$ defined in Fourier domain by

$$\widehat{t_{\text{lum}}} = e^{-i\varphi}\hat{h} \ ,$$

with $\varphi = \mathrm{Arg}(\widehat{h_{\mathrm{lum}}})$ and where the luminance channel of h is defined by

$$h_{\mathrm{lum}} = 0.299h_1 + 0.587h_2 + 0.114h_3 \, .$$

This choice of coefficients for the definition of the luminance channel follows [ITU 2011] but other choices are possible.

As shown in [Desolneux *et al.* 2012], the canonical and luminance textons are interesting because in general they provide a very concentrated summary of the covariance function (and thus of a Gaussian texture). As shown in [Desolneux *et al.* 2012] and [Desolneux *et al.* 2015], the canonical texton is, among all the textons associated to a model, the solution of different optimization problems linked to the concentration at the spatial point 0. We refer to these two articles for the description and properties of the canonical and luminance textons, for several examples of textons associated to Gaussian textures, and also for a discussion about the definition of textons in the multi-channel case.

However, we will see in Section 3.3 (dedicated to the so-called synthesis-oriented texton) that if we seek a texton with a limited compact support, then cropping the luminance texton is not the optimal solution in terms of model error. In order to define such a model error, we need to explain how to measure the distance between two ADSN fields, which is the object of the next subsection.

2.1.6 Optimal Transport Distance

In this subsection, we first recall the definition of the L^2 optimal transport distance [Villani 2003]. This distance has been used in [Xia *et al.* 2014] in order to define barycenters of Gaussian texture models (with application to texture mixing). In particular, [Xia *et al.* 2014] give an expression for the distance between two ADSN models. In Theorem 2.1.4, we generalize this expression to the case of two rank-one colored noises driven by the same white noise process; in this case, we provide a simplified proof based only on the spectral representation of the random field.

Let us denote by μ_0, μ_1 the probability distributions of two random fields on Θ. Let us recall that a coupling of (μ_0, μ_1) is a couple (F, G) of random fields defined on the same probability space and such that $F \sim \mu_0$ and $G \sim \mu_1$.

We shall denote by $\|\cdot\|_\Theta$ the normalized ℓ^2-norm of a function $f : \Theta \longrightarrow \mathbb{C}^d$,

$$\|f\|_\Theta^2 = \frac{1}{|\Theta|} \sum_{\mathbf{x} \in \Theta} \|f(\mathbf{x})\|^2 \, .$$

Definition 2.1.12. The L^2-optimal transport distance (OTD) between μ_0 and μ_1 is defined by

$$d_{\mathrm{OT}}(\mu_0, \mu_1)^2 = \inf \mathbb{E}\Big(\|F - G\|_\Theta^2 \Big) \, . \tag{2.10}$$

where the infimum is taken over couplings (F, G) of (μ_0, μ_1).

We begin by a simple proposition showing that the computation of such distances can be reduced to the case of random fields with zero mean.

Proposition 2.1.7. *If F, G are random fields of distributions μ_0, μ_1 with mean fields $m_0, m_1 : \Theta \longrightarrow \mathbb{R}^d$ and if ν_0, ν_1 refer to the distributions of $F - m_0, G - m_1$, then*

$$d_{OT}(\mu_0, \mu_1)^2 = d_{OT}(\nu_0, \nu_1)^2 + \sum_{\mathbf{x}\in\Theta} \|m_0(\mathbf{x}) - m_1(\mathbf{x})\|^2 .$$

Proof. A simple calculation gives

$$\begin{aligned}
&\mathbb{E}\left(\sum_{\mathbf{x}} \|F(\mathbf{x}) - G(\mathbf{x})\|^2\right) \\
&= \sum_{\mathbf{x}} \mathbb{E}\Big(\|(F(\mathbf{x}) - m_0(\mathbf{x})) - (G(\mathbf{x}) - m_1(\mathbf{x})) + (m_0(\mathbf{x}) - m_1(\mathbf{x}))\|^2\Big) \\
&= \sum_{\mathbf{x}} \mathbb{E}\Big(\|(F(\mathbf{x}) - m_0(\mathbf{x})) - (G(\mathbf{x}) - m_1(\mathbf{x}))\|^2\Big) + \|m_0(\mathbf{x}) - m_1(\mathbf{x})\|^2 \\
&\qquad + 2(m_0(\mathbf{x}) - m_1(\mathbf{x}))^T \mathbb{E}\Big((F(\mathbf{x}) - m_0(\mathbf{x})) - (G(\mathbf{x}) - m_1(\mathbf{x}))\Big) \\
&= \sum_{\mathbf{x}} \mathbb{E}\Big(\|(F(\mathbf{x}) - m_0(\mathbf{x})) - (G(\mathbf{x}) - m_1(\mathbf{x}))\|^2\Big) + \|m_0(\mathbf{x}) - m_1(\mathbf{x})\|^2 ,
\end{aligned}$$

and taking the infimum on (F, G) leads to the desired result. □

The next proposition is another simple result which allows to better understand the optimal transport distance in the case of stationary random fields. The result will not be used in the rest of the paragraph, but it will be helpful to generalize the optimal transport distance to stationary random fields over $\mathbb{Z}^2$.

Proposition 2.1.8. *Let μ_0, μ_1 be two circular stationary random fields on Θ, then*

$$d_{OT}(\mu_0, \mu_1)^2 = \inf \mathbb{E}\Big(\|F(0) - G(0)\|^2\Big) ,$$

where the infimum is taken over all the stationary couplings (F, G) of μ_0, μ_1.

Proof. Let us first show that, since μ_0, μ_1 are stationary, the infimum in (2.10) can be restricted to stationary couplings (F, G) (F and G are stationary but (F, G) may not be). Indeed, let $F \sim \mu_0$ and $G \sim \mu_1$. Let us introduce a uniform translation τ of the domain Θ which is independent of (F, G) and let us consider $(F', G') = (F, G) \circ \tau$. Since μ_0 is stationary, $F' = F \circ \tau \sim \mu_0$, and also $G' \sim \mu_1$. Besides, we have $\|F' - G'\|^2 = \|F - G\|^2$. This shows that we can indeed restrict to stationary couplings.

Furthermore, if (F, G) is a stationary coupling of (μ_0, μ_1) then

$$\mathbb{E}[\|F - G\|_\Theta^2] = \frac{1}{|\Theta|} \sum_{\mathbf{x}\in\Theta} \mathbb{E}\Big(\|F(\mathbf{x}) - G(\mathbf{x})\|^2\Big) = \mathbb{E}\Big(\|F(0) - G(0)\|^2\Big) .$$

□

The authors of [Xia *et al.* 2014] give an explicit formula for the optimal transport distance between circular ADSN fields. Their proof is based upon the main theorem of [Dowson & Landau 1982] which expresses the distance between two centered Gaussian random vectors in terms of their covariance functions. We show in

the next theorem that this result can be obtained directly using the spectral representation of the circular ADSN fields, and thus can be generalized to a wider class of random fields.

Theorem 2.1.4. *Let $h_0, h_1 : \Theta \longrightarrow \mathbb{R}^d$ be two kernel functions and let μ_0, μ_1 be the distributions of $h_0 * W$ and $h_1 * W$ where W is a white noise on Θ with mean 0, variance 1. We assume that W has uniform phase. Then the L^2-optimal transport distance between μ_0 and μ_1 is given by*

$$d_{OT}(\mu_0, \mu_1)^2 = \frac{1}{|\Theta|} \sum_{\boldsymbol{\xi} \in \Theta} \|\hat{h}_0(\boldsymbol{\xi})\|^2 + \|\hat{h}_1(\boldsymbol{\xi})\|^2 - 2|\hat{h}_0(\boldsymbol{\xi})^* \hat{h}_1(\boldsymbol{\xi})| \; . \tag{2.11}$$

Proof. Let us first prove that the right-hand side is a lower bound of the squared distance. For that, let (F, G) be a coupling of (μ_0, μ_1). Thanks to Parseval's formula, one gets

$$\begin{aligned}
\mathbb{E}\left(\sum_{\mathbf{x} \in \Theta} \|F(\mathbf{x}) - G(\mathbf{x})\|^2 \right) &= \frac{1}{MN} \mathbb{E}\left(\sum_{\boldsymbol{\xi} \in \Theta} \|\widehat{F}(\boldsymbol{\xi}) - \widehat{G}(\boldsymbol{\xi})\|^2 \right) \\
&= \frac{1}{MN} \sum_{\boldsymbol{\xi} \in \Theta} \mathbb{E}\Big(\|\widehat{F}(\boldsymbol{\xi}) - \widehat{G}(\boldsymbol{\xi})\|^2 \Big) \\
&= \frac{1}{MN} \sum_{\boldsymbol{\xi} \in \Theta} \mathbb{E}\Big(\|\widehat{F}(\boldsymbol{\xi})\|^2 + \|\widehat{G}(\boldsymbol{\xi})\|^2 - 2\,\mathrm{Re}(\widehat{F}(\boldsymbol{\xi})^* \widehat{G}(\boldsymbol{\xi})) \Big) \\
&= \frac{1}{MN} \sum_{\boldsymbol{\xi} \in \Theta} \mathbb{E}\Big(\|\widehat{F}(\boldsymbol{\xi})\|^2 \Big) + \mathbb{E}\Big(\|\widehat{G}(\boldsymbol{\xi})\|^2 \Big) - 2\,\mathrm{Re}\left(\mathbb{E}(\widehat{F}(\boldsymbol{\xi})^* \widehat{G}(\boldsymbol{\xi})) \right)
\end{aligned}$$

Since $F \sim h_0 * W$,

$$\mathbb{E}\Big(\|\widehat{F}(\boldsymbol{\xi})\|^2 \Big) = \mathbb{E}\Big(\|\widehat{h_0 * W}(\boldsymbol{\xi})\|^2 \Big) = \mathbb{E}\Big(|\widehat{W}(\boldsymbol{\xi})|^2 \|\hat{h}_0(\boldsymbol{\xi})\|^2 \Big) \; ,$$

and since W is a normalized white noise, we get

$$\mathbb{E}\Big(\|\widehat{F}(\boldsymbol{\xi})\|^2 \Big) = MN \|\hat{h}_0(\boldsymbol{\xi})\|^2 \; .$$

Similarly,

$$\mathbb{E}\Big(\|\widehat{G}(\boldsymbol{\xi})\|^2 \Big) = MN \|\hat{h}_1(\boldsymbol{\xi})\|^2 \; .$$

Moreover, we have almost surely $\widehat{F}(\boldsymbol{\xi}) \in \mathbb{C}\hat{h}_0(\boldsymbol{\xi})$. If $\hat{h}_0(\boldsymbol{\xi}) = 0$, if follows that $\mathbb{E}\Big(\widehat{F}(\boldsymbol{\xi})^* \widehat{G}(\boldsymbol{\xi}) \Big) = 0$. Otherwise, we have

$$\widehat{F}(\boldsymbol{\xi}) = \frac{\hat{h}_0(\boldsymbol{\xi}) \hat{h}_0(\boldsymbol{\xi})^*}{\|\hat{h}_0(\boldsymbol{\xi})\|^2} \widehat{F}(\boldsymbol{\xi})$$

because the right-hand side is the orthogonal projection of $\widehat{F}(\boldsymbol{\xi})$ on $\mathbb{C}\hat{h}_0(\boldsymbol{\xi})$. Thus

$$\widehat{F}(\boldsymbol{\xi})^* \widehat{G}(\boldsymbol{\xi}) = \widehat{F}(\boldsymbol{\xi})^* \frac{\hat{h}_0(\boldsymbol{\xi}) \hat{h}_0(\boldsymbol{\xi})^*}{\|\hat{h}_0(\boldsymbol{\xi})\|^2} \widehat{G}(\boldsymbol{\xi})$$

Therefore, using Cauchy-Schwarz inequality, we get

$$\begin{aligned}\Big|\operatorname{Re}\Big(\mathbb{E}\Big(\widehat{F}(\boldsymbol{\xi})^*\widehat{G}(\boldsymbol{\xi})\Big)\Big)\Big| &\leqslant \left|\mathbb{E}\left(\widehat{F}(\boldsymbol{\xi})^*\frac{\hat{h}_0(\boldsymbol{\xi})\hat{h}_0(\boldsymbol{\xi})^*}{\|\hat{h}_0(\boldsymbol{\xi})\|^2}\widehat{G}(\boldsymbol{\xi})\right)\right| \\ &\leqslant \mathbb{E}\left(\left|\frac{\widehat{F}(\boldsymbol{\xi})^*\hat{h}_0(\boldsymbol{\xi})}{\|\hat{h}_0(\boldsymbol{\xi})\|}\right|\left|\frac{\hat{h}_0(\boldsymbol{\xi})^*\widehat{G}(\boldsymbol{\xi})}{\|\hat{h}_0(\boldsymbol{\xi})\|}\right|\right) \\ &\leqslant \left[\mathbb{E}\left(\left|\frac{\widehat{F}(\boldsymbol{\xi})^*\hat{h}_0(\boldsymbol{\xi})}{\|\hat{h}_0(\boldsymbol{\xi})\|}\right|^2\right)\mathbb{E}\left(\left|\frac{\hat{h}_0(\boldsymbol{\xi})^*\widehat{G}(\boldsymbol{\xi})}{\|\hat{h}_0(\boldsymbol{\xi})\|}\right|^2\right)\right]^{1/2}.\end{aligned}$$

Now, since $G \sim h_1 * W$, we have

$$\begin{aligned}\mathbb{E}\left(\left|\frac{\widehat{G}(\boldsymbol{\xi})^*\hat{h}_0(\boldsymbol{\xi})}{\|\hat{h}_0(\boldsymbol{\xi})\|}\right|^2\right) &= \mathbb{E}\left(\left|\frac{\widehat{W}(\boldsymbol{\xi})^*\hat{h}_1(\boldsymbol{\xi})^*\hat{h}_0(\boldsymbol{\xi})}{\|\hat{h}_0(\boldsymbol{\xi})\|}\right|^2\right) \\ &= \mathbb{E}\Big(|\widehat{W}(\boldsymbol{\xi})|^2\Big)\frac{|\hat{h}_1(\boldsymbol{\xi})^*\hat{h}_0(\boldsymbol{\xi})|^2}{\|\hat{h}_0(\boldsymbol{\xi})\|^2} = MN\frac{|\hat{h}_1(\boldsymbol{\xi})^*\hat{h}_0(\boldsymbol{\xi})|^2}{\|\hat{h}_0(\boldsymbol{\xi})\|^2}\end{aligned}$$

and by a similar calculation

$$\mathbb{E}\left(\left|\frac{\widehat{F}(\boldsymbol{\xi})^*\hat{h}_0(\boldsymbol{\xi})}{\|\hat{h}_0(\boldsymbol{\xi})\|}\right|^2\right) = MN\|\hat{h}_0(\boldsymbol{\xi})\|^2 .$$

Therefore,

$$\Big|\operatorname{Re}\Big(\mathbb{E}[\widehat{F}(\boldsymbol{\xi})^*\widehat{G}(\boldsymbol{\xi})]\Big)\Big| \leqslant MN|\hat{h}_1(\boldsymbol{\xi})^*\hat{h}_0(\boldsymbol{\xi})| .$$

It thus follows that

$$\mathbb{E}\Big(\|F-G\|^2\Big) \geqslant \frac{1}{|\Theta|}\sum_{\boldsymbol{\xi}\in\Theta}\|\hat{h}_0(\boldsymbol{\xi})\|^2 + \|\hat{h}_1(\boldsymbol{\xi})\|^2 - 2|\hat{h}_0(\boldsymbol{\xi})^*\hat{h}_1(\boldsymbol{\xi})| ,$$

which entails

$$d_{\mathrm{OT}}(\mu_0,\mu_1)^2 \geqslant \frac{1}{|\Theta|}\sum_{\boldsymbol{\xi}\in\Theta}\|\hat{h}_0(\boldsymbol{\xi})\|^2 + \|\hat{h}_1(\boldsymbol{\xi})\|^2 - 2|\hat{h}_0(\boldsymbol{\xi})^*\hat{h}_1(\boldsymbol{\xi})| .$$

To end the proof, we will exhibit one coupling that achieves the lower bound. For that, we first define a texton $g_1 : \Theta \longrightarrow \mathbb{R}^d$ associated to the model μ_1 by

$$\hat{g}_1 = \frac{\hat{h}_1\hat{h}_1^*\hat{h}_0}{|\hat{h}_1^*\hat{h}_0|}\mathbf{1}_{\hat{h}_1^*\hat{h}_0\neq 0} + \hat{h}_1\mathbf{1}_{\hat{h}_1^*\hat{h}_0=0} .$$

Notice that $\hat{g}_1(\boldsymbol{\xi})$ is a projection of $\hat{h}_0(\boldsymbol{\xi})$ on the circle $\{e^{i\theta}\hat{h}_1(\boldsymbol{\xi}) \ ; \ \theta \in \mathbb{T}\}$ for the ℓ^2 distance, which is uniquely defined by

$$\frac{\hat{h}_1(\boldsymbol{\xi})^*\hat{h}_0(\boldsymbol{\xi})}{|\hat{h}_1(\boldsymbol{\xi})^*\hat{h}_0(\boldsymbol{\xi})|}\hat{h}_1(\boldsymbol{\xi})$$

as soon as $\hat{h}_1(\boldsymbol{\xi})^*\hat{h}_0(\boldsymbol{\xi}) \neq 0$, that is, $\hat{h}_1(\boldsymbol{\xi})$ not orthogonal to $\hat{h}_0(\boldsymbol{\xi})$ (see the Remark 2.1.3 below about this projection).

Let us consider the coupling $(h_0 * W, g_1 * W)$ of (μ_0, μ_1). By construction,

$$\begin{aligned}\mathbb{E}\Big(\sum_{\mathbf{x}\in\Theta} \|F(\mathbf{x}) - G(\mathbf{x})\|^2\Big) &= \frac{1}{MN}\sum_{\boldsymbol{\xi}\in\Theta} \mathbb{E}\Big(\|\widehat{F}(\boldsymbol{\xi}) - \widehat{G}(\boldsymbol{\xi})\|^2\Big) \\ &= \frac{1}{MN}\sum_{\boldsymbol{\xi}\in\Theta} \mathbb{E}\Big(|\widehat{W}(\boldsymbol{\xi})|^2\Big)\|\hat{h}_0(\boldsymbol{\xi}) - \hat{g}_1(\boldsymbol{\xi})\|^2 \\ &= \sum_{\boldsymbol{\xi}\in\Theta} \|\hat{h}_0(\boldsymbol{\xi}) - \hat{g}_1(\boldsymbol{\xi})\|^2 \,.\end{aligned}$$

And finally, in either case $\hat{h}_1^*(\boldsymbol{\xi})\hat{h}_0(\boldsymbol{\xi}) = 0$ or $\hat{h}_1^*(\boldsymbol{\xi})\hat{h}_0(\boldsymbol{\xi}) \neq 0$, one has

$$\|\hat{h}_0(\boldsymbol{\xi}) - \hat{g}_1(\boldsymbol{\xi})\|^2 = \|\hat{h}_0(\boldsymbol{\xi})\|^2 + \|\hat{h}_1(\boldsymbol{\xi})\|^2 - 2|\hat{h}_1(\boldsymbol{\xi})^*\hat{h}_0(\boldsymbol{\xi})| \,,$$

which completes the proof. □

We would like to emphasize that the last proof is truly a problem of geometry. Its main ingredient is indeed the Cauchy-Schwarz inequality. Besides, the right-hand side of Equation (2.11) admits a geometric interpretation given by the next Remark.

Remark 2.1.3. For $a \in \mathbb{C}^d$, let us denote $\mathcal{C}_a = \{e^{i\theta}a \,;\, \theta \in \mathbb{T}\}$. If $a, b \in \mathbb{C}^d$, then

$$\|a\|^2 + \|b\|^2 - 2|a^*b| = d(\mathcal{C}_a, \mathcal{C}_b)^2 := \inf_{\substack{z\in\mathcal{C}_a \\ w\in\mathcal{C}_b}} \|z - w\|^2 \,.$$

Indeed, if $a^*b = 0$ then the circles $\mathcal{C}_a$ and $\mathcal{C}_b$ are contained in two orthogonal subspaces of $\mathbb{C}^d$, so that for all $z \in \mathcal{C}_a$ and $w \in \mathcal{C}_b$,

$$\|z - w\|^2 = \|z\|^2 + \|w\|^2 = \|a\|^2 + \|b\|^2 \,.$$

Let us now assume that $a^*b \neq 0$. Let $z \in \mathcal{C}_a$ and $w \in \mathcal{C}_b$. Let us denote by p the orthogonal projection of w on $\mathbb{C}a = \mathbb{C}z$, which is given by

$$p = \frac{z^*w}{\|z\|^2}\, z \,. \tag{2.12}$$

We thus have

$$\|z - w\|^2 = \|z - p\|^2 + \|p - w\|^2 \,.$$

Next, as one can see in Fig. 2.3, in $\mathbb{C}z$ (which can be seen as a plane), the projection q of p on $\mathcal{C}_a = \mathcal{C}_z$ is given by

$$q = \frac{\|z\|}{\|p\|}\, p \tag{2.13}$$

(because it is the only point of $\mathbb{R}_+p$ which has norm $\|z\|$). Since $z \in \mathcal{C}_a$ and since q is the projection of p on $\mathcal{C}_a$, we have

$$\|z - p\| \geqslant \|q - p\| \,.$$

Besides, grouping (2.12) and (2.13), we get

$$q = \frac{z^*w}{|z^*w|}\, z \,. \tag{2.14}$$

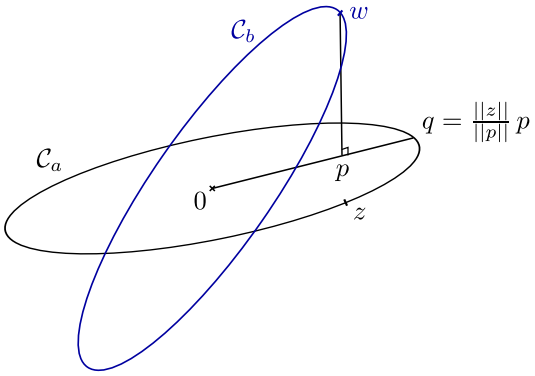


Figure 2.3: **Projection on a circle.** After the projection on $\mathbb{C}z$, it only amounts to correct the norm, leading to Equation (2.14).

Hence, recalling that $w - p$ is orthogonal to $q - p \in \mathbb{C}a$, we get

$$\|z - w\|^2 \geqslant \|q - p\|^2 + \|p - w\|^2 = \|q - w\|^2 \ .$$

Furthermore,

$$\|q-w\|^2 = \|q\|^2 + \|w\|^2 - 2\,\mathrm{Re}(q^* w) = \|q\|^2 + \|w\|^2 - 2|z^* w| = \|a\|^2 + \|b\|^2 - 2|a^* b| \ .$$

where the last equality holds because $z \in \mathcal{C}_a$, $w \in \mathcal{C}_b$, and $q \in \mathcal{C}_a$. Therefore,

$$d(\mathcal{C}_a, \mathcal{C}_b)^2 = \|a\|^2 + \|b\|^2 - 2|a^* b| \ .$$

More precisely, we have shown that for each $w \in \mathcal{C}_b$, there exists a point $q \in \mathcal{C}_a$ such that

$$\|q - w\|^2 = d(\mathcal{C}_a, w)^2 = d(\mathcal{C}_a, \mathcal{C}_b)^2 = \|a\|^2 + \|b\|^2 - 2|a^* b|$$

and that this point q can be obtained as $\frac{z^* w}{|z^* w|}\, z$ where z is any point in $\mathcal{C}_a$.

Let us mention that the expression (2.14) of the projection of w on the circle $\mathcal{C}_z$ was proved in [Tartavel *et al.* 2014, Appendix A.1] with a method based on calculus.

Corollary 2.1.5 ([Xia *et al.* 2014])**.** *For $i = 0, 1$, let $\mu_i = \mathcal{N}(m_i, h_i * \tilde{h}_i^T)$ where $m_i \in \mathbb{R}^d$ and where $h_i : \Theta \longrightarrow \mathbb{R}^d$ is a kernel function. Then*

$$d_{OT}(\mu_0, \mu_1)^2 = \|m_0 - m_1\|^2 + \frac{1}{|\Theta|} \sum_{\boldsymbol{\xi} \in \Theta} \|\hat{h}_0(\boldsymbol{\xi})\|^2 + \|\hat{h}_1(\boldsymbol{\xi})\|^2 - 2|\hat{h}_0(\boldsymbol{\xi})^* \hat{h}_1(\boldsymbol{\xi})| \ .$$

Proof. The proof follows from the last theorem, and from Proposition 2.1.7, noticing that in this case, the mean fields of μ_0, μ_1 are constant. □

Corollary 2.1.6. *Let us assume that $d = 1$ and let us consider two circular ADSN fields μ_0, μ_1 with kernel functions h_0, h_1. If t_0, t_1 denote the canonical textons of the Gaussian models μ_0, μ_1, then*

$$d_{OT}(\mu_0, \mu_1)^2 = \sum_{\mathbf{x} \in \Theta} (t_0(\mathbf{x}) - t_1(\mathbf{x}))^2 \ ,$$

which means that the distance is exactly the non-normalized ℓ^2 distance between the corresponding canonical textons.

Proof. Since $d = 1$, the result of Theorem 2.1.4 simply writes

$$d_{\mathrm{OT}}(\mu_0, \mu_1)^2 = \frac{1}{|\Theta|} \sum_{\boldsymbol{\xi} \in \Theta} |\hat{h}_0(\boldsymbol{\xi})|^2 + |\hat{h}_1(\boldsymbol{\xi})|^2 - 2|\hat{h}_0(\boldsymbol{\xi})||\hat{h}_1(\boldsymbol{\xi})|$$
$$= \frac{1}{|\Theta|} \sum_{\boldsymbol{\xi} \in \Theta} \Big(|\hat{h}_0(\boldsymbol{\xi})| - |\hat{h}_1(\boldsymbol{\xi})| \Big)^2 .$$

Besides, since t_i is the canonical texton associated to h_i, we have $\hat{t}_i = |\hat{h}_i|$ which entails

$$d_{\mathrm{OT}}(\mu_0, \mu_1)^2 = \frac{1}{|\Theta|} \sum_{\boldsymbol{\xi} \in \Theta} \Big(\hat{t}_0(\boldsymbol{\xi}) - \hat{t}_1(\boldsymbol{\xi}) \Big)^2 ,$$

and the desired result follows from Parseval's formula. □

Let us emphasize that Theorem 2.1.4 gives the optimal transport distance between rank-one colored noises generated with the same white noise process with uniform phase. It includes the circular ADSN case (already studied in [Xia *et al.* 2014]) and the RPN case (mentioned in [Desolneux *et al.* 2015]). Let us mention that an analog result will be presented in Subsection 2.2.7 in the case of stationary random fields over $\mathbb{Z}^2$.

Further generalizations of this theorem could be useful. For example, it would be interesting to express the optimal transport distance between rank-one colored noises generated with different white noise processes. This would give a way to measure the global convergence speed of the DSN to its Gaussian limit. But in contrast with the extension to $\mathbb{Z}^2$ presented in Subsection 2.2.7, this new generalization would certainly be much more difficult (see [Huesmann & Sturm 2013] for the case of a Poisson process and a uniform process).

To end this paragraph, let us mention the following generalization of Theorem 2.1.4 with one ADSN and one general Gaussian random field.

Theorem 2.1.7. *Let $F : \Theta \longrightarrow \mathbb{R}^d$ be a circular stationary Gaussian field of mean 0 and covariance C, and let us denote by μ the distribution of F. Let $h : \Theta \longrightarrow \mathbb{R}^d$ be a kernel function and let ν be the distribution of the circular ADSN with spot h. The L^2-optimal transport distance between μ and ν is given by*

$$d_{OT}(\mu, \nu)^2 = \frac{1}{|\Theta|} \sum_{\boldsymbol{\xi} \in \Theta} \mathrm{Tr}(\hat{C}(\boldsymbol{\xi})) + \|\hat{h}(\boldsymbol{\xi})\|^2 - 2\sqrt{\hat{h}(\boldsymbol{\xi})^* \widehat{C}(\boldsymbol{\xi}) \hat{h}(\boldsymbol{\xi})} \, . \tag{2.15}$$

Besides, if C is fixed, the optimal transport distance between μ and ν is minimized as soon as

$$\forall \boldsymbol{\xi} \in \Theta, \quad \hat{h}(\boldsymbol{\xi}) = e^{i\theta(\boldsymbol{\xi})} a(\boldsymbol{\xi}) p(\boldsymbol{\xi})$$

where $\theta(\boldsymbol{\xi}) \in \mathbb{R}$ and where $p(\boldsymbol{\xi})$ is a normalized eigenvector associated to the largest eigenvalue $a(\boldsymbol{\xi})$ of $\widehat{C}(\boldsymbol{\xi})$.

Proof. The formula (2.15) is a special case of [Dowson & Landau 1982] because

$$\mathrm{Tr}\left(\left(\frac{\hat{h}(\boldsymbol{\xi}) \hat{h}(\boldsymbol{\xi})^*}{\|\hat{h}(\boldsymbol{\xi})\|} \hat{C}(\boldsymbol{\xi}) \frac{\hat{h}(\boldsymbol{\xi}) \hat{h}(\boldsymbol{\xi})^*}{\|\hat{h}(\boldsymbol{\xi})\|} \right)^{1/2} \right) = \sqrt{\hat{h}(\boldsymbol{\xi})^* \widehat{C}(\boldsymbol{\xi}) \hat{h}(\boldsymbol{\xi})} \, .$$

Let us remark that we have found no simpler path than using the very general result by Dowson and Landau. Indeed, the inequality $\geqslant$ can still be shown by adapting the proof of Theorem 2.1.4. However, exhibiting one optimal coupling seems more difficult in this case.

For the second part of the theorem, the problem boils down to the minimization of the function

$$\hat{h}(\boldsymbol{\xi}) \in \mathbb{C}^3 \longmapsto \frac{1}{2}\|\hat{h}(\boldsymbol{\xi})\|^2 - \sqrt{\hat{h}(\boldsymbol{\xi})^* \widehat{C}(\boldsymbol{\xi}) \hat{h}(\boldsymbol{\xi})} \ .$$

Using an eigenvector basis of $\widehat{C}(\boldsymbol{\xi})$ and standard calculus, one can show that the minimum of this function is $-a(\boldsymbol{\xi})$ (where $a(\boldsymbol{\xi})$ is the largest eigenvalue of $\widehat{C}(\boldsymbol{\xi})$) and that it is reached at any eigenvector of $\widehat{C}(\boldsymbol{\xi})$ associated to $a(\boldsymbol{\xi})$. □

This last result allows to interpret in terms of optimal transport distance the canonical decomposition of a Gaussian model that was presented in Subsection 2.1.4. Indeed, the decomposition amounts to progressively extract from the Gaussian model the ADSN components that best approximate the residual Gaussian field in terms of optimal transport distance.

Of course, in general the realizations of a Gaussian field and the realizations of its principal ADSN component will not look alike. Indeed, since we extract the principal component on each frequency, there will be in general a severe loss of variance (it can be seen for example on the case of a Gaussian field obtained by concatenation of three independent ADSN fields). However, a more interesting question is to ask whether the realizations of a general color Gaussian field are always visually similar to the realizations of well-chosen color ADSN field. We will see in Section 3.1 how such an ADSN model can be estimated. So far we have found no counter-example to that assertion.

2.2 Random Fields on $\mathbb{Z}^2$

In many applications (in particular texture modelling), the circular framework is not natural. Here, we extend the tools presented in the last section to the case of random fields defined on the whole discrete plane $\mathbb{Z}^2$. In this setting the Fourier transform of a random field is not necessarily sampled on a discrete grid. It is one reason that makes this setting more convenient for the comparison of the different texture models (moving average, autoregressive, discrete spot noise, Gaussian, Markovian).

2.2.1 General Definitions

Random Fields and Covariance Functions

Definition 2.2.1. A random field on $\mathbb{Z}^2$ is a random process F on $\mathbb{Z}^2$ with values in $\mathbb{R}^d$. Again, omitting the random sample ω, we will often denote F as a random function

$$F : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d \ .$$

The random field F is said to be of order p if for all $\mathbf{x} \in \mathbb{Z}^2$, $\mathbb{E}(\|F(\mathbf{x})\|^p) < \infty$.

The basic definitions that we introduced in the circular case can be easily adapted to the infinite-lattice case, by considering now the addition of $\mathbb{Z}^2$ instead of the addition on $\Theta = \mathbb{Z}/M\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$. For that reason, we will not repeat all the definitions.

Definition 2.2.2. A random field $F : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$ is said to be stationary if for every $n \geqslant 1$, $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{Z}^2$, $\mathbf{v} \in \mathbb{Z}^2$, $(F(\mathbf{x}_1), \ldots, F(\mathbf{x}_n))$ has the same distribution as $(F(\mathbf{x}_1 + \mathbf{v}), \ldots, F(\mathbf{x}_n + \mathbf{v}))$. A second-order stationary random field on $\mathbb{Z}^2$ has a constant expected value

$$m = \mathbb{E}(F(\mathbf{x})) \in \mathbb{R}^d .$$

Besides, there exists a function $C_F : \mathbb{Z}^2 \longrightarrow \mathbb{R}^{d\times d}$ called covariance function of F, such that

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{Z}^2, \quad C_F(\mathbf{x} - \mathbf{y}) = \mathrm{Cov}(F(\mathbf{x}), F(\mathbf{y})) = \mathbb{E}((F(\mathbf{x}) - m)(F(\mathbf{y}) - m)^T) .$$

We will write $F \sim \mathcal{N}(m, C)$ if F is a stationary Gaussian random field with mean m and covariance function C.

Spectral Representation

The spectral representation is now more difficult to obtain because the Fourier series associated to a function $f : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$ may not converge. Nevertheless, the Herglotz theorem ensures that the covariance function of a stationary process on $\mathbb{Z}^2$ is the inverse Fourier transform of a non-negative finite measure on $\mathbb{T}^2$. Based on this theorem, one can build a spectral representation of the process: F can be written as the inverse Fourier transform of a random measure, understood as a stochastic integral. But in this thesis, we will not need the Fourier transform of F, so that we only include here the Herglotz theorem that gives the Fourier transform of the covariance function. A wider discussion about spectral representations can be found in the following references. The existence of the spectral representation is proved in a very general framework in [Rozanov 1967] (see also the seminal paper [Cramer 1940]). The case of one discrete variable is more simply discussed in [Brémaud 1993] and [Cabral 2010]. Let us also mention that the spectral representation of continuous generalized random fields is thoroughly studied in [Gelfand & Vilenkin 1967].

Theorem 2.2.1 (Herglotz)**.** *Let $F : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$ be a second-order stationary random field. Then there exists a unique finite measure μ_F on $\mathbb{T}^2$ with values in the set $\mathcal{H}_d^{\geqslant 0}$ of non-negative Hermitian matrices, such that*

$$\forall \mathbf{x} \in \mathbb{Z}^2, \quad C_F(\mathbf{x}) = \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} e^{i\boldsymbol{\xi}.\mathbf{x}} d\mu_F(\boldsymbol{\xi}) .$$

This measure μ_F is called spectral measure of F.

Since μ_F is matrix-valued, its components are real-valued finite Borelian measures on $\mathbb{T}^2$. We say that F admits a spectral density if μ_F admits a density φ_F with respect to the Lebesgue measure on $\mathbb{T}^2$. This spectral density φ_F is a function on $\mathbb{T}^2$ such that for almost all $\mathbf{x} \in \mathbb{T}^2$, $\varphi(\mathbf{x}) \in \mathcal{H}_d^{\geqslant 0}$.

If F, G are two stationary random fields on $\mathbb{Z}^2$, one can consider the spectral measure $\mu_{(F;G)}$ associated to $(F;G)$. If F and G have zero mean (for the sake of simplicity), it satisfies

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{Z}^2, \quad \mathbb{E}\left[\begin{pmatrix} F(\mathbf{x}) \\ G(\mathbf{x}) \end{pmatrix} (F(\mathbf{y})^T G(\mathbf{y})^T)\right] = \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} e^{i\boldsymbol{\xi}.(\mathbf{x}-\mathbf{y})} d\mu_{(F;G)}(\boldsymbol{\xi}) \ .$$

One can see that

$$\mu_{(F;G)} = \begin{pmatrix} \mu_F & \nu_{F,G} \\ \nu_{G,F} & \mu_G \end{pmatrix}$$

where μ_F and μ_G are the spectral measures of F and G. The measure $\nu_{F,G}$ is called the cross-spectral measure of F and G.

Let us now assume that F and G are real-valued, and that they admit spectral densities φ_F and φ_G. Then μ_F, μ_G are finite non-negative measures. Besides, if A is a Borelian set of $\mathbb{T}^2$, since $\mu_{(F;G)}(A)$ is a Hermitian matrix, we have

$$|\nu_{F,G}(A)|^2 \leqslant \mu_F(A)\mu_G(A) \ . \tag{2.16}$$

Therefore, if $\lambda(A) = 0$, then $\nu_{F,G}(A) = 0$ which shows that $\nu_{F,G}$ is absolutely continuous with respect to the Lebesgue measure. Thus $\nu_{F,G}$ admits a density $\psi_{F,G} \in L^1(\mathbb{T})$ with respect to Lebesgue measure on $\mathbb{T}^2$. This function is called the cross-spectral density of F, G. From Equation (2.16), it follows that we have for almost all $\boldsymbol{\xi} \in \mathbb{T}^2$,

$$|\psi_{F,G}(\boldsymbol{\xi})|^2 \leqslant \varphi_F(\boldsymbol{\xi})\varphi_G(\boldsymbol{\xi}) \ . \tag{2.17}$$

More generally, if F and G are $\mathbb{R}^d$-valued and admit spectral densities φ_F, φ_G, then $\nu_{F,G}$ still admit a matrix-valued density called the cross spectral density of F, G. Indeed, it is enough to show that each component of $\nu_{F,G}$ admits a spectral density, which follows from the real-valued case: indeed, for $k, l = 1, \ldots, d$, we still have

$$|\nu_{F_k,G_l}(A)|^2 \leqslant \mu_{F_k}(A)\mu_{G_l}(A) \ .$$

Remark 2.2.1. A circular stationary random field F on $\Theta = \mathbb{Z}/M\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$ can be identified to a (M, N)-periodic stationary random field on $\mathbb{Z}^2$, and its covariance function $C : \Theta \longrightarrow \mathbb{R}^{d\times d}$ to a periodic covariance function on $\mathbb{Z}^2$. Besides, we have

$$\forall \mathbf{x} \in \mathbb{Z}^2, \quad C(\mathbf{x}) = \frac{1}{MN} \sum_{\boldsymbol{\xi}\in\Theta} \widehat{C}(\boldsymbol{\xi}) \exp\left(2i\pi\Big(\frac{\xi_1 x_1}{M} + \frac{\xi_2 x_2}{N}\Big)\right) \ ,$$

where the DFT of C satisfies $\widehat{C}(\boldsymbol{\xi}) \in \mathcal{H}_d^{\geqslant 0}$. Therefore, the spectral measure of the stationary process $F : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$ is

$$\mu_F = \frac{(2\pi)^2}{MN} \sum_{\boldsymbol{\xi}\in\Theta} \widehat{C}(\boldsymbol{\xi}) \delta_{2\pi\left(\frac{\xi_1}{M}, \frac{\xi_2}{N}\right)} \ .$$

As expected, the periodicity of F entails that its spectral measure is sampled on the discrete subgroup

$$\left\{ \ 2\pi\Big(\frac{\xi_1}{M}, \frac{\xi_2}{N}\Big) \mod 2\pi\mathbb{Z}^2 \ ; \ \boldsymbol{\xi} \in \mathbb{Z}^2 \right\}$$

of $\mathbb{T}^2$. Such a measure is not absolutely continuous with respect to Lebesgue measure.

Definition 2.2.3. A stationary random field $F : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$ with covariance function C is said to be a white noise if it has a constant spectral density (or equivalently if $C = C(0)\delta_0$). We say that F is a pure white noise if $(F(\mathbf{x}))_{\mathbf{x}\in\mathbb{Z}^2}$ are independent.

Since the convolution of two functions f, g on $\mathbb{Z}^2$ is not always defined, the definition of colored noise processes is not as simple as in the case of circular random fields on a finite domain. We end this paragraph by giving two classical results [Doob 1990] which ensure that the convolution $h*W$ defines a second-order stationary process, modulo certain assumptions on the kernel function $h : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$. In the following, we denote by $L^2(\mathbb{P})$ the Banach space of $\mathbb{R}^d$-valued random variables V (defined on the probability space associated to $\mathbb{P}$) such that $\mathbb{E}(\|V\|^2) < \infty$. Thus, by definition, V_n converges to V in $L^2(\mathbb{P})$ if

$$\mathbb{E}(\|V_n - V\|^2) \xrightarrow[n\to\infty]{} 0 \ .$$

Proposition 2.2.1. *If W is a scalar white noise on $\mathbb{Z}^2$ with mean m and variance σ^2 and if $h \in \ell^1(\mathbb{Z}^2, \mathbb{R}^d)$, then, for all $\mathbf{x} \in \mathbb{Z}^2$, the series*

$$\sum_{\mathbf{y}\in\mathbb{Z}^2} h(\mathbf{y})W(\mathbf{x}-\mathbf{y}) \tag{2.18}$$

converges absolutely almost surely and in $L^2(\mathbb{P})$ and defines a second-order stationary process $h*W$ whose mean is $m\sum h$ and whose covariance function is $\sigma^2\, h*\tilde{h}^T$.

Proof. Since W is stationary, $a = \mathbb{E}(|W(\mathbf{x})|)$ does not depend on $\mathbf{x}$, and thus

$$\mathbb{E}\left(\sum_{\mathbf{y}\in\mathbb{Z}^2} \|h(\mathbf{y})W(\mathbf{x}-\mathbf{y})\|\right) = \sum_{\mathbf{y}\in\mathbb{Z}^2} \|h(\mathbf{y})\|\,\mathbb{E}(|W(\mathbf{x}-\mathbf{y})|) = a\sum|h| < \infty$$

In particular, we almost surely have

$$\sum_{\mathbf{y}\in\mathbb{Z}^2} \|h(\mathbf{y})W(\mathbf{x}-\mathbf{y})\| < \infty$$

so that the series (2.18) is almost surely absolutely convergent.

Also, since $b = \|W(\mathbf{x})\|_{L^2(\mathbb{P})}$ is constant, we have similarly

$$\sum_{\mathbf{y}\in\mathbb{Z}^2} \|h(\mathbf{y})W(\mathbf{x}-\mathbf{y})\|_{L^2(\mathbb{P})} = b\sum|h| < \infty \ ,$$

and the series (2.18) is absolutely convergent in $L^2(\mathbb{P})$. Thus we can define a second-order random field $h*W$ by setting

$$h*W(\mathbf{x}) = \sum_{\mathbf{y}\in\mathbb{Z}^2} h(\mathbf{y})W(\mathbf{x}-\mathbf{y}) \ .$$

The stationarity of $h * W$ derives from the one of W. Besides, thanks to Fubini theorem, we get

$$\mathbb{E}(h * W(\mathbf{x})) = m \sum h \, .$$

Also, we can thus write

$$h * W(\mathbf{x}) - \mathbb{E}(h * W(\mathbf{x})) = \sum_{\mathbf{z} \in \mathbb{Z}^2} h(\mathbf{z})(W(\mathbf{x} - \mathbf{z}) - m) \, ,$$

and therefore,

$$\begin{aligned}
&\mathbb{E}\Big(\big(h * W(\mathbf{x}) - m \sum h\big) \big(h * W(\mathbf{y}) - m \sum h\big)^T \Big) \\
&= \sum_{\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{Z}^2} h(\mathbf{z}_1) h(\mathbf{z}_2)^T \mathrm{Cov}(W(\mathbf{x} - \mathbf{z}_1), W(\mathbf{y} - \mathbf{z}_2)) \\
&= \sum_{\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{Z}^2} h(\mathbf{z}_1) h(\mathbf{z}_2)^T \sigma^2 \delta_{\mathbf{x} - \mathbf{z}_1 = \mathbf{y} - \mathbf{z}_2} \\
&= \sigma^2 \sum_{\mathbf{z} \in \mathbb{Z}^2} h(\mathbf{z}) h(\mathbf{y} - \mathbf{x} + \mathbf{z})^T \\
&= \sigma^2 h * \tilde{h}^T(\mathbf{x} - \mathbf{y}) \, .
\end{aligned}$$

□

The next proposition ensures the existence of $h * W(\mathbf{x})$ in $L^2(\mathbb{P})$ under the weaker assumption that $h \in \ell^2$ but requiring the noise W to have zero mean.

Proposition 2.2.2. *Let W be a scalar white noise with mean 0 and variance σ^2, and let $h \in \ell^2(\mathbb{Z}^2, \mathbb{R}^d)$. Then the series*

$$\sum_{\mathbf{y} \in \mathbb{Z}^2} h(\mathbf{y}) W(\mathbf{x} - \mathbf{y}) \tag{2.19}$$

converges in $L^2(\mathbb{P})$ (in the sense of summable families) and thus defines a second-order stationary process $h * W$ with mean zero and covariance function $\sigma^2 \, h * \tilde{h}^T$.

Proof. Let $\mathbf{x} \in \mathbb{Z}^2$. Since $L^2(\mathbb{P})$ is a Banach space, to prove the convergence of the series (2.19), we only have to show that it satisfies the Cauchy criterion for summable families. But if A is a finite subset of $\mathbb{Z}^2$, using the orthogonality of the random variables $W(\mathbf{x})$, we have

$$\mathbb{E}\left\| \sum_{\mathbf{y} \in A} h(\mathbf{y}) W(\mathbf{x} - \mathbf{y}) \right\|^2 = \sum_{\mathbf{y} \in A} \|h(\mathbf{y})\|^2 \mathbb{E}(W(\mathbf{x} - \mathbf{y})^2) = \sigma^2 \sum_{\mathbf{y} \in A} \|h(\mathbf{y})\|^2 \, .$$

Thus it suffices to check the Cauchy criterion for $\sum_{\mathbf{y}} \|h(\mathbf{y})\|^2$ which follows from the fact that $h \in \ell^2$. The covariance function can be obtained as in the proof of Proposition 2.2.1. □

Remark 2.2.2. In Proposition 2.2.2, we required W to be of mean zero. In the case where W has a mean value $m \neq 0$, then the convergence of the series (2.19) in $L^2(\mathbb{P})$

(in the sense of summable families) implies that $h \in \ell^1$. Indeed, for a fixed $\mathbf{x} \in \mathbb{Z}^2$, Proposition 2.2.2 shows that

$$h(\mathbf{y})(W(\mathbf{x}-\mathbf{y}) - m)$$

is a summable family. Therefore, the summability in $L^2(\mathbb{P})$ of $\mathbf{y} \mapsto h(\mathbf{y})W(\mathbf{x}-\mathbf{y})$ is equivalent to the one of $\mathbf{y} \mapsto mh(\mathbf{y})$, which is also equivalent to $h \in \ell^1$ since $\mathbf{y} \mapsto mh(\mathbf{y})$ is deterministic. To sum up, in the case where W has a non-zero mean value, the convergence in $L^2(\mathbb{P})$ cannot happen if $\ell^2 \backslash \ell^1$ and in the case where $h \in \ell^1$ there is a stronger convergence given by Proposition 2.2.1.

Remark 2.2.3. Studying the almost sure convergence of (2.19) in the case where $h \in \ell^2 \backslash \ell^1$ is a more difficult problem. Several results exist for the almost sure convergence of a series $\sum X_n$ of independent random variables X_n in $L^2(\mathbb{P})$, and Kolmogorov's three-series theorem gives a famous necessary and sufficient condition for convergence (see [Billingsley 2012, Chap. 22]). We would like to mention in particular the following result based on a maximal inequality: if (X_n) is an independent sequence of random variables with zero mean, and such that $\sum \mathrm{Var}(X_n) < \infty$, then $\sum X_n$ converges almost surely. But it is not clear that this result extends to the case of random series with indices $\mathbf{x} \in \mathbb{Z}^2$ because the convergence and the sum of the series may now depend on the method of summation. See [Ronsin *et al.* 2013] for the example of random Fourier series.

In view of the results of Proposition 2.2.1 and Proposition 2.2.2, we adopt the following definition.

Definition 2.2.4. A stationary random field $F : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$ is called colored noise on $\mathbb{Z}^2$ if it can be written $F = m + h * \mathcal{W}$ where $m \in \mathbb{R}^d$, where $h : \mathbb{Z}^2 \longrightarrow \mathbb{R}^{d\times d}$, where $\mathcal{W}$ is a $\mathbb{R}^d$-valued white noise on $\mathbb{Z}^2$, and where

$$h * \mathcal{W}(\mathbf{x}) = \sum_{\mathbf{y}\in\mathbb{Z}^2} h(\mathbf{y})\mathcal{W}(\mathbf{x}-\mathbf{y})$$

is defined as a convergent series in $L^2(\mathbb{P})$ (in the sense of summable families). A stationary random field $F : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$ is called rank-one colored noise on $\mathbb{Z}^2$ if it can be written $F = m + h * W$ where $m \in \mathbb{R}^d$, where $h : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$, where W is a scalar white noise on $\mathbb{Z}^2$, and where

$$h * W(\mathbf{x}) = \sum_{\mathbf{y}\in\mathbb{Z}^2} h(\mathbf{y})W(\mathbf{x}-\mathbf{y})$$

is defined as a convergent series in $L^2(\mathbb{P})$ (in the sense of summable families).

In order to obtain the spectral measure of the above-mentioned colored noises, we need to recall a convolution lemma. In this lemma which focuses on a convergence issue, we do not need to precise a specific norm $\|\cdot\|$ on $\mathbb{R}^{p\times q}$ (because they are all equivalent).

Lemma 2.2.2. *Let $h : \mathbb{Z}^2 \longrightarrow \mathbb{R}^{p\times q}$ and $k : \mathbb{Z}^2 \longrightarrow \mathbb{R}^{q\times r}$ be two kernel functions such that*

$$\|h\|_{\ell^2}^2 = \sum_{\mathbf{x}\in\mathbb{Z}^2} \|h(\mathbf{x})\|^2 < \infty \quad \text{and} \quad \|k\|_{\ell^2}^2 = \sum_{\mathbf{x}\in\mathbb{Z}^2} \|k(\mathbf{x})\|^2 < \infty \,.$$

Then the series

$$h * k(\mathbf{x}) = \sum_{\mathbf{y}\in\mathbb{Z}^2} h(\mathbf{y})k(\mathbf{x}-\mathbf{y}) \tag{2.20}$$

is absolutely convergent for all $\mathbf{x} \in \mathbb{Z}^2$ and defines a bounded function $h * k : \mathbb{Z}^2 \longrightarrow \mathbb{R}^{p\times r}$. Besides, $h * k$ is the inverse Fourier transform of $\hat{h}\hat{k}$ in the sense that

$$\forall \mathbf{x} \in \mathbb{Z}^2, \quad h * k(\mathbf{x}) = \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} \hat{h}(\boldsymbol{\xi})\hat{k}(\boldsymbol{\xi}) e^{i\boldsymbol{\xi}.\mathbf{x}} d\boldsymbol{\xi} \,. \tag{2.21}$$

Proof. The first assertion is a direct application of the Cauchy-Schwarz inequality:

$$\sum_{\mathbf{y}\in\mathbb{Z}^2} \|h(\mathbf{y})\|\|k(\mathbf{x}-\mathbf{y})\| \leqslant \|h\|_{\ell^2}\|k\|_{\ell^2} \,.$$

It remains to prove (2.21). Let us fix $\mathbf{x} \in \mathbb{Z}^2$. The beginning of the proof shows that $(h,k) \longmapsto h*k(\mathbf{x})$ is a continuous bilinear application from $\ell^2\times\ell^2$ to $\mathbb{R}^{p\times r}$. Thanks to the Cauchy-Schwarz inequality in $L^2\Big(\mathbb{T}^2, \frac{d\boldsymbol{\xi}}{(2\pi)^2}\Big)$, and thanks to Plancherel's formula, we also have

$$\begin{aligned}\left\| \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} \hat{h}(\boldsymbol{\xi})\hat{k}(\boldsymbol{\xi}) e^{i\boldsymbol{\xi}.\mathbf{x}} d\boldsymbol{\xi} \right\|^2 &\leqslant \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} \|\hat{h}(\boldsymbol{\xi})\|^2 d\boldsymbol{\xi} \times \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} \|\hat{k}(\boldsymbol{\xi})\|^2 d\boldsymbol{\xi} \\ &= \|h\|_{\ell^2}^2 \|k\|_{\ell^2}^2 \,,\end{aligned}$$

so that

$$(h,k) \longmapsto \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} \hat{h}(\boldsymbol{\xi})\hat{k}(\boldsymbol{\xi}) e^{i\boldsymbol{\xi}.\mathbf{x}} d\boldsymbol{\xi}$$

is also a continuous bilinear application. Since the compactly-supported sequences are dense in $\ell^2(\mathbb{Z}^2)$, we only have to prove that (2.21) holds if h and k have compact support. In this case, thanks to the orthogonality of the functions $\boldsymbol{\xi} \longmapsto e^{i\boldsymbol{\xi}.\mathbf{x}}$ one can write

$$\begin{aligned}\frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} \hat{h}(\boldsymbol{\xi})\hat{k}(\boldsymbol{\xi}) e^{i\boldsymbol{\xi}.\mathbf{x}} d\boldsymbol{\xi} &= \sum_{\mathbf{y},\mathbf{z}\in\mathbb{Z}^2} h(\mathbf{y})k(\mathbf{z}) \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} e^{i\boldsymbol{\xi}.(\mathbf{x}-\mathbf{y}-\mathbf{z})} d\boldsymbol{\xi} \\ &= \sum_{\mathbf{y},\mathbf{z}\in\mathbb{Z}^2} h(\mathbf{y})k(\mathbf{z})\delta_{\mathbf{x}=\mathbf{y}+\mathbf{z}} \\ &= \sum_{\mathbf{y}\in\mathbb{Z}^2} h(\mathbf{y})k(\mathbf{x}-\mathbf{y}) = h*k(\mathbf{x}) \,,\end{aligned}$$

where the sums on $\mathbf{y}$ and $\mathbf{z}$ contains only a finite number of terms. □

Therefore, under hypotheses of Proposition 2.2.1 or Proposition 2.2.2, $h * W$ is a rank-one colored noise with covariance function $\mathrm{Var}(W(0))\, h * \tilde{h}^T$. Since we have $h \in \ell^2$ in both cases, Lemma 2.2.2 gives that the spectral measure of $h * W$ is the integrable function $\hat{h}\hat{h}^*$. In the case of Proposition 2.2.1, notice that $h \in \ell^1$ entails that $\hat{h}$ is continuous on $\mathbb{T}^2$ so that $\hat{h} \in \ell^p$ for every $p \in [1,\infty]$, and that $\hat{h}\hat{h}^*$ is also a continuous function.

2.2.2 Discrete Spot Noise on $\mathbb{Z}^2$

In this subsection, we recall the definition and properties of discrete spot noise processes [Papoulis 1971], [Rice 1977], [Van Wijk 1991].

Let $h \in \ell^1(\mathbb{Z}^2, \mathbb{R}^d)$ denote a kernel function and let us denote $\sum h = \sum_{\mathbf{x}\in\mathbb{Z}^2} h(\mathbf{x})$. Let $\lambda > 0$ and let Π_λ denote a Poisson point process on $\mathbb{Z}^2$ with intensity λ. The Poisson point process Π_λ can be represented by an infinite sequence $(\mathbf{X}_i)_{i\geqslant 1}$ of random points of $\mathbb{Z}^2$. For $\mathbf{x} \in \Theta$, let us introduce

$$P_\lambda(\mathbf{x}) = |\{i \geqslant 1 \text{ such that } \mathbf{X}_i = \mathbf{x}\}|$$

which follows the Poisson distribution $\mathcal{P}(\lambda)$. Notice that, in this case, the $\mathbf{X}_i$'s cannot be supposed to be independent and identically distributed because there is no uniform probability distribution on $\mathbb{Z}^2$.

Since $h \in \ell^1$ and since P_λ is a Poisson white noise, Proposition 2.2.2 shows that

$$\sum_{\mathbf{y}} h(\mathbf{x}-\mathbf{y})P_\lambda(\mathbf{y})$$

is almost surely absolutely convergent. Thus we can give the following definition.

Definition 2.2.5. The discrete spot noise (DSN) (or discrete Poisson spot noise) of intensity λ associated to h is the random field $F_{h,\lambda} : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$ defined by

$$\forall \mathbf{x} \in \mathbb{Z}^2, \quad F_{h,\lambda}(\mathbf{x}) = \sum_{i\geqslant 1} h(\mathbf{x}-\mathbf{X}_i)\ . \tag{2.22}$$

or equivalently by

$$\forall \mathbf{x} \in \mathbb{Z}^2, \quad F_{h,\lambda}(\mathbf{x}) = \sum_{\mathbf{y}} h(\mathbf{x}-\mathbf{y})P_\lambda(\mathbf{y}) = h * P_\lambda(\mathbf{x})\ .$$

Notice that $F_{h,\lambda}$ is a rank-one colored noise in the sense of Definition 2.2.4. Thanks to Proposition 2.2.1, it has mean $\lambda \sum h$ and covariance function $\lambda\, h * \tilde{h}^T$ which is the autocorrelation of h multiplied by λ.

Gaussian convergence at high intensity

Definition 2.2.6. The renormalized discrete spot noise of intensity λ associated to h is the random field $G_{h,\lambda} : \mathbb{Z}^2 \longrightarrow \mathbb{R}$ defined by

$$G_{h,\lambda} = \frac{F_{h,\lambda} - \mathbb{E}(F_{h,\lambda})}{\sqrt{\lambda}} = \frac{1}{\sqrt{\lambda}}\left(h * P_\lambda - \lambda \sum h\right) = \frac{1}{\sqrt{\lambda}}\, h * (P_\lambda - \lambda)\ .$$

It is a stationary random field with mean 0 and covariance function $h * \tilde{h}^T$.

Theorem 2.2.3 ([Papoulis 1971])**.** *When $\lambda \to \infty$, $G_{h,\lambda}$ converges in distribution (in the sense of finite-dimensional marginal distributions) to the Gaussian random field of mean 0 and covariance function $h * \tilde{h}^T$.*

Proof. The proof follows the same sketch as the one given in [Rice 1977]. It amounts to show that the marginal distribution of $G_{h,\lambda}$ converges to the one of $\mathcal{N}(0, h * \tilde{h}^T)$ using Lévy's theorem about characteristic functions, and then the convergence of all the finite-dimensional marginal distributions follows from the linearity of the spot noise process. But here, since we only deal with discrete spot noises, computing the characteristic function of $G_{h,\lambda}(0)$ does not require Campbell's theorem.

So, first, let us show that

$$G_{h,\lambda}(0) \xrightarrow[\lambda\to+\infty]{(d)} \mathcal{N}(0, h * \tilde{h}^T(0)) \ .$$

Thanks to Lévy's theorem, we only have to show that for a fixed $u \in \mathbb{R}^d$,

$$\mathbb{E}\left[\exp\left(iu^T G_{h,\lambda}(0)\right)\right] \xrightarrow[\lambda\to+\infty]{(d)} \exp\left(-\frac{1}{2}u^T h * \tilde{h}^T(0)u\right) .$$

Let us compute the characteristic function of $G_{h,\lambda}(0)$. By definition,

$$G_{h,\lambda}(0) = \sum_{\mathbf{x}\in\mathbb{Z}^2} h(\mathbf{x})\left(\frac{P_\lambda(-\mathbf{x})-\lambda}{\sqrt{\lambda}}\right)$$

where the series converges almost surely absolutely. Let A be a finite subset of $\mathbb{Z}^2$. Since the $(P_\lambda(\mathbf{x}))_{\mathbf{x}\in A}$ are independent, we have

$$\mathbb{E}\left[\exp\left(iu^T \sum_{\mathbf{x}\in A} h(\mathbf{x})\ \frac{P_\lambda(-\mathbf{x})-\lambda}{\sqrt{\lambda}}\right)\right] = \prod_{\mathbf{x}\in A} \mathbb{E}\left[\exp\left(iu^T h(\mathbf{x})\ \frac{P_\lambda(-\mathbf{x})-\lambda}{\sqrt{\lambda}}\right)\right]$$

Moreover, since the characteristic function of a real random variable $Q \sim \mathcal{P}(\lambda)$ is given by $t \mapsto e^{\lambda(e^{it}-1)}$, we have for all $s \in \mathbb{R}$,

$$\begin{aligned}
\mathbb{E}\left[\exp\left(is\ \frac{Q-\lambda}{\sqrt{\lambda}}\right)\right] &= e^{-is\sqrt{\lambda}}\,\mathbb{E}\left[\exp\left(i\,\frac{s}{\sqrt{\lambda}}\,Q\right)\right] \\
&= \exp\left(-is\sqrt{\lambda}\right)\exp\left(\lambda\left(e^{i\frac{s}{\sqrt{\lambda}}}-1\right)\right) \\
&= \exp\left[\lambda\left(e^{i\frac{s}{\sqrt{\lambda}}}-1-i\,\frac{s}{\sqrt{\lambda}}\right)\right] .
\end{aligned}$$

Applying this equality for each $s = u^T h(\mathbf{x})$, and taking the product on $\mathbf{x}$, we get

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(iu^T \sum_{\mathbf{x}\in A} h(\mathbf{x})\,\frac{P_\lambda(-\mathbf{x})-\lambda}{\sqrt{\lambda}}\right)\right] \\
&\qquad = \exp\left(\lambda \sum_{\mathbf{x}\in A} \exp\left(i\ \frac{u^T h(\mathbf{x})}{\sqrt{\lambda}}\right) - 1 - i\,\frac{u^T h(\mathbf{x})}{\sqrt{\lambda}}\right) . \qquad (2.23)
\end{aligned}$$

We will let $A \to \mathbb{Z}^2$ in this equality. Since $\sum_{\mathbf{x}\in\mathbb{Z}^2} h(\mathbf{x})\,\frac{P_\lambda(-\mathbf{x})-\lambda}{\sqrt{\lambda}}$ is almost surely absolutely convergent, the dominated convergence theorem ensures that the left-hand side tends to

$$\mathbb{E}\left[\exp\left(iu^T \sum_{\mathbf{x}\in\mathbb{Z}^2} h(\mathbf{x})\,\frac{P_\lambda(-\mathbf{x})-\lambda}{\sqrt{\lambda}}\right)\right] .$$

Moreover, using

$$\forall t \in \mathbb{R}, \quad |e^{it} - 1 - it| \leqslant \frac{t^2}{2} ,$$

we get that

$$\sum_{\mathbf{x}\in\mathbb{Z}^2} \left| \exp\left(i \, \frac{u^T h(\mathbf{x})}{\sqrt{\lambda}} \right) - 1 - i \, \frac{u^T h(\mathbf{x})}{\sqrt{\lambda}} \right| \leqslant \frac{1}{2\lambda} \sum_{\mathbf{x}\in\mathbb{Z}^2} (u^T h(\mathbf{x}))^2 \tag{2.24}$$

$$\leqslant \frac{1}{2\lambda} \|u\|^2 \sum_{\mathbf{x}\in\mathbb{Z}^2} \|h(\mathbf{x})\|^2 < \infty , \tag{2.25}$$

so that

$$\sum_{\mathbf{x}\in\mathbb{Z}^2} \exp\left(i \, \frac{u^T h(\mathbf{x})}{\sqrt{\lambda}} \right) - 1 - i \, \frac{u^T h(\mathbf{x})}{\sqrt{\lambda}}$$

is absolutely convergent. Letting $A \to \mathbb{Z}^2$ in (2.23), we obtain the characteristic function of $G_{h,\lambda}(0)$ as

$$\mathbb{E}\left[\exp\left(i u^T G_{h,\lambda}(0) \right) \right] = \exp\left(\lambda \sum_{\mathbf{x}\in\mathbb{Z}^2} \left(\exp\left(i \, \frac{u^T h(\mathbf{x})}{\sqrt{\lambda}} \right) - 1 - i \, \frac{u^T h(\mathbf{x})}{\sqrt{\lambda}} \right) \right) \tag{2.26}$$

A Taylor-Young expansion ensures that for each $\mathbf{x} \in \mathbb{Z}^2$,

$$\lambda \left(\exp\left(i \, \frac{u^T h(\mathbf{x})}{\sqrt{\lambda}} \right) - 1 - i \, \frac{u^T h(\mathbf{x})}{\sqrt{\lambda}} \right) \xrightarrow[\lambda\to+\infty]{} -\frac{1}{2} (u^T h(\mathbf{x}))^2 .$$

and with the domination given by (2.24), the dominated convergence theorem gives

$$\mathbb{E}\left[\exp\left(i u^T G_{h,\lambda}(0) \right) \right] \xrightarrow[\lambda\to+\infty]{} \exp\left(-\frac{1}{2} \sum_{\mathbf{x}\in\mathbb{Z}^2} (u^T h(\mathbf{x}))^2 \right) = \exp\left(-\frac{1}{2} u^T h * \tilde{h}^T(0) u \right) ,$$

and the right-hand side is exactly the characteristic function of $\mathcal{N}(0, h * \tilde{h}(0))$.

To show that all finite-dimensional marginal distributions converge, we only need to show that for any compactly-supported kernel $k : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$, the random variable $k^T * G_{h,\lambda}(0)$ converges in distribution to a Gaussian. But, thanks to the linearity of the spot noise construction, we have

$$k^T * G_{h,\lambda} \overset{(d)}{=} G_{k^T * h,\lambda}$$

and therefore, the first part of the proof shows that $k^T * G_{h,\lambda}(0)$ converges in distribution to $\mathcal{N}(0, k^T * h * \tilde{h}^T * \tilde{k}(0))$. In conclusion, when $\lambda \to +\infty$, $G_{h,\lambda}$ converges in finite-dimensional marginal distributions to $\mathcal{N}(0, h * \tilde{h}^T)$. □

Definition 2.2.7. The asymptotic discrete spot noise (ADSN) on $\mathbb{Z}^2$ associated to the kernel function $h \in \ell^1$ is the random field

$$G_h = h * W \sim \mathcal{N}(0, h * \tilde{h}^T) ,$$

where W is a scalar Gaussian white noise on $\mathbb{Z}^2$ with mean 0 and variance 1.

Notice that G_h is another example of rank-one colored noise. Notice also that, thanks to Proposition 2.2.2, the random fields $h * W$ is defined as soon as $h \in \ell^2$. But if $h \in \ell^2 \backslash \ell^1$ the discrete spot noise associated to h is not defined so that $h * W$ cannot be considered as an ADSN. That is why, in the case $h \in \ell^2 \backslash \ell^1$, $h * W$ will not be called ADSN, but only Gaussian random field with mean 0 and covariance function $h * \tilde{h}^T$.

2.2.3 Can we define a Random Phase Noise on $\mathbb{Z}^2$?

As we have seen, the construction of an ADSN field on $\mathbb{Z}^2$ is nearly as easy as the one of the circular ADSN on a finite domain Θ. In contrast, the existence of a RPN field on $\mathbb{Z}^2$ is a more difficult question.

Let us observe that, in the simple case of a function $h : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$ which is (M, N)-periodic, h identifies to a function h^Θ defined on $\Theta = \mathbb{Z}/M\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$ and the RPN associated to h^Θ on Θ gives a natural definition of a (M, N)-periodic RPN on $\mathbb{Z}^2$. According to Remark 2.2.1, its spectral measure is

$$\frac{(2\pi)^2}{MN} \sum_{\boldsymbol{\xi} \in \Theta} \widehat{h^\Theta}(\boldsymbol{\xi}) \widehat{h^\Theta}(\boldsymbol{\xi})^* \delta_{2\pi\left(\frac{\xi_1}{M}, \frac{\xi_2}{N}\right)} \ , \tag{2.27}$$

where $\widehat{h^\Theta}$ is the DFT of h^Θ.

More generally, let us assume that $\hat{h}$ is a Borelian measure $\nu = \sum_{\boldsymbol{\xi} \in \Xi} \nu(\boldsymbol{\xi}) \delta_{\boldsymbol{\xi}}$ supported on a finite subset Ξ of $\mathbb{T}^2$, which means that h is a sum of pure waves:

$$\forall \mathbf{x} \in \mathbb{Z}^2, \quad h(\mathbf{x}) = \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} e^{i\boldsymbol{\xi}.\mathbf{x}} d\nu(\boldsymbol{\xi}) = \frac{1}{(2\pi)^2} \sum_{\boldsymbol{\xi} \in \Xi} \nu(\boldsymbol{\xi}) e^{i\boldsymbol{\xi}.\mathbf{x}} \ .$$

Then one can still define a RPN associated to h over $\mathbb{Z}^2$. Indeed, using a family $(\psi(\boldsymbol{\xi}))_{\boldsymbol{\xi} \in \Xi}$ of uniform random variables on $\mathbb{T}$ which satisfies

- $\forall \boldsymbol{\xi} \in \Xi, \quad \psi(-\boldsymbol{\xi}) = -\psi(\boldsymbol{\xi})$,
- $\psi(\boldsymbol{\xi}) \sim \mathcal{U}(\mathbb{T})$ if $\boldsymbol{\xi} \neq -\boldsymbol{\xi}$ and $\psi(\boldsymbol{\xi}) \sim \mathcal{U}(\{0, \pi\})$ if $\boldsymbol{\xi} = -\boldsymbol{\xi}$,
- $(\psi(\boldsymbol{\xi}))_{\boldsymbol{\xi} \in A}$ are independent as soon as $A \cap (-A) = \varnothing$,

we can define $R_h : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$ by

$$\forall \mathbf{x} \in \mathbb{Z}^2, \quad R_h(\mathbf{x}) = \frac{1}{2\pi} \sum_{\boldsymbol{\xi} \in \Xi} \nu(\boldsymbol{\xi}) e^{i\boldsymbol{\xi}.\mathbf{x} + i\psi(\boldsymbol{\xi})} \ . \tag{2.28}$$

One can show with a proof similar to the one of Proposition 2.1.2 that R_h is a stationary random field on $\mathbb{Z}^2$. Since $\mathbb{E}(e^{i\psi(\boldsymbol{\xi})}) = 0$, we get that $\mathbb{E}(R_h) = 0$. Furthermore, using that $\mathbb{E}(e^{i\psi(\boldsymbol{\xi})} e^{-i\psi(\boldsymbol{\zeta})}) = 0$ as soon as $\boldsymbol{\xi} \neq \boldsymbol{\zeta}$, we get

$$\begin{aligned} \mathbb{E}\Big(R_h(\mathbf{x}) R_h(\mathbf{y})^T\Big) &= \frac{1}{(2\pi)^2} \sum_{\boldsymbol{\xi}, \boldsymbol{\zeta} \in \Xi} \nu(\boldsymbol{\xi}) \nu(\boldsymbol{\zeta})^* e^{i\boldsymbol{\xi}.\mathbf{x}} e^{-i\boldsymbol{\zeta}.\mathbf{y}} \mathbb{E}\Big(e^{i\psi(\boldsymbol{\xi})} e^{-i\psi(\boldsymbol{\zeta})}\Big) \ . \\ &= \frac{1}{(2\pi)^2} \sum_{\boldsymbol{\xi} \in \Xi} \nu(\boldsymbol{\xi}) \nu(\boldsymbol{\xi})^* e^{i\boldsymbol{\xi}.(\mathbf{x}-\mathbf{y})} \ , \end{aligned}$$

and thus, the covariance function of R_h is

$$C_{R_h}(\mathbf{v}) = \frac{1}{(2\pi)^2} \sum_{\boldsymbol{\xi} \in \Xi} \nu(\boldsymbol{\xi}) \nu(\boldsymbol{\xi})^* e^{i\boldsymbol{\xi}.\mathbf{v}}$$

(which actually defines the autocorrelation $h * \tilde{h}^T$ of h), and its spectral measure is

$$\sum_{\boldsymbol{\xi} \in \Xi} \nu(\boldsymbol{\xi}) \nu(\boldsymbol{\xi})^* \delta_{\boldsymbol{\xi}} \, . \tag{2.29}$$

Notice that, again in this case, the random field R_h was obtained by multiplying the Fourier transform of h not only by $e^{i\psi(\boldsymbol{\xi})}$ but by $2\pi e^{i\psi(\boldsymbol{\xi})}$ where 2π is the square root of the volume of $\mathbb{T}^2$. One must also be aware of the difference in normalization of the spectral measures given by (2.27) and (2.29). They come from the different normalizations adopted for the continuous and discrete Fourier transforms.

But is it possible to give a RPN definition for a wider class of kernel functions? For now, we have not a clear answer to this question but the following of this subsection shows, based on a result of [Desolneux *et al.* 2015], that a frequency-sampling approach would result only in a globally Gaussian process.

Notice that if $\hat{h}$ is a continuous function on $\mathbb{T}^2$, it can be sampled on any discrete subset of $\mathbb{T}^2$. Thus, if $\Xi = \Xi_{M,N}$ is the finite subgroup of $\mathbb{T}^2$ of size $M \times N$ spanned by $(\frac{2\pi}{M}, 0)$ and $(0, \frac{2\pi}{N})$, we can define a function $g_\Xi : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$ by

$$\forall \mathbf{x} \in \mathbb{Z}^2, \quad g_\Xi(\mathbf{x}) = \frac{1}{2\pi} \frac{1}{\sqrt{|\Xi|}} \sum_{\boldsymbol{\xi} \in \Xi} \hat{h}(\boldsymbol{\xi}) e^{i\boldsymbol{\xi}.\mathbf{x}} \, . \tag{2.30}$$

Let us recall that g_Ξ can be seen as a (M, N)-periodic version of h. The above discussion ensures that R_{g_Ξ} is a stationary random field on $\mathbb{Z}^2$ with zero mean and covariance function

$$C_{R_{g_\Xi}}(\mathbf{v}) = \frac{1}{(2\pi)^2} \frac{1}{|\Xi|} \sum_{\boldsymbol{\xi} \in \Xi} \hat{h}(\boldsymbol{\xi}) \hat{h}(\boldsymbol{\xi})^* e^{i\boldsymbol{\xi}.\mathbf{v}} \, .$$

Notice that the right-hand side is a Riemann sum, and thus, when M and N tends to infinity, this covariance function tends to

$$C(\mathbf{v}) = \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} \hat{h}(\boldsymbol{\xi}) \hat{h}(\boldsymbol{\xi})^* e^{i\boldsymbol{\xi}.\mathbf{v}} d\boldsymbol{\xi} = h * \tilde{h}^T(\mathbf{v}) \, .$$

Therefore, when the frequency sampling gets finer and finer, the second order moments of the corresponding RPN converge. The next theorem, whose proof can be found in [Desolneux *et al.* 2015], shows that these RPN converge in a stronger sense.

Theorem 2.2.4 ([Desolneux *et al.* 2015])**.** *Let $h : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$ such that $\hat{h}$ is a continuous function. Let us consider the finite subgroup $\Xi_{M,N}$ of $\mathbb{T}^2$ spanned by $(\frac{2\pi}{M}, \frac{2\pi}{N})$. Recall that $g_{\Xi_{M,N}} : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$ is defined by (2.30) and that the corresponding RPN is defined by (2.28). If $M, N \to \infty$, then the RPN associated to $g_{\Xi_{M,N}}$ converges in finite-dimensional marginal distributions to $\mathcal{N}(0, h * \tilde{h}^T)$.*

This theorem shows that a process over $\mathbb{Z}^2$ obtained as a limit in distribution of densely frequency-sampled RPN would necessarily be Gaussian. Indeed, let us assume that for a certain class $\mathfrak{R}$ of random fields, we can build, for each finite subset Ξ of $\mathbb{T}^2$, an operator $\mathcal{L}_\Xi : \mathfrak{R} \longrightarrow \mathfrak{R}$ that would perform frequency-sampling on the grid Ξ in such a manner that for each $F \in \mathfrak{R}$, the random field $\mathcal{L}_\Xi F$ converges in distribution to F when the sampling becomes denser and denser. Then, a natural condition that could be required for a RPN process R associated to h over $\mathbb{Z}^2$ would be that its frequency-sampled version $R_{M,N} = \mathcal{L}_{\Xi_{M,N}} R$ over Ξ_n coincides with the RPN associated with the frequency-sampled kernel function $g_{\Xi_{M,N}}$. The last result shows that in these conditions, we would necessarily have $R \sim \mathcal{N}(0, h * \tilde{h}^T)$.

2.2.4 Gaussian Models on $\mathbb{Z}^2$

In this subsection, we present and discuss different stationary Gaussian models that appear in the texture modelling literature. Since Gaussian random fields are characterized by their covariance, the comparison only concerns the second-order moments and the spectral measures.

Gaussian Moving Average Fields

The term "moving average" originally appeared in time series analysis [Yule 1921], [Wold 1938, p.51]. It then referred to random sequences $F : \mathbb{Z} \longrightarrow \mathbb{R}$ of the form

$$F(t) = h(0)W(t) + h(1)W(t-1) + \ldots + h(q)W(t-q) \ ,$$

where $W : \mathbb{Z} \longrightarrow \mathbb{R}$ is such that the $W(t)$ are independent and where $h(0), \ldots, h(q)$ are deterministic real coefficients. Such processes are thus particular cases of rank-one colored noises $h * W$ where $h : \mathbb{Z} \longrightarrow \mathbb{R}$ is a causal compactly-supported filter (meaning that $\mathrm{Supp}(h)$ is a finite subset of $\mathbb{N}$).

But unfortunately, authors of later articles on this topic did not agree on a universal definition of moving average processes. For example, the author of [Doob 1949] and [Doob 1990] calls moving average process any random sequence $F : \mathbb{Z} \longrightarrow \mathbb{R}$ which can be written $h * W$ where $h \in \ell^2$, and where W is a white noise. This corresponds to the univariate case of the processes given by Proposition 2.2.2. For the multivariate case, one can find in [Helson & Lowdenslager 1961, p.201] a definition of vectorial moving average processes as $h * \mathcal{W}$ where $h \in \ell^2(\mathbb{Z}^2, \mathbb{R}^{d \times d})$ and where $\mathcal{W}$ is a multi-channel white noise. The author of [Eom 2000] also uses a moving average model but only in a circular framework.

In the univariate case, the causality of the filter h has a simple chronological interpretation. In the two-variable case, since there is no canonical order on $\mathbb{Z}^2$, the causality assumption has no temporal interpretation and in particular is *a priori* not relevant in an image processing context. However, the article [Francos & Friedlander 1998] still gives a definition of moving average fields on $\mathbb{Z}^2$ with a filter h that is causal for a particular order on $\mathbb{Z}^2$; these causal moving average fields plays a crucial role in the prediction theory, in relation with the Wold decomposition [Francos *et al.* 1993].

In this thesis (and particularly in Section 3.3), we will frequently encounter Gaussian random fields of the form $h * W$ where $h : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$ has a compact support and where W is a Gaussian white noise of variance 1. Since it is a very simple case of colored noise, such a moving average random field has mean zero, covariance function $h*\tilde{h}^T$, and spectral density $\hat{h}\hat{h}^*$. Notice that $\hat{h}$ is a trigonometric polynomial function whose degree depends on the size of the support of h. It would be very convenient for this thesis to call these random fields Gaussian moving average fields but since the term is already quite ambiguous and overloaded, we prefer not to use it.

The following proposition (which is the Gaussian multivariate case of a result given in [Doob 1949, p.327] and [Doob 1990, p.498]) characterizes the stationary Gaussian random fields with spectral density.

Proposition 2.2.3 (Doob)**.** *A Gaussian random process $F : \mathbb{Z}^2 \longrightarrow \mathbb{R}$ with mean zero and covariance function C admits a spectral density if and only if it has the same distribution as $h * W$ where $h \in \ell^2$ and W is a Gaussian white noise.*

Proof. The reciprocal implication was already shown in the end of Subsection 2.2.1. For the direct one, let us assume that F admits a spectral density φ (which is necessarily a non-negative integrable function over $\mathbb{T}^2$), that is

$$\forall \mathbf{v} \in \mathbb{Z}^2, \quad C(\mathbf{v}) = \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} \varphi(\boldsymbol{\xi}) e^{i\boldsymbol{\xi}.\mathbf{x}} d\boldsymbol{\xi} \ .$$

Let us remark that $\sqrt{\varphi} \in L^2(\mathbb{T}^2)$ so that

$$h(\mathbf{x}) = \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} \sqrt{\varphi(\boldsymbol{\xi})}\ e^{i\boldsymbol{\xi}.\mathbf{x}} d\boldsymbol{\xi}$$

defines $h \in \ell^2$ such that $\hat{h} \geqslant 0$, $\hat{h}^2 = \widehat{C}$. Thus, $\tilde{h} = h$ and $h * h = C$. Since F is Gaussian, we get that F has the same distribution as $h * W$ where W is a Gaussian white noise on $\mathbb{Z}^2$ with variance 1. □

Gaussian Autoregressive Fields

Several authors suggest to model random fields with spatial interactions as autoregressive fields [Mead 1971], [Ord 1975]. These random fields have been studied for example in [Doob 1944], [Whittle 1954] and [Besag & Kooperberg 1995]. These models have been applied to texture synthesis in [Chellappa & Kashyap 1985]. Since these articles concentrate on real-valued random fields, in this paragraph we will assume $d = 1$. We will see that the stationary autoregressive fields with spectral density can be written in the form $h * W$ where $h \in \ell^2$ has a Fourier transform given by an inverse trigonometric polynomial.

As for the moving average model, the literature seems a bit confuse about the definition of autoregressive model. Let us only mention here that the authors of [Chellappa & Kashyap 1985] define a (non-causal) autoregressive random field to be a stationary random field $F : \mathbb{Z}^2 \longrightarrow \mathbb{R}$, which satisfies

$$\forall \mathbf{x} \in \mathbb{Z}^2, \quad F(\mathbf{x}) = \sum_{\mathbf{y} \in \mathbb{Z}^2} \theta(\mathbf{y}) F(\mathbf{x} - \mathbf{y}) \ + \ W(\mathbf{x}) \ , \tag{2.31}$$

where $\theta : \mathbb{Z}^2 \longrightarrow \mathbb{R}$ has a finite support with $\theta(0) = 0$, and where W is a (non-necessarily Gaussian) pure white noise with mean zero and $\beta = \mathbb{E}(W(0)^2) > 0$. In other words,

$$F = \theta * F + W \ .$$

In particular, when W is a white Gaussian noise, we get a Gaussian autoregressive field on $\mathbb{Z}^2$. Notice that in this case, F is not directly given by a convolution $h * W'$ of h with a Gaussian white noise W'.

Taking the covariance functions, we get

$$(\delta_0 - \theta) * (\delta_0 - \tilde{\theta}) * C_F = \beta \delta_0 \ .$$

This implies that the spectral measure μ_F satisfies

$$|1 - \hat{\theta}(\boldsymbol{\xi})|^2 \mu_F(d\boldsymbol{\xi}) = \beta d\boldsymbol{\xi} \ .$$

If F has a spectral density φ, we thus have for almost all $\boldsymbol{\xi} \in \mathbb{T}^2$,

$$\varphi(\boldsymbol{\xi}) = \frac{\beta}{|1 - \hat{\theta}(\boldsymbol{\xi})|^2} \ , \tag{2.32}$$

so that φ is the inverse of the squared modulus of a trigonometric polynomial (in two variables).

Let us remark that the existence of such an autoregressive process with spectral density imposes some constraints on θ so that (2.32) actually defines an integrable function. For example, if $\theta = \frac{1}{2}(\delta_{(-1,0)} + \delta_{(1,0)})$, then $\hat{\theta}(\boldsymbol{\xi}) = \cos(\xi_1)$ so that when $\boldsymbol{\xi} \to 0$,

$$|1 - \hat{\theta}(\boldsymbol{\xi})|^2 = (1 - \cos(\xi_1))^2 \sim \frac{\xi_1^4}{4},$$

and thus $\frac{1}{|1-\hat{\theta}|^2} \notin L^1(\mathbb{T}^2)$.

As mentioned in [Woods 1972] a sufficient condition for the existence of a Gaussian autoregressive process F satisfying (2.31) and which has a spectral density φ is that

$$\forall \boldsymbol{\xi} \in \mathbb{T}^2, \quad 1 - \hat{\theta}(\boldsymbol{\xi}) \neq 0 \ . \tag{2.33}$$

Indeed, in that case, $\frac{1}{1-\hat{\theta}}$ is a continuous function on $\mathbb{T}^2$ and in particular square-integrable, so that we can define a kernel function $h \in \ell^2$ by its Fourier transform

$$\hat{h} = \frac{1}{1 - \hat{\theta}} \in \ell^2 \ .$$

Then, thanks to Proposition 2.2.2, if W is a Gaussian white noise with variance $\sqrt{\beta}$, $h * W$ is a well-defined Gaussian process with mean zero and its spectral density is exactly the right-hand side of (2.32). Using Lemma 2.2.2, we get $(1 - \theta) * h = \delta_0$, and thus

$$(1 - \theta) * F = (1 - \theta) * h * W = \delta_0 * W = W$$

so that F satisfies (2.31).

In conclusion, we see that a Gaussian autoregressive model (2.31) such that $\hat{\theta}(\boldsymbol{\xi}) \neq 1$ for all $\boldsymbol{\xi} \in \mathbb{T}^2$ is a particular case of Gaussian rank-one colored noise $h * W$ where $h \in \ell^2$.

Gaussian Markov Random Fields

In this paragraph, we investigate, in the case of zero mean real-valued stationary Gaussian random fields, the difference between autoregressive fields and Markov random fields. The following argument was already suggested in [Chellappa 1985] but was not used for comparison with the autoregressive model.

Let $F : \mathbb{Z}^2 \longrightarrow \mathbb{R}$ be a stationary Gaussian random field with zero mean and covariance function C. Assume that F satisfies the Markov property which means that there exists a finite $N \subset \mathbb{Z}^2$ such that $0 \notin N$ and such that the distribution of $F(\mathbf{x})$ conditionally to $(F(\mathbf{y}))_{\mathbf{y}\notin\mathbf{x}}$ only depends on $(F(\mathbf{y}))_{\mathbf{y}\in\mathbf{x}-N}$. Since $(F(\mathbf{y}), \mathbf{y} \in \{\mathbf{x}\} \cup (\mathbf{x} - N))$ is a Gaussian vector, there exist coefficients $(\eta_{\mathbf{x}}(\mathbf{y}))_{\mathbf{y}\in N}$ such that

$$\mathbb{E}\Big(F(\mathbf{x}) \mid F(\mathbf{y}), \mathbf{y} \in \mathbf{x} - N\Big) = \sum_{\mathbf{y}\in N} \eta_{\mathbf{x}}(\mathbf{y}) F(\mathbf{x} - \mathbf{y}) \ .$$

Besides, since F is stationary, the coefficients $\eta_{\mathbf{x}}$ do not depend on $\mathbf{x}$. Therefore, there exists a function $\eta : \mathbb{Z}^2 \longrightarrow \mathbb{R}$ with support in N such that for all $\mathbf{x} \in \mathbb{Z}^2$,

$$\mathbb{E}\Big(F(\mathbf{x}) \mid F(\mathbf{y}), \mathbf{y} \neq \mathbf{x}\Big) = \mathbb{E}\Big(F(\mathbf{x}) \mid F(\mathbf{y}), \mathbf{y} \in \mathbf{x} - N \Big) = \sum_{\mathbf{y}\in\mathbb{Z}^2} \eta(\mathbf{y}) F(\mathbf{x} - \mathbf{y}) \ .$$

In particular, for all $\mathbf{x} \in \mathbb{Z}^2$ and $\mathbf{v} \neq 0$,

$$\begin{aligned}
\mathbb{E}(F(\mathbf{x})F(\mathbf{x}+\mathbf{v})) &= \mathbb{E}\Big[\mathbb{E}\big(F(\mathbf{x}) \mid F(\mathbf{y}), \mathbf{y} \neq \mathbf{x}\big) F(\mathbf{x}+\mathbf{v})\Big] \\
&= \mathbb{E}\left[\left(\sum_{\mathbf{y}\in\mathbb{Z}^2} \eta(\mathbf{y}) F(\mathbf{x}-\mathbf{y})\right) F(\mathbf{x}+\mathbf{v})\right] \\
&= \sum_{\mathbf{y}\in\mathbb{Z}^2} \eta(\mathbf{y}) \mathbb{E}(F(\mathbf{x}-\mathbf{y})F(\mathbf{x}+\mathbf{v})) \ ,
\end{aligned}$$

so that

$$\forall \mathbf{v} \neq 0, \quad C(\mathbf{v}) = \sum_{\mathbf{y}\in\mathbb{Z}^2} \eta(\mathbf{y}) C(\mathbf{v}+\mathbf{y}) = \tilde{\eta} * C(\mathbf{v}) \ .$$

In other words, $(\delta_0 - \tilde{\eta}) * C = C - \tilde{\eta} * C = \lambda\delta_0$ where $\lambda = C(0) - \tilde{\eta} * C(0)$. As in the case of the autoregressive fields, this implies that the spectral measure μ_F satisfies

$$(1 - \hat{\eta}(\boldsymbol{\xi})^*)\mu_F(d\boldsymbol{\xi}) = \lambda d\boldsymbol{\xi} \ .$$

And therefore, if F has a spectral density φ, it is given by

$$\varphi(\boldsymbol{\xi}) = \frac{\lambda}{1 - \hat{\eta}(\boldsymbol{\xi})^*} \ ,$$

which is the inverse of a trigonometric polynomial. Again in this case, the integrability of $\frac{1}{1-\hat{\eta}^*}$ is a necessary condition for the existence of such a Markov random field with spectral density.

The last calculation showed that if $\lambda \neq 0$, the covariance function C admits a convolution inverse given by $\frac{1}{\lambda}(\delta_0 - \tilde{\eta})$ which has compact support. By analogy with the term "precision matrix" that is sometimes used in the literature, the inverse

(for the convolution) of the covariance function can be called precision function. Therefore, we have seen that, except in the case $\lambda = 0$ (which can be thought of as a degenerate case), a stationary Gaussian random field on $\mathbb{Z}^2$ which has the Markov property admits a compactly-supported precision function which does not vanish at 0.

To end this paragraph, we study the links between autoregressive models and Markov models. With the same notations as above, let us now consider the stationary Gaussian random field

$$U(\mathbf{x}) = F(\mathbf{x}) - \sum_{\mathbf{y}\in\mathbb{Z}^2} \eta(\mathbf{y})F(\mathbf{x}-\mathbf{y}) = (\delta_0 - \eta) * F(\mathbf{x}) \ .$$

Therefore, we have

$$\forall \mathbf{x} \in \mathbb{Z}^2, \quad F(\mathbf{x}) = \sum_{\mathbf{y}\in\mathbb{Z}^2} \eta(\mathbf{y})F(\mathbf{x}-\mathbf{y}) + U(\mathbf{x}) \ , \tag{2.34}$$

where U is a stationary Gaussian random field with covariance function

$$C_U = (\delta_0 - \tilde{\eta}) * (\delta_0 - \eta) * C = \lambda(\delta_0 - \eta) \ .$$

Besides, thanks to the definition of the conditional expectation, we know that for every $\mathbf{x} \in \mathbb{Z}^2$, $U(\mathbf{x})$ is orthogonal to

$$\mathrm{Span}\Big(F(\mathbf{y}) \ , \ \mathbf{y} \neq \mathbf{x}\Big) \subset L^2(\mathbb{P}) \ .$$

Since $(F, U(\mathbf{x}))$ is jointly Gaussian, we get that for all $\mathbf{x}$, $U(\mathbf{x})$ is independent of $(F(\mathbf{y}))_{\mathbf{y}\notin\mathbf{x}}$.

Let us notice that Equation (2.34) looks like Equation (2.31) but is actually very different. Indeed, as can be seen on the expression of its covariance, U is never a white noise except in the trivial case $\eta = 0$ for which $F = U$ is also a Gaussian white noise. Actually, if F satisfies the autoregression equations (2.31), then $\sum_{\mathbf{y}\in\mathbb{Z}^2} \theta(\mathbf{y})F(\mathbf{x}-\mathbf{y})$ is *not* the expectation of $F(\mathbf{x})$ conditionally on $(F(\mathbf{y}))_{\mathbf{y}\neq\mathbf{x}}$, because we cannot say that $W(\mathbf{x})$ is independent of $(F(\mathbf{y}))_{\mathbf{y}\neq\mathbf{x}}$. Therefore, there is no reason to have θ and η equal. If it was the case, $U = F - \eta * F = F - \theta * F = W$ would be a white noise, and so would F as we have just said. For the same reason, it is unclear that an autoregressive field defined by (2.31) has the Markov property.

According to [Besag & Kooperberg 1995, Section 2], defining a Gaussian random vector $(X_1, \ldots, X_n)$ by conditional autoregression (or auto-normal formulation) amounts to suppose that for every i, the conditional distribution of X_i given all the other values $(X_j)_{j\neq i}$ is

$$X_i \mid (X_j)_{j\neq i} \ \sim \ \mathcal{N}\Big(\mu_i + \sum_j \beta_{ij} X_j, \kappa_i\Big) \ ,$$

with $\kappa_i > 0$. Setting $Q_{ii} = \frac{1}{\kappa_i} > 0$ and for $i \neq j$, $Q_{ij} = -\frac{\beta_{ij}}{\kappa_i}$, one can show that this Gaussian random vector has covariance matrix Q. Actually, as mentioned at the end of [Li 2009, Section 2.2], we can obtain with this construction any multivariate

Gaussian random vector so that the conditional autoregression point of view is just another way to describe the class of Gaussian random vectors. However, it may not be the case for random fields defined on the infinite lattice $\mathbb{Z}^2$, and as we have seen above, the Gaussian autoregressive model and Gaussian Markov random field model must not be confused.

Additional remarks

In general, moving averages are adapted to the case of a smooth spectral density (given by a trigonometric polynomial), whereas autoregressive or Markov random fields have irregular spectral measures ((squared) inverse of a trigonometric polynomial). For the case of real-valued Gaussian fields with spectral density, those models can be represented as $\mathcal{N}(0, h * \tilde{h})$ for a well-chosen $h \in \ell^2$.

Notice that the terms "moving average" and "autoregressive" are not limited to the case of Gaussian random fields. But for a general moving average or autoregressive random field, the final distribution not only depends on the corresponding filters but also on the excitation function. So one cannot analyze clearly the model limits until a proper form of excitation function is specified. Thus, for the sake of simplicity, we chose to restrict here to Gaussian excitation functions which already leads to a quite wide texture model. Notice that in [Chellappa & Kashyap 1985], even if non-Gaussian excitation functions are mentioned in Subsection II.A, the simulation in Subsection II.B is restricted to the Gaussian case.

Let us also remark that the articles [Chellappa & Kashyap 1985], [Cadzow *et al.* 1993], and [Eom 2000] restrict to the case of circular random fields when it comes to practical aspects (simulation or estimation of the parameters). We have seen that in this circular setting, every Gaussian random fields is a sum of at most d circular ADSN, which are nothing more than a circular moving average field. This remark has two consequences: first, it is not worth distinguishing in the circular setting subclasses of Gaussian random fields (as autoregressive or moving average fields) unless one specifies a small compact support for the autoregressive or moving average filter (as will be done in Section 3.3); and second, this proves that for circular gray-level random fields, the circular ADSN model presented in [Galerne *et al.* 2011b] generalizes autoregressive and moving average random fields with a much clearer (and non-parametric) analysis scheme.

Let us finally mention that autoregressive models are especially useful when the recursion is based on the time variable. In particular, a wide class of dynamic textures can be successfully synthesized by using a Gaussian chronologically causal autoregressive model [Xia *et al.* 2014].

2.2.5 Compactly-Supported Textons

We have presented in Subsection 2.1.5 the definition of the textons associated to a circular Gaussian model [Desolneux *et al.* 2012]. Here we propose to extend this definition to the case of a stationary Gaussian random field on $\mathbb{Z}^2$.

Definition 2.2.8. A matrix texton associated to a stationary Gaussian random

field $\mathcal{N}(m, C)$ on $\mathbb{Z}^2$ is any function $t \in \ell^2(\mathbb{Z}^2, \mathbb{R}^{d\times d})$ such that

$$t * \tilde{t}^T = C \ .$$

Notice that, thanks to Lemma 2.2.2, the convolution $t * \tilde{t}$ is well defined and its Fourier transform is $\hat{t}\,\hat{t}^*$. Thus, a Gaussian random field that admits a texton necessarily has a spectral density.

Definition 2.2.9. Let us consider a stationary Gaussian random field $\mu = \mathcal{N}(m, h * \tilde{h}^T)$ with $h \in \ell^2(\mathbb{Z}^2, \mathbb{R}^d)$. A texton associated to μ is any function $t \in \ell^2(\mathbb{Z}^2, \mathbb{R}^d)$ such that $t * \tilde{t}^T = h * \tilde{h}^T$. Equivalently, in Fourier domain, a texton satisfies

$$\hat{h}\hat{h}^* = \hat{t}\hat{t}^* \quad \text{a.e.} \ .$$

Besides,

- if $d = 1$, the canonical texton of μ is the texton $t \in \ell^2$ defined in Fourier domain by
$$\hat{t} = |\hat{h}| \ .$$
- if $d = 3$, the luminance texton of μ is the texton $t \in \ell^2$ defined in Fourier domain by
$$\hat{t} = e^{-i\varphi_{\text{lum}}}\,\hat{h} \ ,$$
where $\varphi = \text{Arg}(\widehat{h_{\text{lum}}})$ and where the luminance channel of h is defined by
$$h_{\text{lum}} = 0.299h_1 + 0.587h_2 + 0.114h_3 \ .$$
Again, the choice of coefficients is motivated by [ITU 2011].

As in the circular case, if t is a texton of $\mu = \mathcal{N}(m, h * \tilde{h}^T)$ and if W is a scalar Gaussian white noise with variance 1, then $m + t * W$ follows the distribution μ. This remark is particularly interesting in the case where t has a compact support because, as we will see in Subsection 2.2.6, we thus have at our disposal a simple simulation scheme. Therefore, the question is raised to find necessary and sufficient conditions of existence of a compactly-supported texton. In the rest of this subsection, we will discuss this question in the simple case $d = 1$.

Let us first remark that in general, the canonical texton associated to a Gaussian model has not a compact support. Indeed, in the real-valued case, if one can write $|\hat{h}| = \hat{t}$ for a compactly-supported t, this implies that $|\hat{h}|$ is a $\mathscr{C}^\infty$ function, which is not always the case. For example, consider the real-valued filter $h = \delta_{(0,0)} - \delta_{(1,0)}$ whose Fourier transform is given by

$$\hat{h}(\boldsymbol{\xi}) = 1 - e^{i\xi_1} \ .$$

When $\boldsymbol{\xi} \to 0$, we have $|\hat{h}(\boldsymbol{\xi})| \sim |\xi_1|$ so that $|\hat{h}(\boldsymbol{\xi})|$ is not differentiable at 0. However, there may well be a non-canonical texton with compact support.

It is clear that a necessary condition for the existence of a compactly-supported texton $h : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$ is that the covariance $C = h * \tilde{h}^T$ has compact support. One

can wonder if this condition is also sufficient. In Fourier domain, the problem is to know if the non-negative trigonometric polynomial $\hat{C}$ admits a factorization $|\hat{h}|^2$ where $\hat{h}$ is a trigonometric polynomial. If we were dealing with functions with only one variable $x \in \mathbb{Z}$, the answer would be true: indeed, the Fejér-Riesz theorem [Daubechies 1988] for the real-valued case (which admits a matrix extension due to Rosenblum, see [Rosenblum & Rovnyak 1997] and [Dritschel 2004]) ensures that any trigonometric polynomial

$$\varphi(\xi) = \sum_{|k| \leqslant n} c(k) e^{ik\xi} \quad (\xi \in \mathbb{T})$$

such that for all $\xi \in \mathbb{T}$, $\varphi(\xi) \geqslant 0$, can be written $\varphi = |p|^2$ where p is a trigonometric polynomial

$$p(\xi) = \sum_{k=0}^{n} h(k) e^{ik\xi}$$

of same degree n. The proof relies on the fundamental theorem of algebra.

Unfortunately, the condition is not sufficient when we have more than one variable. Hilbert indeed showed that there exist non-negative trigonometric polynomials in several variables that cannot be written as sums of squared moduli of trigonometric polynomials. Nevertheless, if we consider only trigonometric polynomials that are positive (instead of non-negative), then there always exist a decomposition in a sum of $|p_i|^2$. One can refer to [Dritschel & Woerdeman 2005] or [Dumitrescu 2006] for a large discussion on this issue. Notice also that [Dritschel 2004, Theorem 5.1] indicates that there always exists a small perturbation of the non-negative trigonometric polynomial that can be factorized as a squared modulus of one trigonometric polynomial. In other words, for a compactly-supported covariance function $C : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$, there always exist an approximate compactly-supported texton t such that $t * \tilde{t}^T \approx C$.

2.2.6 Simulation

In this subsection, we explain how to sample on a rectangle a Gaussian random field that admits a compactly-supported texton.

Let $h : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$ be a kernel function with compact support S_h. We will sample the ADSN $G_h = h * W$ associated to h (where W is a white Gaussian noise with zero mean and variance 1) on a finite rectangle $\Omega \subset \mathbb{Z}^2$. For the sake of clarity, we can assume that $0 \in S_h$.

Let us remark that

$$\forall \mathbf{x} \in \Omega, \quad G_h(\mathbf{x}) = \sum_{\mathbf{y} \in S_h} h(\mathbf{y}) W(\mathbf{x} - \mathbf{y}) \ .$$

Thus, when h has a very small support, one only has to sample W on $\Omega - S_h$ and the convolution can be performed by direct summation. But when S_h gets bigger, it is well-known that the direct summation method becomes inefficient to compute the convolution, and that it may be computed more quickly using the DFT.

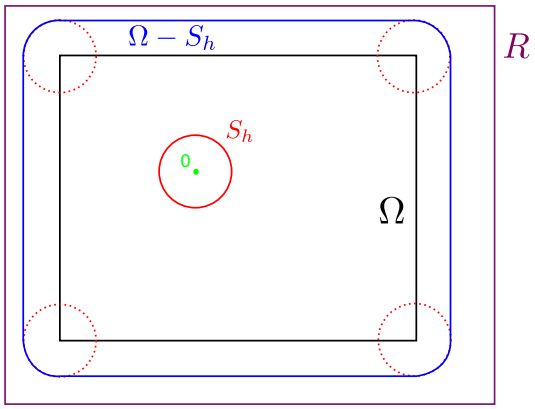


Figure 2.4: **Use of Circular Convolution.** If Ω is augmented by $-S_h$, then the restriction on Ω of the convolution by h of a function on Ω is the restriction on Ω of the circular convolution performed on a rectangle R containing $\Omega - S_h$.

This method amounts to augment Ω by the size of the kernel function, so that the restriction on Ω of the convolution by h is the same as the restriction of the circular convolution by h on the augmented domain (see Fig. 2.4). For that, let us introduce a rectangle $R \subset \mathbb{Z}^2$ of size $M \times N$ such that $\Omega - S_h \subset R$, and also the circular domain $\Theta = \mathbb{Z}/M\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$. Then Ω and S_h naturally identify to subsets Ω' and S'_h of Θ. One can thus remark that the restriction of the ADSN G_h to Ω has the same distribution as the restriction of the circular ADSN $G^{\Theta}_{h'}$ to Ω'. As we said in Subsection 2.1.2, the simulation of $G^{\Theta}_{h'}$ can be done in $\mathscr{O}(|\Theta| \log |\Theta|)$ using a DFT, and therefore, we obtain a way to simulate the restriction of G_h on Ω in $\mathscr{O}(k \log k)$ where $k = |\Omega - S_h|$. Let us also notice that this remark extends to DSN with finite intensities (the only difference is that the white noise is Poisson and not Gaussian).

To sum up, a restriction of the (asymptotic) DSN associated to a compactly-supported h can be obtained by cropping a circular (asymptotic) DSN. Of course, as one can see in Fig. 2.5 this implies that the resulting restrictions are not tileable, but the advantage is that they can naturally be extended afterwards to a wider domain. When S_h is relatively small and the intensity λ relatively low, the DSN can be simulated by direct summation. We will take profit of this remark in Section 3.3 where we will design compactly-supported textons associated to a Gaussian model such that the ADSN can be approximated by a low intensity DSN. But let us mention immediately that the direct summation method allows for very efficient on-demand synthesis: indeed, it can be parallelized using a grid-based simulation scheme for the Poisson point process explained in [Lagae *et al.* 2009] and illustrated in Fig. 2.6

2.2.7 Optimal Transport Distance

In Subsection 2.1.6, we have recalled the expression for the optimal transport distance between two circular ADSN fields given in [Xia *et al.* 2014]. Here, we show

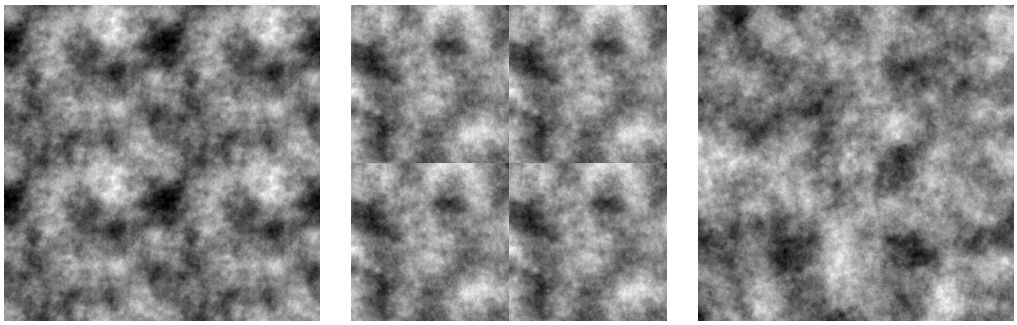


Figure 2.5: **ADSN versus circular ADSN.** From left to right, one can see a $(2,2)$-tiling of a circular ADSN of size 256×256, a $(2,2)$-tiling of a non-circular ADSN of size 256×256, and a non-circular ADSN of size 512×512. The ADSN are computed with the same kernel (indicator function of a disc of radius 30). As expected, the circular ADSN is naturally tileable. But the advantage of the non-circular ADSN model is that it can be extended to any domain without repetition (right image).

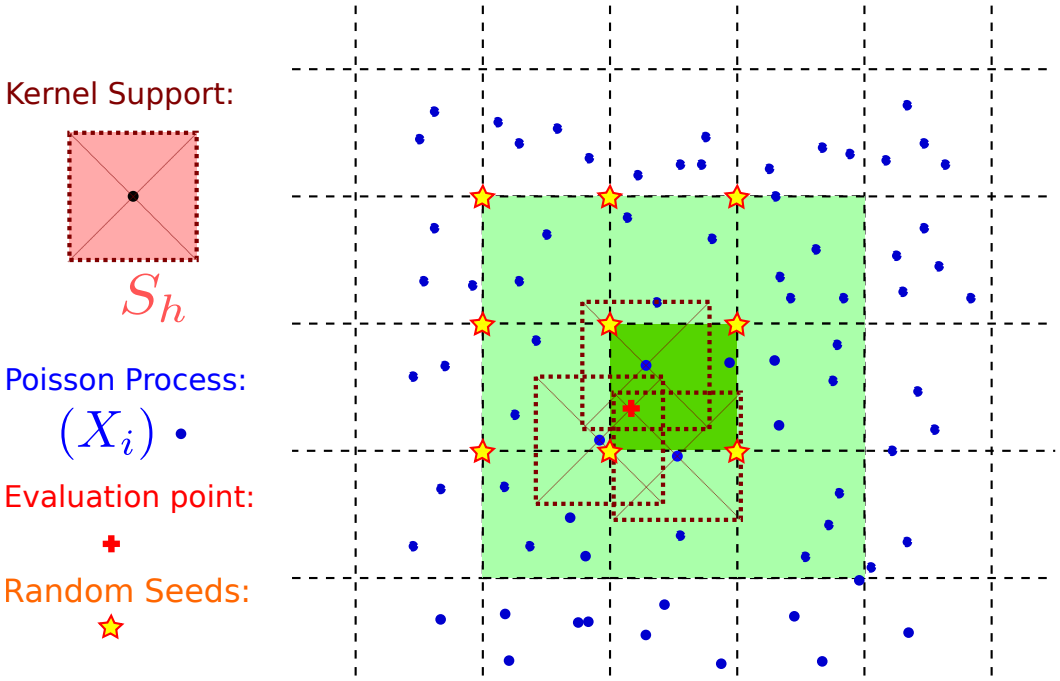


Figure 2.6: **Parallel local evaluation of a low-intensity DSN**. We consider here a DSN associated to a kernel h (with square support S_h shown in pink) and with underlying Poisson process (X_i). The synthesis domain is divided into square cells having the same size as S_h. If one wants to sample the DSN on a point of the dark green cell, then one must compute the Poisson process in the nine green cells. In each cell, the Poisson process can be sampled by using a random seed given as a function of the up-left coordinate of the cell. With this precaution, multiple evaluations of the Poisson process will be coherent. This method thus allows for parallel local evaluation of the DSN.

that it is possible to define the optimal transport distance between two **stationary** random fields on $\mathbb{Z}^2$, and we get an analog formula for the distance between two ADSN fields defined over $\mathbb{Z}^2$.

Let μ_0 and μ_1 be the probability distributions of two **stationary** random fields on $\mathbb{Z}^2$. Inspired by [Gray *et al.* 1975] and by the result of Proposition 2.1.8, we adopt the following definition.

Definition 2.2.10. The L^2-optimal transport distance between the stationary random fields μ_0 and μ_1 on $\mathbb{Z}^2$ is defined by

$$d_{\mathrm{OT}}(\mu_0,\mu_1)^2 = \inf \mathbb{E}\Big(\|F(0)-G(0)\|^2\Big)\ ,$$

where the infimum is taken over all the stationary couplings (F,G) such that $F\sim\mu_0$ and $G\sim\mu_1$

With this definition, we shall give an analog of Theorem 2.1.4 in the case of random fields on $\mathbb{Z}^2$. This result was already given in [Gray *et al.* 1975] for the case of real-valued Gaussian stationary random fields. For $d>1$, we need the following lemma.

Lemma 2.2.5. *Let $M\in\mathbb{C}^{d\times d}$ and let $a,b\in\mathbb{C}^d$. We assume that*

$$H=\begin{pmatrix} aa^* & M\\ M^* & bb^*\end{pmatrix}\in\mathcal{H}_{2d}^{\geqslant 0}\ .$$

Then we have $|\mathrm{Tr}(M)|\leqslant|a^*b|$.

Proof. The proof of this lemma relies on the following observation: if K is a non-negative Hermitian matrix, then

$$\forall i,j,\quad |K_{i,j}|^2\leqslant K_{i,i}K_{j,j}\ . \tag{2.35}$$

In the simple case $a=0$, we have $aa^*=0$, and since H is non-negative, we get from (2.35) that $M=0$ so that $|\mathrm{Tr}(M)|\leqslant|a^*b|$.

We can thus assume $a\neq 0$. In this case, there exists a unitary matrix P of size $d\times d$ whose first column is $\frac{a}{\|a\|}$. We thus have

$$P^*a=\begin{pmatrix}\|a\|\\0\\\vdots\\0\end{pmatrix}\quad\text{and}\quad P^*b=\begin{pmatrix}\alpha\\*\\\vdots\\*\end{pmatrix}\quad\text{with }\alpha=\Big(\frac{a}{\|a\|}\Big)^*b\ .$$

Let us consider

$$Q=\begin{pmatrix}P&0\\0&P\end{pmatrix}$$

which is a unitary matrix of size $2d\times 2d$. Since $H\in\mathcal{H}_{2d}^{\geqslant 0}$, we also have

$$K:=Q^*HQ=\begin{pmatrix}P^*aa^*P & P^*MP\\ P^*M^*P & P^*bb^*P\end{pmatrix}\in\mathcal{H}_{2d}^{\geqslant 0}\ .$$

Let us remark that

$$P^*aa^*P = \begin{pmatrix} \|a\|^2 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & 0 & \\ 0 & & & \end{pmatrix} \quad \text{and} \quad P^*bb^*P = \begin{pmatrix} |\alpha|^2 & * & \cdots & * \\ * & & & \\ \vdots & & * & \\ * & & & \end{pmatrix} .$$

Thanks to (2.35), we get

$$\forall i \in \{1, \dots d\}, \quad |K_{i,i+d}|^2 \leqslant K_{i,i} K_{i+d,i+d} \ ,$$

which rewrites

$$\forall i \in \{1, \dots d\}, \quad |(P^*MP)_{i,i}|^2 \leqslant (P^*aa^*P)_{i,i} (P^*bb^*P)_{i,i} \ ,$$

Besides, for $2 \leqslant i \leqslant d$, we have $(P^*aaP)_{i,i} = 0$ so that $(P^*MP)_{i,i} = 0$. And for $i = 1$, we get

$$|(P^*MP)_{1,1}|^2 \leqslant \|a\|^2 |\alpha|^2 = \|a\|^2 \left| \left(\frac{a}{\|a\|} \right)^* b \right|^2 = |a^*b|^2 \ .$$

Finally,

$$\mathrm{Tr}(M) = \mathrm{Tr}(P^*MP) = \sum_{i=1}^{d} (P^*MP)_{i,i} = (P^*MP)_{1,1}$$

and thus

$$|\mathrm{Tr}(M)| \leqslant |a^*b| \ .$$

□

Theorem 2.2.6. *Let $h_0, h_1 \in \ell^2(\mathbb{Z}^2, \mathbb{R}^d)$ be two kernel functions and let μ_0, μ_1 be the distributions of $h_0 * W$ and $h_1 * W$ where W is a white noise on $\mathbb{Z}^2$ with variance 1. We suppose that for $i = 0, 1$, μ_i has uniform random phase in the following sense:*

$$\forall g_i \in \ell^2, \quad \hat{g}_i \hat{g}_i^* = \hat{h}_i \hat{h}_i^* \implies g_i * W \overset{(d)}{=} h_i * W \ . \tag{2.36}$$

Then the L^2-optimal transport distance between μ_0 and μ_1 is given by

$$d_{OT}(\mu_0, \mu_1)^2 = \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} \left(\|\hat{h}_0(\boldsymbol{\xi})\|^2 + \|\hat{h}_1(\boldsymbol{\xi})\|^2 - 2|\hat{h}_0(\boldsymbol{\xi})^* \hat{h}_1(\boldsymbol{\xi})| \right) d\boldsymbol{\xi} \ . \tag{2.37}$$

Proof. The proof follows the same scheme as in the circular case. Let (F, G) be any stationary coupling of (μ_0, μ_1). Let us recall that F has spectral density $\varphi_F = \hat{h}_0 \hat{h}_0^*$ and G has spectral density $\varphi_G = \hat{h}_1 \hat{h}_1^*$. Therefore, as we have said in Subsection 2.2.1, F, G admit a cross-spectral density $\psi_{F,G}$. The spectral density of $(F; G)$

$$\begin{pmatrix} \varphi_F & \psi_{F,G} \\ \psi_{F,G}^* & \varphi_G \end{pmatrix} = \begin{pmatrix} \hat{h}_0 \hat{h}_0^* & \psi_{F,G} \\ \psi_{F,G}^* & \hat{h}_1 \hat{h}_1^* \end{pmatrix}$$

is $\mathcal{H}_{2d}^{\geqslant 0}$-valued so that, with Lemma 2.2.5 we get

$$\mathrm{Tr}(\psi_{F,G}) \leqslant |\hat{h}_0^* \hat{h}_1| \ .$$

By definition of the spectral densities, we thus have

$$\begin{aligned}
\mathbb{E}\Big[\|F(0)-G(0)\|^2\Big] &= \mathbb{E}\Big[\|F(0)\|^2\Big] + \mathbb{E}\Big[\|G(0)\|^2\Big] - 2\mathbb{E}\Big[G(0)^T F(0)\Big] \\
&= \mathrm{Tr}\Big(\mathbb{E}\Big[F(0)F(0)^T\Big]\Big) + \mathrm{Tr}\Big(\mathbb{E}\Big[G(0)G(0)^T\Big]\Big) - 2\mathrm{Tr}\Big(\mathbb{E}\Big[F(0)G(0)^T\Big]\Big) \\
&= \frac{1}{(2\pi)^2}\int_{\mathbb{T}^2}\Big(\mathrm{Tr}(\varphi_F(\boldsymbol{\xi})) + \mathrm{Tr}(\varphi_G(\boldsymbol{\xi})) - 2\mathrm{Tr}(\psi_{F,G}(\boldsymbol{\xi}))\Big)d\boldsymbol{\xi} \\
&\geqslant \frac{1}{(2\pi)^2}\int_{\mathbb{T}^2}\Big(\mathrm{Tr}(\hat{h}_0(\boldsymbol{\xi})\hat{h}_0(\boldsymbol{\xi})^*) + \mathrm{Tr}(\hat{h}_1(\boldsymbol{\xi})\hat{h}_1(\boldsymbol{\xi})^*) - 2|\hat{h}_0^*(\boldsymbol{\xi})\hat{h}_1(\boldsymbol{\xi})|\Big)d\boldsymbol{\xi} \\
&= \frac{1}{(2\pi)^2}\int_{\mathbb{T}^2}\Big(\|\hat{h}_0(\boldsymbol{\xi})\|^2 + \|\hat{h}_1(\boldsymbol{\xi})\|^2 - 2|\hat{h}_0(\boldsymbol{\xi})^*\hat{h}_1(\boldsymbol{\xi})|\Big)d\boldsymbol{\xi}
\end{aligned}$$

Therefore,

$$d_{\mathrm{OT}}(\mu_0,\mu_1)^2 \geqslant \frac{1}{(2\pi)^2}\int_{\mathbb{T}^2}\Big(|\hat{h}_0(\boldsymbol{\xi})|^2 + |\hat{h}_1(\boldsymbol{\xi})|^2 - 2|\hat{h}_0(\boldsymbol{\xi})^*\hat{h}_1(\boldsymbol{\xi})|\Big)d\boldsymbol{\xi}\ .$$

Now let us exhibit a coupling that achieves the lower bound. For that, we define $g_1 \in \ell^2$ by its Fourier transform

$$\hat{g}_1 = \frac{\hat{h}_1\hat{h}_1^*\hat{h}_0}{|\hat{h}_1^*\hat{h}_0|}\mathbf{1}_{\hat{h}_1^*\hat{h}_0\neq 0} + \hat{h}_1\mathbf{1}_{\hat{h}_1^*\hat{h}_0=0}\ .$$

The function $\hat{g_1}$ is measurable (because $\hat{h}_0$ and $\hat{h}_1$ are measurable), and we have $\|\hat{g}_1\| = \|\hat{h}_1\|$ so that $g_1 \in \ell^2$. We then define the coupling

$$F = h_0 * W \quad \text{and} \quad G = g_1 * W\ .$$

of μ_0, μ_1. We see that $F-G = (h_0-g_1)*W$ has spectral density $(\hat{h}_0-\hat{g}_1)(\hat{h}_0-\hat{g}_1)^*$ and

$$\mathrm{Tr}\Big((\hat{h}_0-\hat{g}_1)(\hat{h}_0-\hat{g}_1)^*\Big) = (\hat{h}_0-\hat{g}_1)^*(\hat{h}_0-\hat{g}_1) = \|\hat{h}_0\|^2 + \|\hat{h}_1\|^2 - 2|\hat{h}_0^*\hat{h}_1|\ .$$

We thus have

$$\begin{aligned}
\mathbb{E}\Big(\|F(0)-G(0)\|^2\Big) &= \frac{1}{(2\pi)^2}\int_{\mathbb{T}^2}\mathrm{Tr}\Big((\hat{h}_0-\hat{g}_1)(\hat{h}_0-\hat{g}_1)^*\Big)(\boldsymbol{\xi})d\boldsymbol{\xi} \\
&= \frac{1}{(2\pi)^2}\int_{\mathbb{T}^2}\Big(\|\hat{h}_0(\boldsymbol{\xi})\|^2 + \|\hat{h}_1(\boldsymbol{\xi})\|^2 - 2|\hat{h}_0(\boldsymbol{\xi})^*\hat{h}_1(\boldsymbol{\xi})|\Big)d\boldsymbol{\xi}\ .
\end{aligned}$$

which concludes the proof. □

Remark 2.2.4. It was conjectured in [Gray *et al.* 1975, Remark after Corollary 3] that the equality given by the last theorem is true only in the case of Gaussian processes. Actually, the result of Theorem 2.2.6 shows that this conjecture is false because the uniform random phase hypothesis (2.36) does not imply that W (or $h*W$) follows a Gaussian distribution. Indeed, let us consider two independent white Gaussian noises W_0, W_1 on $\mathbb{Z}^2$ of variances $\alpha \neq \beta$ and an independent Bernoulli variable B of parameter $\frac{1}{2}$. Then let us consider

$$W = (1-B)W_0 + BW_1\ .$$

Then W is a stationary process which is non-Gaussian because the probability density function of $W(0)$ is the average of the Gaussian densities with mean 0 and variances α and β. Besides, W is a white noise on $\mathbb{Z}^2$ because for all $\mathbf{x}, \mathbf{y} \in \mathbb{Z}^2$,

$$\begin{aligned}\mathbb{E}(W(\mathbf{x})W(\mathbf{y})) &= \mathbb{E}(W_0(\mathbf{x})W_0(\mathbf{y}))\mathbb{P}(B=0) + \mathbb{E}(W_1(\mathbf{x})W_1(\mathbf{y}))\mathbb{P}(B=1)\\ &= \Big(\frac{\alpha+\beta}{2}\Big)\,\delta_0(\mathbf{x}-\mathbf{y})\ .\end{aligned}$$

Besides, if $\hat{h}\hat{h}^* = \hat{g}\hat{g}^*$, then we have $h * W_0 \sim g * W_0$ and $h * W_1 \sim g * W_1$ (because these processes are Gaussian with the same spectral density), and therefore

$$h * W = (1-B)h * W_0 + Bh * W_1 \overset{(d)}{=} (1-B)g * W_0 + Bg * W_1 = g * W\ .$$

Remark 2.2.5. If $\hat{h}_0, \hat{h}_1$ are continuous, the integral (2.37) can be seen as the limit of the Riemann sums (2.11). In particular, if h_0 and h_1 are two compactly-supported functions on $\mathbb{Z}^2$, they define circular models μ_0^Θ μ_1^Θ on a domain $\Theta \subset \mathbb{Z}^2$ containing their support, and we thus see that when $\Theta \to \mathbb{Z}^2$,

$$d_{\mathrm{OT}}(\mu_0^\Theta, \mu_1^\Theta) \xrightarrow[\Theta\to\mathbb{Z}^2]{} d_{\mathrm{OT}}(\mu_0, \mu_1)\ .$$

Even if $\hat{h}_0$ and $\hat{h}_1$ may have large oscillations, we observe in practice that the approximation by the Riemann sums is in general quite good (with a relative error of order 1% as soon as Θ is 4 times larger than the support of the h_0 and h_1 as one can see in Fig. 2.7). This remark *a posteriori* justifies the use of the optimal transport distance on the circular models as a good approximation of the distance between models defined on $\mathbb{Z}^2$. Let us also mention that the author of [Gray *et al.* 1975] mentions another convergence result: the optimal transport distance between the stationary models μ_0, μ_1 on $\mathbb{Z}^2$ is also the limit of the optimal transport distance of the restrictions of μ_0, μ_1 to a domain Ω, when the dimensions of Ω tend to infinity.

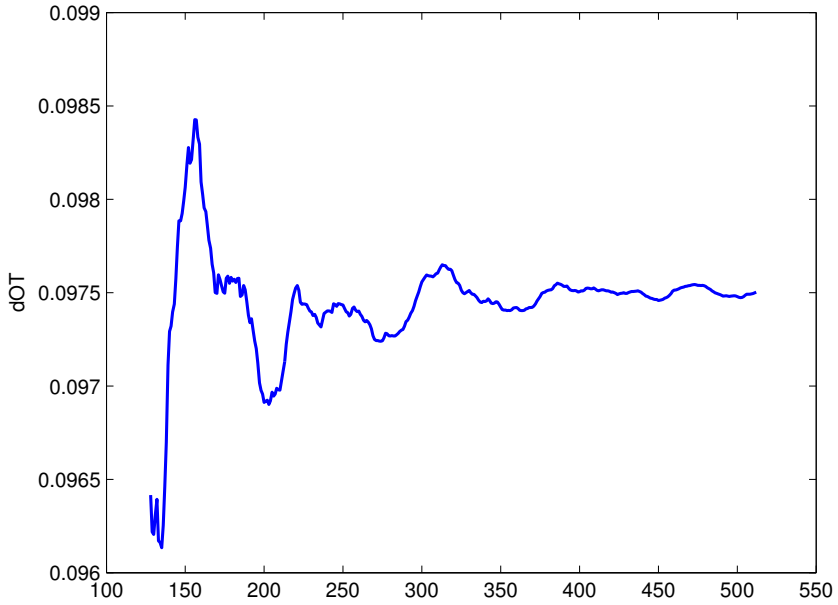


Figure 2.7: **Limit of optimal transport distances of large circular models.** To obtain this graph, we fixed two compactly-supported textons t_0, t_1 (with a support of size 128×128) associated to two natural textures (see Subsection 3.1.1) and we plotted the OTD between the circular models defined by t_0, t_1 on $(\mathbb{Z}/N\mathbb{Z})^2$ as a function of N (which is represented on the x-axis). One can see that the convergence of the OTD values to the OTD of the models on $\mathbb{Z}^2$ is quite fast: the relative error is about 1% when the dimensions of Θ are 4 times larger than the ones of the support of h_0, h_1.

Chapter 3

Random Phase Texture Synthesis by Example

Contents

In this chapter, we explain how the random phase models presented in Chapter 2 can be used to perform fast and flexible by-example texture synthesis.

As we have already said, the articles [Lewis 1984], [Lewis 1989] and [Van Wijk 1991] proposed to use random phase fields in texture synthesis. One main contribution of these articles is to demonstrate the richness and practical flexibility of random phase models; in particular, the wide possibilites offered by the interactive design of the kernel function has largely and undoubtedly contributed to the success of the spot noise model. But they did not propose a clear framework for by-example synthesis. Let us recall the following sentence of [Van Wijk 1991, §6.3 Further Work].

> *"An approach to gain more insight in the relation between the shape of the spot and the resulting texture is to attempt to derive spots from sampled real-world textures. This step is the inverse from spot to texture. A spot $h(\mathbf{x})$ has to be constructed such that its energy spectrum is the*

> *same as the power spectrum of the texture, and such that it corresponds to the notion of the spot used here, i.e. satisfies some criterion such as minimal size or minimal variance."*

By-example textures synthesis has been sucessfully addressed in many later articles; we refer the reader to [Wei *et al.* 2009] for a detailed survey on this topic. Since the steerable pyramid (and more generally a wavelet frame) is a way to extract the spectral content, the method of [Heeger & Bergen 1995] can be understood as spectrum approximation, but the distribution of the synthesized random field is not clearly formulated. As regards models that are closer to random phase fields, let us mention the articles [Ghazanfarpour & Dischler 1995], [Lagae *et al.* 2010b], and [Galerne *et al.* 2012] whose authors explain the use of noise functions for by-example texture synthesis. The methods proposed in these three articles rely on the extraction of the more significant parts of the texture power spectrum. They are limited to parametric estimation of the texture spectrum because their main objective is to perform procedural synthesis (which imposes a non-periodic, continuous, and randomly-accessible model).

If one is not particularly interested in a procedural synthesis method, then one can lead a non-parametric estimation of the texture spectrum based on one single exemplar, as was done in [Galerne *et al.* 2011b]. In the Gaussian case, this leads to a clear analysis-synthesis pipeline, and the synthesized random field has a simple distribution because it is an asymptotic discrete spot noise. This method outperforms those mentioned in the last paragraph because it preserves the whole power spectrum of the input texture. Besides, the textures that are well-reproduced by an ADSN are exactly the textures whose aspect is characterized by the power spectrum, i.e. the ones for which the Fourier phase does not convey any relevant information.

The estimation method of [Galerne *et al.* 2011b] allows to derive from an exemplar texture a spot that can be used for synthesis of this texture. But let us remark that the spot suggested in [Galerne *et al.* 2011b] does not fit the "minimal size" constraint mentioned by van Wijk. The authors of [Desolneux *et al.* 2012], [Xia *et al.* 2014] and [Desolneux *et al.* 2015] brought an elegant solution to this problem, with the notion of texton associated to a Gaussian texture model (which we already introduced in Chapter 2).

The word texton was actually introduced by Julesz in his seminal paper [Julesz 1981] to refer to the features (statistical features or geometrical features) that are involved in the perception of textures (and in particular which allow preattentive texture discrimination). In the class of stationary Gaussian textures, the second-order moments (or equivalently the Fourier spectrum) suffice to characterize a texture, hence the idea to define a texton for a Gaussian texture as a particular representative of the class of images having the same Fourier spectrum. Therefore, the authors of [Desolneux *et al.* 2012] suggest to define the canonical texton associated to a gray-level Gaussian texture by the only image with same Fourier spectrum and zero phase. Beyond the characterization of a Gaussian texture, this canonical texton has several interesting properties. For example, it is spatially con-

centrated around zero in the sense that, among the images having the same Fourier modulus, it is the unique solution of a class of optimization problems that translate the concentration at zero by a spatially-weighted ℓ^2-norm [Desolneux *et al.* 2015]. Besides, in contrast with covariance functions, it becomes natural to approximate a Gaussian model by cropping the texton[1]. This texton is thus closer to the idea of a spot associated to a texture, with "minimal size".

However, as we will see, given a prescribed compact support, it is not true that the cropped canonical texton is the kernel function that realizes the best approximation of the corresponding Gaussian models (measured by the L^2 optimal transport distance). Besides, the canonical texton is not adapted to spot noise synthesis with finite intensity: indeed, the Gaussian convergence of the corresponding spot noise is slow because the canonical texton is very spiky. The main contribution of this chapter is the introduction of the so-called Synthesis-Oriented Texton which solves both these issues. It is obtained by an algorithm of alternating projections on support and spectral constraints, and can be used for very fast Gaussian texture synthesis.

As an application of Gaussian texture synthesis, we will see at the end of this chapter that the Gaussian model can be used to address textural inpainting. Indeed, the texture inpainting problem has a clear formulation in terms of conditional simulation, and in the Gaussian case, the conditional simulation can be simply addressed with a method based on kriging estimation. We will see that this method brings an elegant and efficient solution for microtexture inpainting.

The random phase and Gaussian models used in this chapter have already been presented in Chapter 2. In Section 3.1, we present and discuss the estimation of the Gaussian model suggested in [Galerne *et al.* 2011b], and in particular we question the relevance of the periodic covariance estimator. In Section 3.2 we give examples of Gaussian texture synthesis by example, and discuss the limitations of the Gaussian model. Section 3.3 is devoted to the presentation of the Synthesis-Oriented Texton. Finally, in Section 3.4 we will present the conditional simulation for stationary Gaussian random fields and its application to the inpainting of Gaussian textures.

3.1 Model estimation

Some recent articles about Gaussian texture analysis and synthesis [Galerne *et al.* 2011b], [Desolneux *et al.* 2012], [Xia *et al.* 2014] suggest to estimate the covariance of a Gaussian model by the periodic autocorrelation of the exemplar texture. If we observe a texture $u : \Omega \longrightarrow \mathbb{R}^d$ on a discrete domain

$$\Omega = \{0, \dots M-1\} \times \{0, \dots, N-1\} \ ,$$

this amounts to set

$$\forall \mathbf{v} \in \mathbb{Z}^2, \quad c_u(\mathbf{v}) = \frac{1}{|\Omega|} \sum_{\mathbf{x} \in \Omega} (u(\mathbf{x}) - \bar{u})(\dot{u}(\mathbf{x}+\mathbf{v}) - \bar{u})^T \ , \tag{3.1}$$

[1] even if a crop of a zero-phase texton may not be a zero-phase texton.

where $\dot{u}$ refers to the (M, N)-periodic extension of u defined by

$$\forall (x, y) \in \Omega,\ \forall (k, l) \in \mathbb{Z}^2, \quad \dot{u}(x + kM, y + lN) = u(x, y)\ ,$$

and where

$$\bar{u} = \frac{1}{|\Omega|} \sum_{\mathbf{x} \in \Omega} u(\mathbf{x}) \tag{3.2}$$

refers to the empirical mean.

Because we have $\widehat{c}_u(\boldsymbol{\xi}) = \frac{1}{|\Omega|} \hat{u}(\boldsymbol{\xi}) \hat{u}^*(\boldsymbol{\xi})$ for $\boldsymbol{\xi} \neq 0$, this is equivalent (for $d = 1$) to estimate the power spectrum by the squared modulus of the DFT. In the signal processing community, it is well-known that the squared modulus of the DFT is a very noisy information and that it may not be the best estimator of the power spectrum.

When estimating the covariance of circular stationary random fields, the use of the periodic autocorrelation is natural: amongst other reasons, the relation (2.2) indicates that the estimator is asymptotically unbiased. Another reason, that was mentioned in [Xia *et al.* 2014] and that we will prove in Subsection 3.1.1, is that this estimator is indeed the one that maximizes the likelihood function.

But in practical cases of texture synthesis by example, there is no reason for the exemplar texture to be periodic, so that the periodic autocorrelation is not legitimate *a priori*. Actually, this is not natural to model a non-periodic texture by a circular stationary random field. The author of [Moisan 2011] proposed an elegant solution to extract from any image a "periodic component" whose geometry is very similar in the domain interior and which has a very attenuated border-to-border discontinuity. As we will see in Subsection 3.1.2, this periodic component makes the circular modelling more legitimate but does not entirely solve the covariance estimation.

Another approach is to use a purely non-periodic estimator of the covariance given by

$$\forall \mathbf{v} \in \mathbb{Z}^2, \quad \gamma_u(\mathbf{v}) = \frac{1}{|\Omega \cap (\Omega - \mathbf{v})|} \sum_{\mathbf{x} \in \Omega \cap (\Omega - \mathbf{v})} (u(\mathbf{x}) - \bar{u})(u(\mathbf{x} + \mathbf{v}) - \bar{u})^T\ , \tag{3.3}$$

with the convention that $\gamma_u(\mathbf{v}) = 0$ as soon as $\Omega \cap (\Omega - \mathbf{v}) = \varnothing$, and where $\bar{u}$ still refers to the empirical mean. A variant is given by

$$\forall \mathbf{v} \in \mathbb{Z}^2, \quad \gamma'_u(\mathbf{v}) = \frac{1}{|\Omega|} \sum_{\mathbf{x} \in \Omega \cap (\Omega - \mathbf{v})} (u(\mathbf{x}) - \bar{u})(u(\mathbf{x} + \mathbf{v}) - \bar{u})^T\ . \tag{3.4}$$

Such non-periodic estimators will be studied in Subsection 3.1.3. On synthetic cases, we will see that, contrary to what we hoped, they do not perform better than the periodic autocorrelation. In turn, our discussion also explains *a posteriori* the success of the traditionnaly used periodic estimator.

For the sake of simplicity, the discussion of this section will be restricted to the gray-level case, that is, $d = 1$.

3.1.1 Maximum Likelihood Estimation in the Circular case

We show in this subsection that, in the gray-level case, the traditional periodic estimator c_u defined by (3.1) can be seen as a maximum likelihood estimator.

Let us consider a stationary Gaussian random field $U : \Theta \longrightarrow \mathbb{R}$ defined on a circular domain

$$\Theta = \mathbb{Z}/M\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$$

with mean $m \in \mathbb{R}^d$ and covariance function $C : \Theta \longrightarrow \mathbb{R}$.

As we have said in Subsection 2.1.1, the covariance

$$\begin{array}{rccc} \Gamma : & \Theta \times \Theta & \longrightarrow & \mathbb{R} \\ & (\mathbf{x}, \mathbf{y}) & \longmapsto & C(\mathbf{x} - \mathbf{y}) \end{array}$$

induces a bilinear non-negative form on $\mathbb{R}^\Theta$ and thus a non-negative linear operator from $\mathbb{R}^\Theta$ to $\mathbb{R}^\Theta$, that we will still denote by Γ. It is well known that the Gaussian vector U has a probability density function if and only if the operator Γ is invertible, and in that case, the probability density function writes

$$u \in \mathbb{R}^\Theta \longmapsto \frac{1}{\sqrt{\det(2\pi\Gamma)}} \exp\left(-\frac{1}{2} \sum_{\mathbf{x},\mathbf{y} \in \Theta} \Gamma^{-1}(\mathbf{x}, \mathbf{y})(u(\mathbf{x}) - m)(u(\mathbf{y}) - m)\right) .$$

As we have seen on Equation (2.3), we have for all $f : \Theta \longrightarrow \mathbb{R}$,

$$\sum_{\mathbf{x},\mathbf{y}} f(\mathbf{x}) f(\mathbf{y}) \Gamma(\mathbf{x}, \mathbf{y}) = \frac{1}{|\Theta|} \sum_{\boldsymbol{\xi}} |\hat{f}(-\boldsymbol{\xi})|^2 \hat{C}(\boldsymbol{\xi}) = \frac{1}{|\Theta|} \sum_{\boldsymbol{\xi}} |\hat{f}(\boldsymbol{\xi})|^2 \hat{C}(\boldsymbol{\xi}) ,$$

where $\hat{f}$ and $\hat{C}$ are the DFTs of f and C. This shows that the operator Γ has an eigenvector basis given by the (unitary) discrete Fourier basis and the corresponding eigenvalues are $\hat{C}(\boldsymbol{\xi})$, $\boldsymbol{\xi} \in \Theta$. Thus, the operator Γ is invertible if and only if $\hat{C}(\boldsymbol{\xi}) \neq 0$ for all $\boldsymbol{\xi}$, and

$$\sum_{\mathbf{x},\mathbf{y}} f(\mathbf{x}) f(\mathbf{y}) \Gamma^{-1}(\mathbf{x}, \mathbf{y}) = \frac{1}{|\Theta|} \sum_{\boldsymbol{\xi}} \frac{|\hat{f}(\boldsymbol{\xi})|^2}{\hat{C}(\boldsymbol{\xi})} .$$

Therefore, the Gaussian random field U has a probability density function if and only if $\hat{C}(\boldsymbol{\xi}) \neq 0$, and in this case, the density function writes

$$u \in \mathbb{R}^\Theta \longmapsto \frac{1}{\sqrt{2\pi|\Theta|\hat{C}(0)}} \exp\left(-\frac{(\hat{u}(0) - m|\Theta|)^2}{2|\Theta|\hat{C}(0)}\right) \prod_{\boldsymbol{\xi} \neq 0} \frac{1}{\sqrt{2\pi|\Theta|\hat{C}(\boldsymbol{\xi})}} \exp\left(-\frac{|\hat{u}(\boldsymbol{\xi})|^2}{2|\Theta|\hat{C}(\boldsymbol{\xi})}\right) .$$

More generally, if $\hat{C}$ takes the value 0, then U admits a probability density function on the subspace

$$\{\ u \in \mathbb{R}^\theta \mid \hat{u}(\boldsymbol{\xi}) = 0 \text{ as soon as } \hat{C}(\boldsymbol{\xi}) = 0\ \} ,$$

and the density function on that subset is the same than before except that the product does not contain the frequencies $\boldsymbol{\xi}$ for which $\hat{C}(\boldsymbol{\xi}) = 0$.

Suppose that we oberve a texture $u : \Theta \longrightarrow \mathbb{R}$ that we want to model by a Gaussian random field $U \sim \mathcal{N}(m, C)$. Then the likelihood function is

$$L(m,C) = \frac{1}{\sqrt{2\pi|\Theta|\hat{C}(0)}} \exp\left(-\frac{(\hat{u}(0) - m|\Theta|)^2}{2|\Theta|\hat{C}(0)}\right) \prod_{\boldsymbol{\xi}\in\Theta\setminus\{0\}} \frac{1}{\sqrt{2\pi|\Theta|\hat{C}(\boldsymbol{\xi})}} \exp\left(-\frac{|\hat{u}(\boldsymbol{\xi})|^2}{2|\Theta|\hat{C}(\boldsymbol{\xi})}\right) .$$

Thus, we have

$$-\log L(m,C) = \text{cst} + \frac{1}{2}\left(\log(\hat{C}(0)) + \frac{(\hat{u}(0) - m|\Theta|)^2}{|\Theta|\hat{C}(0)} + \sum_{\boldsymbol{\xi}\neq 0} \log(\hat{C}(\boldsymbol{\xi})) + \frac{|\hat{u}(\boldsymbol{\xi})|^2}{|\Theta|\hat{C}(\boldsymbol{\xi})}\right) .$$

We want now to compute the maximum likelihood (ML) estimator of (m, C), which means that we want to maximize $L(m, C)$ (or equivalently minimize $-\log L(m, C)$) on the set of couples (m, C) where $m \in \mathbb{R}$ and where C is a circular covariance function on Θ. In the first place, we would like to restrict $L(m, C)$ to non-degenerate covariance functions C (i.e. such that $\hat{C}(\boldsymbol{\xi}) > 0$ for every $\boldsymbol{\xi}$), but then, as soon as there exists $\boldsymbol{\xi}_0$ such that $\hat{u}(\boldsymbol{\xi}_0) = 0$, $L(m, C)$ does not have a finite minimum because

$$\lim_{\hat{C}(\boldsymbol{\xi}_0)\to 0} \log(\hat{C}(\boldsymbol{\xi}_0)) = -\infty .$$

It is thus natural to restrict to covariance functions which satisfy $\hat{C}(\boldsymbol{\xi}) = 0$ as soon as $\hat{u}(\boldsymbol{\xi}) = 0$.

We have a similar problem for $\hat{C}(0)$. Indeed, for a fixed C, $m \mapsto -\log L(m, C)$ always has a unique minimum at

$$\bar{u} = \frac{1}{|\Theta|}\hat{u}(0) = \frac{1}{|\Theta|}\sum_{\mathbf{x}\in\Theta} u(\mathbf{x}) ,$$

and then, the function $C \mapsto -\log(\bar{u}, C)$ does not have a finite minimum (for the same reason as above).

All in all, we will always estimate the mean m by the empirical mean $\bar{u}$, and then we will minimize the function

$$C \longmapsto L(C) = \sum_{\substack{\boldsymbol{\xi}\neq 0 \\ \hat{u}(\boldsymbol{\xi})\neq 0}} \log(\hat{C}(\boldsymbol{\xi})) + \frac{|\hat{u}(\boldsymbol{\xi})|^2}{|\Theta|\hat{C}(\boldsymbol{\xi})}$$

on the set

$$\mathcal{C}_u = \left\{ C : \Theta \longrightarrow \mathbb{R} \;\middle|\; \forall \boldsymbol{\xi},\ \hat{C}(\boldsymbol{\xi}) \geqslant 0\ ,\ \hat{C}(0) = 0\ ,\ \text{and}\ \hat{C}(\boldsymbol{\xi}) = 0\ \text{as soon as}\ \hat{u}(\boldsymbol{\xi}) = 0 \right\} .$$

For that, we noticed that the different terms of the sum can be minimized independently, modulo the fact that the summand is an even function of $\boldsymbol{\xi}$. Besides, one can notice that for a fixed $k > 0$, the function

$$\begin{array}{ccc}]0, +\infty[& \longrightarrow & \mathbb{R} \\ c & \longmapsto & \log(c) + \frac{k}{c} \end{array}$$

has a minimum value which is reached at the point $c = k$. Therefore, L admits a unique minimum on $\mathcal{C}_u$ reached on the covariance function C which satisfies

$$\forall \boldsymbol{\xi} \neq 0, \quad \hat{C}(\boldsymbol{\xi}) = \frac{1}{|\Theta|}|\hat{u}(\boldsymbol{\xi})|^2 \ .$$

Finally, we get the following

Proposition 3.1.1. *If we observe a gray-level texture $u : \Theta \longrightarrow \mathbb{R}$, then the maximum likelihood estimator (in the sense explained above) for the circular Gaussian model $\mathcal{N}(m, C)$ is given by*

$$m_{ML} = \bar{u} = \frac{1}{|\Theta|} \sum_{\mathbf{x} \in \Theta} u(\mathbf{x}) \ , \quad \text{and} \quad C_{ML} = c_u = \frac{1}{|\Theta|}(u - \bar{u}) * (\tilde{u} - \bar{u}) \ ,$$

where $*$ refers to the circular convolution on Θ.

Remark 3.1.1. Let us make three more comments about the ML estimator.

- If $U \sim \mathcal{N}(m, C)$ then $\hat{U}(\boldsymbol{\xi}) \neq 0$ almost surely as soon as $\hat{C}(\boldsymbol{\xi}) \neq 0$. For that reason, it is natural to enforce $\hat{C}(\boldsymbol{\xi}) = 0$ for the estimation of C as soon as we observe $\hat{u}(\boldsymbol{\xi}) = 0$.

- It is quite uncommon to estimate a random model from one single realization. However, we have been able to derive a ML estimator thanks to the stationarity assumption, modulo some constraints on m and C. Notice in particular that the choice $\hat{C}(0) = 0$ is quite arbitrary: it means that the realizations of $\mathcal{N}(m, C)$ have a mean value which is almost surely equal to m. If we had more than one observation, we could compute a ML estimator without imposing $\hat{C}(0) = 0$.

- Let us remark that there exists a generic ML estimation method for a Gaussian random vector. But it is different from the one presented here: indeed, the above ML is restricted to a *circular covariance function* (i.e. the covariance function of a circular stationary random field). In particular, the stationarity assumption modifies the degeneracy condition (because the space of circular covariance functions is much smaller than the space of all covariance matrices).

- The case of vector-valued random fields (i.e. $d > 1$) is a bit more difficult because for each $\boldsymbol{\xi} \in \Theta$, the matrix $\hat{u}(\boldsymbol{\xi})\hat{u}(\boldsymbol{\xi})^*$ has rank one and is thus always degenerate.

3.1.2 Coping with the periodicity assumption

In this subsection, we discuss the practical legitimacy of the periodic covariance estimator c_u defined by (3.1).

The ML estimation that we presented in Subsection 3.1.1 allows to learn a circular stationary Gaussian model from one single realization. But in general, the natural textures that we can observe have no reason to be plausible in a circular

stationary Gaussian model (because they have no reason to be tileable). Therefore, if we observe a natural texture $u : \Omega \longrightarrow \mathbb{R}$ defined on a discrete rectangle $\Omega \subset \mathbb{Z}^2$ (which can be identified to a circular domain $\Theta = \mathbb{Z}/M\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$), and if we consider the covariance ML estimator c_u, then the terms in (3.1) corresponding to positions $\mathbf{x}$ such that $\mathbf{x} \in \Omega$ and $\mathbf{x}+\mathbf{v} \notin \Omega$ cannot be considered as valid samples for the covariance estimation. One can thus wonder if these samples induce a strong bias of the periodic covariance estimator c_u. In Subsection 3.1.3, we will analyze the precision of this estimator and compare it to the precision of another purely non-periodic estimator.

But before that, let us mention a technique introduced in [Moisan 2011] that allows to reduce the border-to-border discontinuity of a natural image without affecting its inner geometry. If $u : \Omega \longrightarrow \mathbb{R}$ is an image defined on

$$\Omega = \{0, \dots M-1\} \times \{0, \dots, N-1\} \ ,$$

then the periodic plus smooth decomposition of u is by definition the unique couple (p, s) of images defined on Ω which minimizes

$$\sum_{\substack{\mathbf{x} \in \Omega,\, \mathbf{y} \in \mathbb{Z}^2 \setminus \Omega \\ |\mathbf{x}-\mathbf{y}|=1}} (p(\mathbf{x}) - \dot{p}(\mathbf{y}))^2 \ + \sum_{\substack{\mathbf{x},\mathbf{y} \in \Omega \\ |\mathbf{x}-\mathbf{y}|=1}} (s(\mathbf{x}) - s(\mathbf{y}))^2$$

(where $|\mathbf{x}|$ is the l^2-norm on $\mathbb{Z}^2$) under the constraints

$$u = p + s \quad \text{and} \quad \sum_{\mathbf{x} \in \Omega} s(\mathbf{x}) = 0 \ .$$

The image p is called the **periodic component** of u, and is also written $\mathrm{per}(u)$. Equivalently, one can express p as a solution of a Poisson equation. Indeed, let us define the (periodic) Laplacian operator $\Delta : \mathbb{R}^\Omega \longrightarrow \mathbb{R}^\Omega$ by

$$\forall \mathbf{x} \in \Omega, \quad \Delta(u)(\mathbf{x}) = -4u(\mathbf{x}) + \sum_{\substack{\mathbf{y} \in \mathbb{Z}^2 \\ |\mathbf{x}-\mathbf{y}|=1}} \dot{u}(\mathbf{y}) \ ,$$

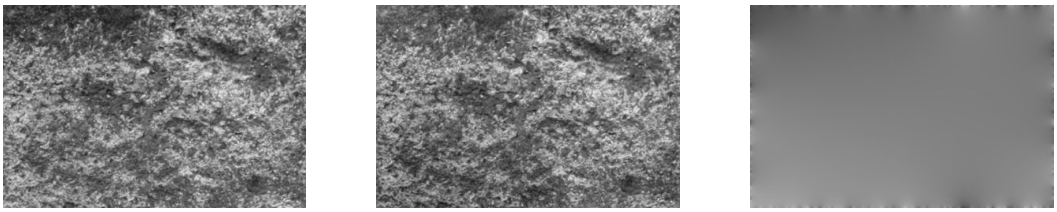
and let us capture the border-to-border discontinuities of u in the image $v = v_1 + v_2$ where for all $(x, y) \in \Omega$ we set

$$v_1(x,y) = \Big(u(M-1-x, y) - u(x,y)\Big)\mathbf{1}_{x=0 \text{ or } x=M-1} \ ,$$
$$v_2(x,y) = \Big(u(x, N-1-y) - u(x,y)\Big)\mathbf{1}_{y=0 \text{ or } y=N-1} \ .$$

Then, one can show that the periodic component p satisfies

$$\Delta p = \Delta u - v \ .$$

Since the Laplacian operator is circulant with respect to the periodic translations of Ω and invertible on the sets of images with zero sum, this equation can be solved in Fourier domain, and thus the periodic component can be obtained in $\mathscr{O}(|\Omega| \log |\Omega|)$ using the fast Fourier transform.



u = p + s

Figure 3.1: **Example of periodic plus smooth decomposition.** Notice that the in the interior of the image domain, the perceived geometrical features are almost the same in u and in p. This experience confirms that s is a smooth image (and actually, s is smoother in the domain interior than near the border).

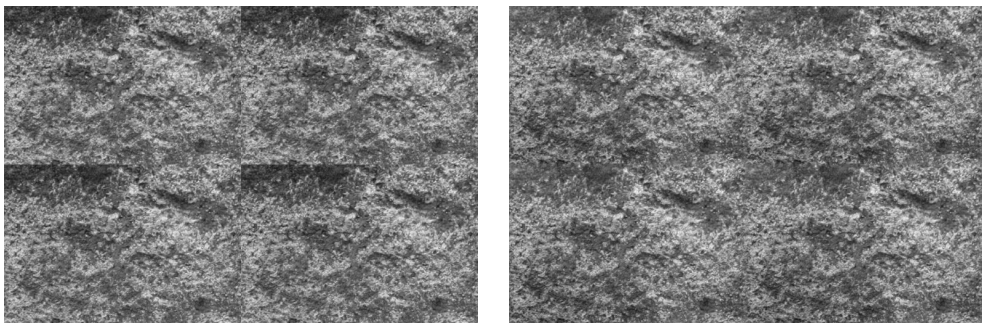


Figure 3.2: **Texture tiling using the periodic component.** On the left, one can see the 2×2 tiling of the texture shown in Fig. 3.1, and on the right, the tiling of its periodic component. On the tiling of the raw texture, the image borders are clearly visible between the tiles, whereas they are more difficult to perceive on the tiling of the periodic component.

As one can see in Fig. 3.1 and Fig. 3.2, taking the periodic component reduces the intensity gaps between two opposite image borders. In several cases (as in Fig. 3.2), this suffices to make the texture tileable. However, let us remark that, even if per can reduce the border to border intensity gaps, nothing ensures that the correlations between values taken on both sides of the border of a tile are compatible with the initial texture autocorrelation (in the interior of the domain). Therefore, as concerns the covariance estimation, one can imagine that the periodic component will reduce the bias induced by the border to border discontinuity, but it does not make the periodic covariance estimator completely legitimate for a non-periodic exemplar texture, as illustrated by the degenerate case given in Fig. 3.3. From these observations, we can draw a simple rule: one can consider that taking the periodic component overcomes the non-periodicity of the exemplar texture as soon as there is no easily-perceptible discontinuity in the textural content in the 2×2 tiling of the exemplar.

Remark 3.1.2. The operator per can be extended to color images by computing the periodic component seperately on each channel.

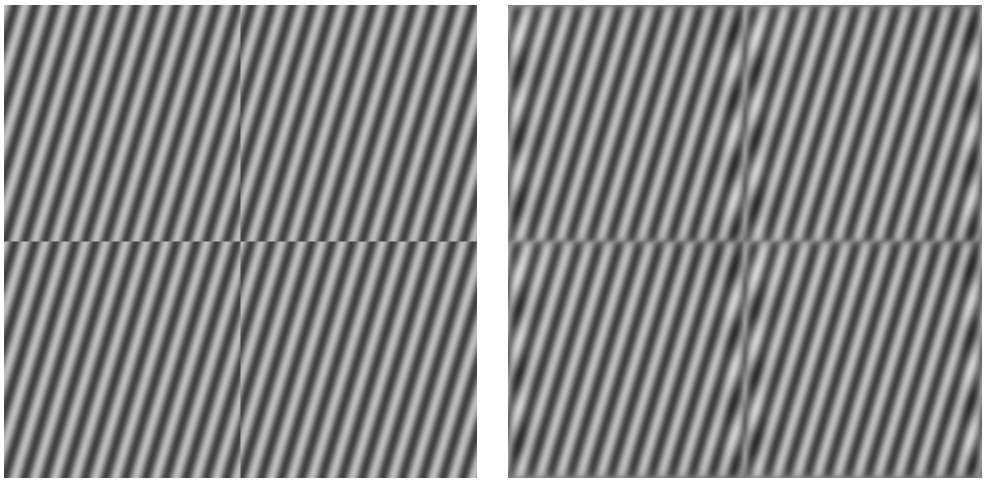


Figure 3.3: **Failure case of texture tiling using the periodic component.** On the left, one can see the 2×2 tiling of a pure wave texture, and on the right, the tiling of its periodic component. In this case, the operator per indeed reduces the intensity gaps, but the borders of the tiles are still clearly visible in the tiling.

3.1.3 Asymptotic performance

Suppose now that we observe $u : \Omega \longrightarrow \mathbb{R}$ and that we want to model u by the restriction to Ω of a stationary Gaussian field $U : \mathbb{Z}^2 \longrightarrow \mathbb{R}$ with mean m and covariance function C. Since we cannot write the density of the random field U (because it has infinitely many components), we cannot follow the maximum likelihood procedure to derive a covariance estimator. Nevertheless, we can still consider the natural (non-periodic) estimator defined by (3.3) and the periodic estimator given by (3.1).

In this setting, we are allowed to study the convergence of the covariance estimators when the size of the observation domain Ω grows to infinity. We will see that, when we restrict to Gaussian models with compactly-supported covariance functions, both the periodic and non-periodic estimators are asymptotically pointwise consistent. After that, we will illustrate on a simple example that the performance of these estimators are comparable.

A first step is to prove the consistency of the mean estimator.

Lemma 3.1.1. *Let $U : \mathbb{Z}^2 \longrightarrow \mathbb{R}$ be a stationary Gaussian field with mean m and covariance function C. Let us assume that C has compact support K. Then the empirical mean (3.2) is a consistent estimator of m, i.e. when the dimensions M and N of the rectangle Ω tend to ∞, we have almost surely*

$$\overline{U_\Omega} \longrightarrow m \,, \tag{3.5}$$

where U_Ω refers to the Ω-restriction of U.

Proof. We will use an ergodicity argument that follows from the compactness of the covariance support. For that, we introduce a rectangle R of size $r \times s$ containing K.

We then consider the partition of Ω defined by the sets

$$\Omega_i = \Omega \cap (i \ + r\mathbb{Z} \times s\mathbb{Z}) \ , \quad (i \in R) \ .$$

Then,

$$\overline{U_\Omega} = \frac{1}{|\Omega|} \sum_{i \in R} \sum_{\mathbf{x} \in \Omega_i} U(\mathbf{x}) = \sum_{i \in R} \frac{|\Omega_i|}{|\Omega|} \cdot \frac{1}{|\Omega_i|} \sum_{\mathbf{x} \in \Omega_i} U(\mathbf{x}) \ .$$

Observe that for each $i \in R$, the random variables $(U(\mathbf{x}))_{\mathbf{x} \in \Omega_i}$ are independent since they are Gaussian with covariance 0, and also identically distributed because U is stationary. Therefore the law of large numbers ensures that when $|\Omega| \longrightarrow \infty$, we have almost surely

$$\forall i \in R, \quad \frac{1}{|\Omega_i|} \sum_{\mathbf{x} \in \Omega_i} U(\mathbf{x}) \longrightarrow m \ ,$$

which entails $\overline{U_\Omega} \longrightarrow m$, because

$$\sum_{i \in R} \frac{|\Omega_i|}{|\Omega|} = 1 \ .$$

□

Based on this result, we can show the consistency of the non-periodic and periodic covariance estimators.

Theorem 3.1.2. *Let $U : \mathbb{Z}^2 \longrightarrow \mathbb{R}$ be a stationary Gaussian field with mean m and covariance function C. Let us assume that C has compact support K. Then the covariance estimators defined by (3.1), (3.3) and (3.4) are pointwise consistent, i.e. when the dimensions M and N of the rectangle Ω tend to ∞, we have almost surely for all $\mathbf{v} \in \mathbb{Z}^2$,*

$$\begin{aligned} \gamma_{U_\Omega}(\mathbf{v}) &\longrightarrow C(\mathbf{v}) \ , \\ \gamma'_{U_\Omega}(\mathbf{v}) &\longrightarrow C(\mathbf{v}) \ , \\ c_{U_\Omega}(\mathbf{v}) &\longrightarrow C(\mathbf{v}) \ , \end{aligned}$$

where U_Ω refers to the Ω-restriction of U.

Proof. We will first prove the convergence for the non-periodic estimators, by employing the same method than in the proof of Lemma 3.1.1. Let $\mathbf{v} \in K$. Expanding (3.3), one gets

$$\begin{aligned} \gamma_{U_\Omega}(\mathbf{v}) = \frac{1}{|\Omega \cap (\Omega - \mathbf{v})|} \sum_{\mathbf{x} \in \Omega \cap (\Omega - \mathbf{v})} U(\mathbf{x}) U(\mathbf{x} + \mathbf{v}) - \overline{U_\Omega} \left(\frac{1}{|\Omega \cap (\Omega - \mathbf{v})|} \sum_{\mathbf{x} \in \Omega \cap (\Omega - \mathbf{v})} U(\mathbf{x}) \right) \\ - \overline{U_\Omega} \left(\frac{1}{|\Omega \cap (\Omega - \mathbf{v})|} \sum_{\mathbf{x} \in \Omega \cap (\Omega - \mathbf{v})} U(\mathbf{x} + \mathbf{v}) \right) + \overline{U_\Omega}^2 \ . \end{aligned}$$

Splitting the sum into sums of independent terms as in the proof of Lemma 3.1.1, we obtain that the sum of the three last terms converges almost surely to

$$-m^2 - m^2 + m^2 = -m^2 \ .$$

The first term is the empirical mean on $\Omega \cap (\Omega - \mathbf{v})$ of the process

$$\mathbf{x} \longmapsto U(\mathbf{x})U(\mathbf{x}+\mathbf{v}) \ .$$

Notice that this process also has a compactly-supported covariance function. More precisely, the samples at pixels separated at least by the size of $K - K + \{-\mathbf{v}, 0, \mathbf{v}\}$ are independent. Therefore, imitating again the proof of Lemma 3.1.1, one can show that when M and N tend to ∞, we have almost surely

$$\frac{1}{|\Omega \cap (\Omega - \mathbf{v})|} \sum_{\mathbf{x} \in \Omega \cap (\Omega - \mathbf{v})} U(\mathbf{x})U(\mathbf{x}+\mathbf{v}) \longrightarrow \mathbb{E}[U(\mathbf{x})U(\mathbf{x}+\mathbf{v})] \ ,$$

so that $\gamma_{U_\Omega}(\mathbf{v}) \longrightarrow C(\mathbf{v})$ almost surely.

The convergence also holds for the renormalized γ'_{U_Ω}. Indeed, we have for each $\mathbf{v}$,

$$\frac{|\Omega \cap (\Omega - \mathbf{v})|}{|\Omega|} \longrightarrow 1$$

when M and N tend to ∞, because of the relation

$$\frac{|\Omega \backslash (\Omega - \mathbf{v})|}{MN} = \frac{|v_1|}{M} + \frac{|v_2|}{N} - \frac{|v_1 v_2|}{MN} \ . \tag{3.6}$$

Now, let us prove the consistency of the periodic estimator. With the first part of the proof, it is enough to show that the difference between γ'_{U_Ω} and c_{U_Ω} tends to zero almost surely. Notice that

$$c_{U_\Omega}(\mathbf{v}) = \gamma'_{U_\Omega}(\mathbf{v}) + \frac{1}{|\Omega|} \sum_{\mathbf{x} \in \Omega \backslash (\Omega - \mathbf{v})} (U(\mathbf{x}) - \overline{U_\Omega})(\dot{U}(\mathbf{x}+\mathbf{v}) - \overline{U_\Omega}) \ . \tag{3.7}$$

The last sum can be rewritten

$$\sum_{\mathbf{x} \in \Omega \backslash (\Omega - \mathbf{v})} U(\mathbf{x})\dot{U}(\mathbf{x}+\mathbf{v}) - \overline{U_\Omega} \left(\sum_{\mathbf{x} \in \Omega \backslash (\Omega - \mathbf{v})} U(\mathbf{x}) \right) - \overline{U_\Omega} \left(\sum_{\mathbf{x} \in \Omega \backslash (\Omega - \mathbf{v})} \dot{U}(\mathbf{x}+\mathbf{v}) \right)$$
$$+ |\Omega \backslash (\Omega - \mathbf{v})| \overline{U_\Omega}^2 \ . \tag{3.8}$$

Notice that

$$|\Omega \backslash (\Omega - \mathbf{v})| = |v_1|N + |v_2|M - |v_1||v_2|$$

so that $|\Omega \backslash (\Omega - \mathbf{v})| \to \infty$ when $|\Omega| \to \infty$. Thanks to the consistency of the empirical mean, we get that

$$-\overline{U_\Omega} \, \frac{\sum_{\mathbf{x} \in \Omega \backslash (\Omega - \mathbf{v})} U(\mathbf{x})}{|\Omega \backslash (\Omega - \mathbf{v})|} - \overline{U_\Omega} \, \frac{\sum_{\mathbf{x} \in \Omega \backslash (\Omega - \mathbf{v})} \dot{U}(\mathbf{x}+\mathbf{v})}{|\Omega \backslash (\Omega - \mathbf{v})|} + \overline{U_\Omega}^2$$

converges almost surely to $-m^2$. And for the first sum of (3.8), we will proceed as in the proof of Lemma 3.1.1. Notice that as soon as Ω is large enough, $U(\mathbf{x})$ and $\dot{U}(\mathbf{x}+\mathbf{v})$ are independent. Indeed, suppose for example that $x_1 + v_1 \notin \Omega_1$ and $x_2 + v_2 \in \Omega_2$, then

$$\dot{U}(\mathbf{x}+\mathbf{v}) = U(x_1 + v_1 - M, x_2 + v_2) \ ,$$

so that the covariance between $U(\mathbf{x})$ and $\dot{U}(\mathbf{x}+\mathbf{v})$ is $C(v_1 - M, v_2)$ which vanishes as soon as M is large enough. Therefore, the sum $\sum_{\mathbf{x}\in\Omega\setminus(\Omega-\mathbf{v})} U(\mathbf{x})\dot{U}(\mathbf{x}+\mathbf{v})$ can be split into several sums of i.i.d. terms, and thus the law of large numbers ensures that when M and N tend to ∞, we have almost surely

$$\frac{1}{|\Omega\setminus(\Omega-\mathbf{v})|}\left(\sum_{\mathbf{x}\in\Omega\setminus(\Omega-\mathbf{v})} U(\mathbf{x})\dot{U}(\mathbf{x}+\mathbf{v})\right) \longrightarrow m^2 \ .$$

Grouping the two parts, we have shown that when M and N tend to ∞,

$$\frac{1}{|\Omega\setminus(\Omega-\mathbf{v})|}\sum_{\mathbf{x}\in\Omega\setminus(\Omega-\mathbf{v})} (U(\mathbf{x}) - \overline{U_\Omega})(\dot{U}(\mathbf{x}+\mathbf{v}) - \overline{U_\Omega}) \xrightarrow{a.s.} 0 \ .$$

Using (3.7) and (3.6), it follows that

$$c_{U_\Omega}(\mathbf{v}) - \gamma'_{U_\Omega}(\mathbf{v}) = o\left(\frac{|\Omega\setminus(\Omega-\mathbf{v})|}{MN}\right) = o\left(\frac{1}{M}+\frac{1}{N}\right) \ . \tag{3.9}$$

Using the consistency result for the non-periodic estimator we get the consistency of the periodic estimator. Notice that we also showed

$$c_{U_\Omega}(\mathbf{v}) - \gamma_{U_\Omega}(\mathbf{v}) = -\frac{|\Omega\setminus(\Omega-\mathbf{v})|}{MN}\gamma_{U_\Omega}(\mathbf{v}) + o\left(\frac{|\Omega\setminus(\Omega-\mathbf{v})|}{MN}\right) \ ,$$

so that when M and N tend to ∞, we have almost surely

$$c_{U_\Omega}(\mathbf{v}) - \gamma_{U_\Omega}(\mathbf{v}) = \mathcal{O}\left(\frac{1}{M}+\frac{1}{N}\right) \ . \tag{3.10}$$

□

Let us add a comment about the last result. Not only have we shown the consistency of the estimators, but we have given the upper bounds (3.9), (3.10) on the difference between the periodic and non-periodic estimators. Unfortunately, this upper bound is not small enough to conclude that both estimators perform equally. Indeed, even if a central limit theorem cannot be applied immediately (because of the dependencies between the terms of the sums defining c_u or γ_u), it could be expected that the estimation errors

$$\gamma_{U_\Omega}(\mathbf{v}) - C(\mathbf{v}) \quad , \quad \gamma'_{U_\Omega}(\mathbf{v}) - C(\mathbf{v}) \quad , \quad c_{U_\Omega}(\mathbf{v}) - C(\mathbf{v})$$

are of order $\frac{1}{\sqrt{MN}}$. Therefore, thanks to the last theorem, we see that for example, if M and N tend to infinity with the constraint $M = N$, then the upper bound $\frac{1}{M} + \frac{1}{N}$ has the same magnitude than the estimation error.

Let us now analyze in a practical asymptotic framework the performance of these estimators. We performed numerical simulations with a Gaussian random field $U = k * W$ where k is a discrete disc of radius r and W a normalized Gaussian white noise on $\mathbb{Z}^2$. We draw a realization u of U on a square of size $N \times N$, estimated

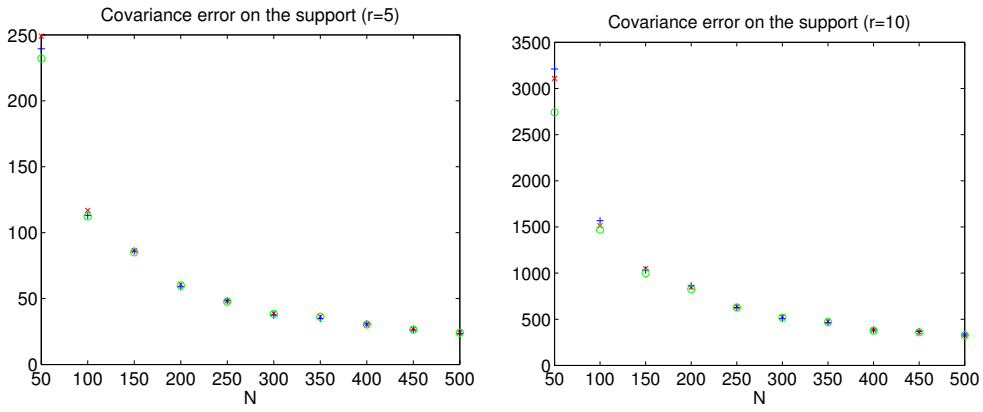


Figure 3.4: **Comparison of the different covariance estimators**. We plotted the estimation errors with the periodic (c_u, $\times$), non-periodic (γ_u, $+$), and non-periodic non-normalized (γ'_u, $\circ$) estimators of the covariance. More precisely, for each size N of the observation, the expectations of $\varepsilon(c_u)$, $\varepsilon(\gamma_u)$, and $\varepsilon(\gamma'_u)$ (which are defined by (3.11)) were estimated by averaging 100 realizations of these random variables. The true covariance is the autocorrelation of a discrete disc of radius r, with $r = 5$ for the left diagram, and $r = 10$ for the right diagram. One can see that the three estimators have approximately equal precision. It is surprising to see that the (natural) non-periodic normalized estimator γ_u does not perform better.

the mean by $\bar{u}$ and the covariance by c_u, γ_u and γ'_u. The estimation error of the estimator c is measured by

$$\varepsilon(c)^2 = \sum_{\mathbf{v}\in K} (c(\mathbf{v}) - C(\mathbf{v}))^2 \, . \tag{3.11}$$

In Fig. 3.4, we plotted an estimation of the expectation of $\varepsilon(c_u)$ and $\varepsilon(\gamma_u)$ and $\varepsilon(\gamma'_u)$ for several values of N (the observation size). This diagram shows that, surprisingly, the non-periodic estimator γ_u (which seemed *a priori* more natural) does not perform better than the periodic one.

Let us remark that the choice of the ℓ^2 norm over K for the comparison of the estimators is quite arbitrary. In the statistical community (see for example [Bickel & Levina 2008]), several articles use the $\ell^2 - \ell^2$ operator norm to study the convergence of an empirical covariance matrix estimate to the true covariance matrix of the model. In our context, the computation of this operator norm is practically difficult because we deal with random models defined over $\mathbb{Z}^2$. However, as in [Xiao & Wu 2012], one could consider the operator norm of the difference between the covariance matrix estimate and the true covariance matrix of the model restricted on the observation domain. But in our case, this operator norm would be difficult to compute (because the covariance matrices cannot be stored, and because the covariance matrix of a restriction is not circulant but only Toeplitz).

Since the non-periodic and periodic estimators perform as well, we may use either one, depending on the applicative context. For synthesis purpose, we will prefer covariance estimators that are true covariance functions associated to a ran-

dom model, because we want to sample the estimated model. Let us mention that c_u is a periodic autocorrelation and thus has a very simple DFT which is a non-negative function; in other words, c_u is a *true* covariance function of a periodic Gaussian random model, which can directly serve for synthesis. In contrast, the Fourier transform of γ_u is not non-negative so that γ_u is not a *true* covariance function. Therefore, one additionnal step would be needed in order to estimate a Gaussian model from γ_u. Actually, the fact that γ_u is not a true covariance function only comes from the normalization. On the other hand, the variant

$$\gamma'_u(\mathbf{v}) = \frac{1}{MN} \sum_{\mathbf{x}\in\Omega\cap(\Omega-\mathbf{v})} (u(\mathbf{x}) - \bar{u})(u(\mathbf{x}+\mathbf{v}) - \bar{u}) \ , \tag{3.12}$$

is the non-periodic autocorrelation of the normalized spot

$$\forall \mathbf{x} \in \mathbb{Z}^2, \quad t_u(\mathbf{x}) = \frac{1}{\sqrt{MN}}(u(\mathbf{x}) - \bar{u})\mathbf{1}_{\mathbf{x}\in\Omega} \ .$$

Therefore, γ'_u has a non-negative Fourier transform and is the covariance function of the $\mathbb{Z}^2$ ADSN field $t_u * W$ which will be used for synthesis in the next section.

3.2 Synthesis examples

In this section, we summarize the method of [Galerne *et al.* 2011b] for by-example synthesis of a microtexture on an arbitrarily large domain, using the ADSN model. We give several synthesis examples that allow us to exhibit the limitations of the ADSN (and RPN) models. Understanding these limitations is an important prerequisite in the design of richer texture models. Finally, we discuss the color handling and in particular compare the different possibilities that were exposed in the literature to synthesize color images with an ADSN or RPN model.

3.2.1 Synthesis Algorithm

Let us assume that we observe an exemplar texture $u : \Omega \longrightarrow \mathbb{R}^d$ on a discrete rectangle $\Omega \subset \mathbb{Z}^2$ of size $M \times N$. For the sake of simplicity, we assume that $0 \in \Omega$. As discussed in Section 3.1, we estimate the mean value by the empirical mean

$$\bar{u} = \frac{1}{|\Omega|} \sum_{\mathbf{x}\in\Omega} u(\mathbf{x}) \ ,$$

and the covariance function by $t_u * \tilde{t}_u^T$ where

$$\forall \mathbf{x} \in \mathbb{Z}^2, \quad t_u(\mathbf{x}) = \frac{1}{\sqrt{MN}}(u(\mathbf{x}) - \bar{u})\mathbf{1}_{\mathbf{x}\in\Omega} \ . \tag{3.13}$$

Then, if W is a Gaussian white noise with variance 1 on $\mathbb{Z}^2$ the Gaussian model

$$\bar{u} + t_u * W$$

can be used for synthesis of the texture u. In order to perform synthesis on a domain Ω_s, one only has to sample the restriction of this Gaussian random field

to Ω_s. For that, as explained in Subsection 2.2.6, one can consider a rectangle R containing $\Omega_s - \Omega$, embed t_u at any position in R, perform on R the circular convolution of t_u with a normalized Gaussian white noise (by using a DFT), add $\bar{u}$, and then crop the result to Ω_s.

As explained in [Galerne *et al.* 2011b, Subsection V.C], high-frequency artifacts can occur when using the normalized spot t_u because of the border discontinuity induced by zero-padding. As suggested in [Galerne *et al.* 2011b], such artifacts can be avoided by multiplying t_u by a smooth window which attenuates this border discontinuity.

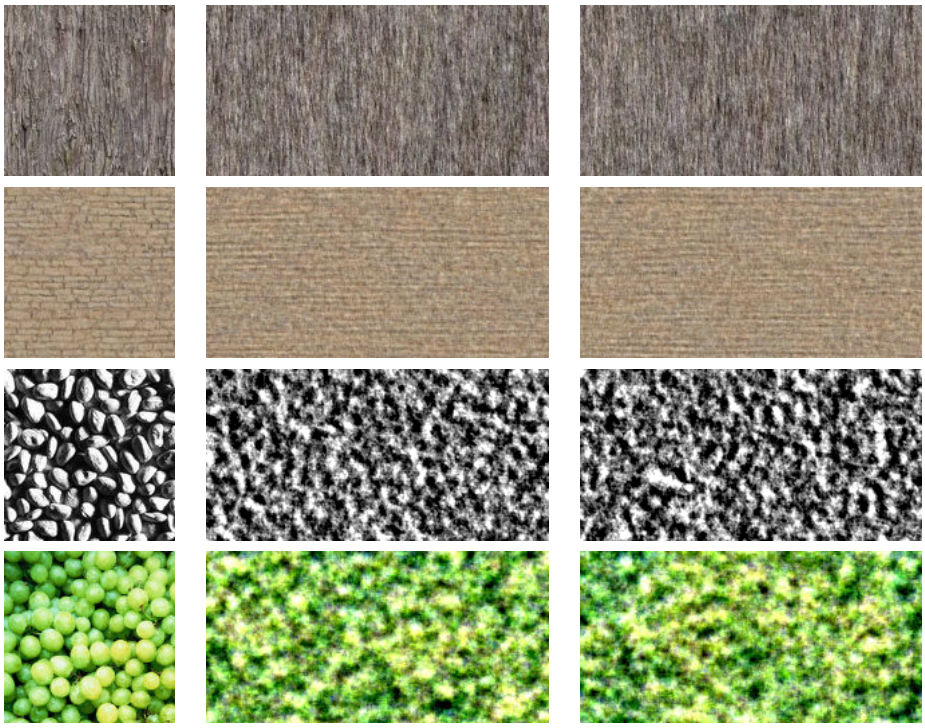
If, for any reason, one wishes to synthesize a texture that is periodic on a domain Ω_s with dimensions larger than Ω, then one only has to embed t_u in Ω_s, and perform the circular convolution of t_u with a white noise on Ω_s by using a DFT (the cropping step is not needed anymore). This corresponds to using a circular ADSN model instead of a non-circular ADSN model. For such a circular synthesis, one can also use the RPN model, which amounts to replace the convolution with a white noise by the multiplication of the DFT with a uniform random phase field (see Subsection 2.1.3). In the case $\Omega_s = \Omega$, then, as discussed in Subsection 3.1.2, one can take the periodic component [Moisan 2011] in order to make the circular model more legitimate.

3.2.2 Examples and Limitations of Uniform Random Phase Synthesis

Let us now give some synthesis examples. In Fig. 3.5, one can observe eight satisfying examples of ADSN and RPN synthesis. These examples can be referred to as microtextures because their perception is governed by non-salient details which are concentrated in the fine scales. As suggested by [Galerne *et al.* 2011b], one can define a class of microtextures as the textures that are characterized by their discrete Fourier spectrum. Such microtextures are exactly the ones for which the Fourier phase does not bear any relevant information, and thus can be changed without affecting the textural aspect. This is equivalent to say that such a microtexture is characterized by its autocorrelation. Since the synthesis with ADSN or RPN models (averagely) preserve the autocorrelation, this explains why such microtextures can be well reproduced through ADSN or RPN synthesis. One can also notice in Fig. 3.5 that in terms of visual perception of the synthesized textures, the RPN and ADSN models are very similar. More experiments would show that it is always the case except for an original texture with a very sparse spectrum (composed of one or two pure sine waves).

Let us now present other synthesis examples which highlight the limitations of uniform random phase synthesis. First, as can be seen in Fig. 3.6, the RPN and ADSN models fail to reproduce textures with large scale geometrical structures. Indeed, as we have seen in Section 2.1 for the circular ADSN and for the RPN, the Fourier coefficients of these random fields are independent modulo the Hermitian symmetry. Therefore, the RPN and ADSN models cannot capture the correlations between the phase coefficients that are needed to produce sharp edges (for example).

Figure 3.5: **ADSN and RPN synthesis of microtextures.** For each line and from left to right, one can observe an original texture image u (of size 128×128) and the results of ADSN and RPN synthesis (of size 256×128). On these microtextures, the synthesis results are satisfactory.



u ADSN(u) RPN(u)

Figure 3.6: **ADSN and RPN synthesis of non-random phase textures.** These uniform random phase models fail to reproduce the large-scale geometrical features which are mainly encoded in the phase information of the original texture.

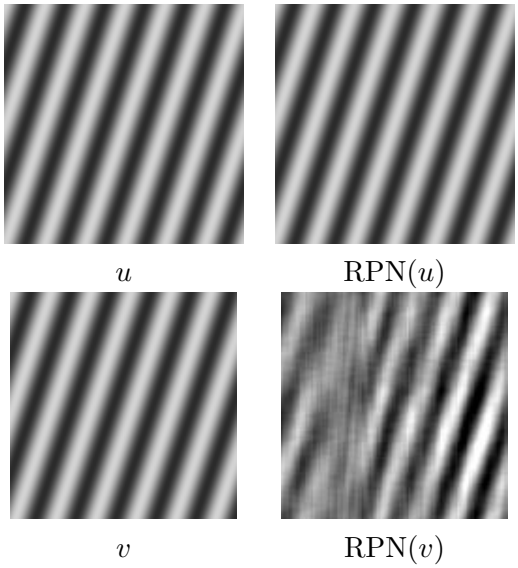
Measuring the coherence of the phase coefficients in an original image is a difficult question for which we will bring a partial answer in Chapter 4. Let us also remark that the first example of Fig 3.6 is not really homogeneous: the real problem of this texture is the wood knots. A basic requirement for RPN or ADSN synthesis is the spatial homogeneity of the exemplar.

The uniform phase constraint also leads to surprising results on nearly periodic textures. Since the Fourier representation is compatible with periodic structures, one could expect the RPN or ADSN models to be very well suited to periodic textures. But in general, the natural textures that are called periodic are generally not pure sine waves because the periodic patterns are often produced with several harmonic frequencies. Therefore, as can be seen in Fig. 3.7, interference patterns can appear on the random phase synthesis of a quasi-periodic texture. Let us also mention a surprising example found by B. Coulange which is shown in Fig. 3.8. This example highlights the fact that the RPN synthesis is very sensitive to the frequency sampling induced by the DFT: a pure sine wave that is not (M, N)-periodic does not have a sparse DFT on a domain of size $M \times N$, and thus the RPN synthesis may fail on such an example.

Another contraint of the RPN or ADSN model is the symmetry of the color distribution. Indeed, if W is a Gaussian white noise, the random field $h * W$ has

u ADSN(u) RPN(u)

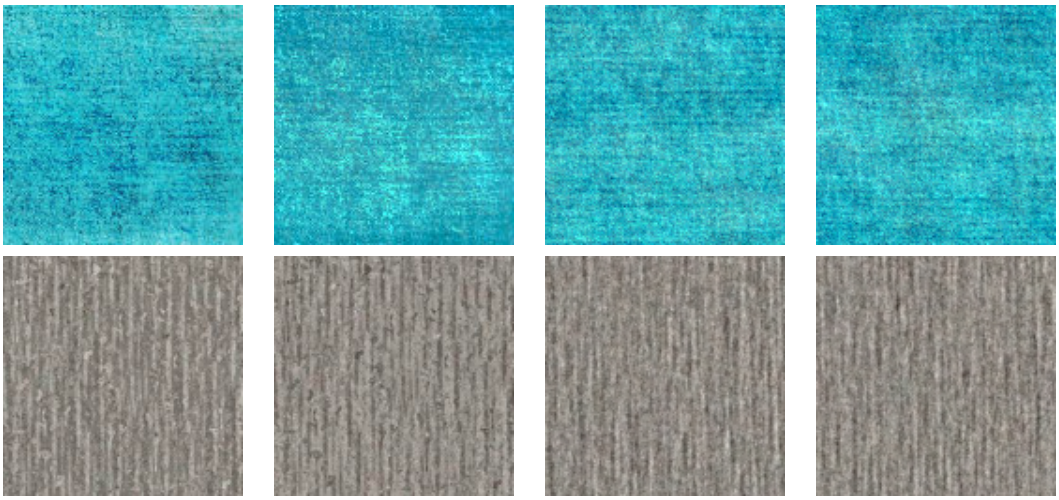
Figure 3.7: **ADSN and RPN synthesis examples with interference artifacts.** A "periodic" natural texture in general does not have a sparse spectrum, but is composed of several coherent harmonic frequencies. The uniform random phase synthesis cannot keep this coherence and thus may lead to undesirable interference patterns.



u RPN(u)

v RPN(v)

Figure 3.8: **A surprising RPN example.** The image u has a DFT supported by $\{0, \boldsymbol{\xi}_0, -\boldsymbol{\xi}_0\}$ and the image v was obtained by cropping 7 rows and 7 columns of u. Therefore, the DFT of v is not sparse anymore, which explains why the RPN of v does not look like a pure sine wave.



u $v = 2\bar{u} - u$ ADSN(u) RPN(u)

Figure 3.9: **ADSN and RPN synthesis of exemplars having a non-symmetrical color distribution.** From each row, and from left to right, the exemplar texture u, its symmetric v with respect to the empirical mean, and the results of ADSN and RPN synthesis. Notice that the assymetry of the exemplar color distribution is not preserved after uniform random phase synthesis; this can be explained by the fact that in these two examples, the textures u and v can be (preattentively) discriminated.

same distribution than $(-h) * W$. Therefore, if we replace the original texture u by

$$\bar{u} - (u - \bar{u}) = 2\bar{u} - u \ ,$$

then we get exactly the same result of ADSN synthesis. A similar remark can be expressed for the RPN model. As can be seen in Fig. 3.9, this implies that an exemplar with a color distribution that is not symmetrical with respect to the empirical mean may not be well reproduced by uniform random phase synthesis. More generally, if T is an image transformation such that $h * W$ and $T(h) * W$ have the same distribution (for example the spatial symmetry $h \mapsto \tilde{h}$), and if one can find an exemplar u such that $T(u)$ does not look like u, then it is likely that the ADSN fails to synthesize u.

We would like to remark that the random phase hypothesis depends on the observation scale of the texture. About that, let us comment and illustrate the following sentence of [Galerne *et al.* 2011b]:

> *"Nonetheless, each textured object has a critical distance at which it becomes a micro-texture."*

Starting from a high-resolution image of a texture, one can simulate the observation at different distances by cropping (with the same number of pixels) subsampled versions of the texture with different sampling rates. Next, one can perform uniform random phase synthesis of the different versions. On the example shown in Fig. 3.10,

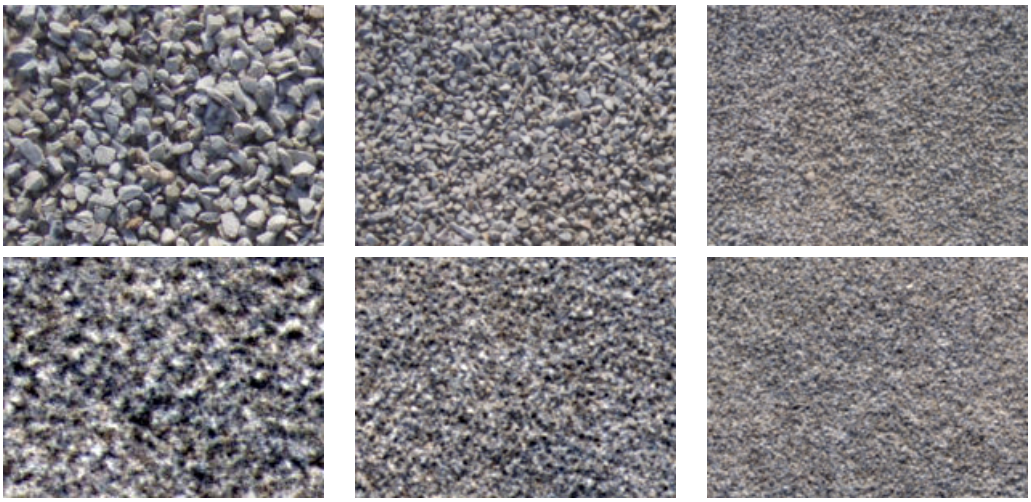


Figure 3.10: **ADSN synthesis through scale.** In the first row, one can see three exemplar textures of size 200 × 150; they have been extracted from three subsampled versions of an image of size 800 × 600 at different sampling rates. In the second row, one can see the corresponding results of ADSN synthesis. Notice that the ADSN synthesis is satisfactory for the right image, nearly satisfactory for the middle image, and not satisfactory at all for the left image. This illustrates that a structured texture may be considered as a random phase texture if it is observed from sufficiently far away.

we see that the ADSN synthesis is more suited to the texture seen from far away because the geometrical details are not salient anymore; this example thus confirms the quoted sentence. However, it is worth questioning this assertion because its validity relies on the (non-clear) definition of "textured object". Actually, one could define an interesting class of textures by considering the homogeneous images that can be considered as random phase as soon as they are seen from sufficiently far away. Such a definition can be precised by using a mathematical operator that emulates the loss of resolution due to the observation distance (a blurring operator for example). The synthesis of such textures will be addressed in Section 5.3.

3.2.3 Handling the Color Distribution

In Subsection 3.2.1, it appeared natural to consider the normalized spot

$$t_u = \frac{1}{\sqrt{MN}}(u - \bar{u})$$

and to synthesize the texture u with $F = \bar{u} + t_u * W$. This way, the expectation of F is exactly the empirical mean of u, and the color covariance $C_F(0)$ of F is exactly the empirical color covariance

$$c_u(0) = \gamma_u(0) = \frac{1}{MN}\sum_{\mathbf{x}\in\Omega}(u(\mathbf{x}) - \bar{u})(u(\mathbf{x}) - \bar{u})^T \ .$$

However, in the texture synthesis literature, several other propositions have been made in order to perform the synthesis of a RGB image

$$u = \begin{pmatrix} u_R \\ u_G \\ u_B \end{pmatrix} : \Omega \longrightarrow \mathbb{R}^3 \ ,$$

and we shall discuss them in this subsection.

First, let us remark that performing independent ADSN synthesis of the three color channels in general leads to incorrect results. Indeed, if W_R, W_G, W_B are three independent normalized Gaussian white noises, the random field

$$\begin{pmatrix} h_R * W_R \\ h_G * W_G \\ h_B * W_B \end{pmatrix}$$

has a marginal color covariance equal to

$$\begin{pmatrix} \|h_R\|^2 & 0 & \\ 0 & \|h_G\|^2 & 0 \\ 0 & 0 & \|h_B\|^2 \end{pmatrix} \ ,$$

which may not lead to a good approximation of the color distribution of the exemplar. This explains why unnatural colors appear when doing independent synthesis of the color channels.

Another method consists in finding a linear transformation of the RGB space which "decorrelate" the channels in a certain sense. For example, the authors of [Heeger & Bergen 1995] suggest to use a transformation matrix $P \in \mathcal{O}_3(\mathbb{R})$ associated to an eigenvector basis of the empirical color covariance $c_u(0)$. After applying this transformation to the exemplar texture, one can perform independent synthesis on each channel, and apply the inverse P^T of P to get the synthesized texture in the original RGB space. Writing p_1, p_2, p_3 the columns of P, this amounts to synthesize u with the Gaussian random field

$$F = m + \sum_{j=1}^{3} p_j (p_j^T t_u) * W_j$$

where W_1, W_2, W_3 are three independent normalized Gaussian white noises. One can see that the color covariance $C_F(0)$ of F is equal to the empirical color covariance of u so that the marginal color distribution is preserved; thus this method in general does not lead to dramatic color artifacts. Notice that in Fourier domain, for $\boldsymbol{\xi} \neq 0$, the covariance of $\hat{F}(\boldsymbol{\xi}) \in \mathbb{C}^3$ is given by

$$\sum_{j=1}^{3} p_j p_j^T \hat{t}_u(\boldsymbol{\xi}) \hat{t}_u^*(\boldsymbol{\xi}) p_j p_j^T \ .$$

The last method allows to decorrelate the color channels in a marginal sense, but one can also apply other color transformation matrices that decorrelate the

color channels in a stronger sense. For example, the authors of [Galerne *et al.* 2012] suggested to find a matrix P which leads to approximate joint diagonalization of all the matrices $c_u(\mathbf{h})$. One could also look for a matrix P which allows approximate joint diagonalization of all the matrices $\hat{c}_u(\boldsymbol{\xi})$. Both these methods based on joint diagonalization seem relevant.

In definitive, since the ADSN synthesis is defined naturally on color images by using the normalized spot t_u, there is no reason to look for a transformation matrix P that authorizes independent synthesis on each channel. In fact since convolution operators are diagonal in the Fourier basis, given an estimator of the texture covariance c, it is relevant to extract, for each frequency $\boldsymbol{\xi}$, the principal components of $\hat{c}(\boldsymbol{\xi})$ and to work independently on these principal components, or even to keep only the first principal component as suggested by the decomposition of Theorem 2.1.2. But, with the periodic estimator c_u of the covariance, for $\boldsymbol{\xi} \neq 0$, the matrix

$$\hat{c}_u(\boldsymbol{\xi}) = \hat{t}_u(\boldsymbol{\xi})\hat{t}_u(\boldsymbol{\xi})^*$$

has only one non-zero principal component, so that proceeding this way only amounts to perform the convolution $t_u * W$ as was initially proposed.

In conclusion, we confirm that synthesizing u by the random Gaussian field $m + t_u * W$ is very natural, very simple, and so far, we have found no better way to perform the synthesis in the Gaussian framework. Finally, as was done in [Galerne *et al.* 2011b], let us emphasize one more time that for ADSN synthesis, it is important to convolve the three channels of t_u with the same white noise; and for RPN synthesis, it is important to multiply the DFTs of the three channels of t_u by $e^{i\psi}$ where ψ is the same uniform random phase.

3.3 A Texton for Fast and Flexible Synthesis

We have seen in Subsection 2.1.2 and Subsection 2.2.2 that a Gaussian texture can be approximated by a high-intensity DSN. The direct simulation of the DSN is simple and allows parallel local evaluation using standard computer graphics techniques for the Poisson process simulation [Lagae *et al.* 2009]. Still, the DSN approximation of a Gaussian texture is satisfactory only for sufficiently high intensity λ, so that the DSN simulation is generally not faster than the spectral simulation. In particular, using the canonical texton introduced in [Desolneux *et al.* 2012] as a kernel for DSN synthesis generally results in a very poor approximation for small values of λ.

In this section we show that, given an exemplar texture image u, it is possible to compute an approximate compactly-supported texton of $\mathcal{N}(0, c_u)$ (Gaussian model estimated from u) having a prescribed small support and for which the realizations of the associated DSN are visually similar to the ones of $\mathcal{N}(0, c_u)$, even for a low intensity λ (see Fig. 3.11). This so-called synthesis-oriented texton (SOT), which can be considered as an inverse texture synthesis solution [Wei *et al.* 2008] for the Gaussian model, is computed using the classical error reduction algorithm, introduced in [Fienup 1982] for phase retrieval, with a uniform random phase initialization, see Subsection 3.3.1. As will be shown in Subsection 3.3.2, for an average number of 30

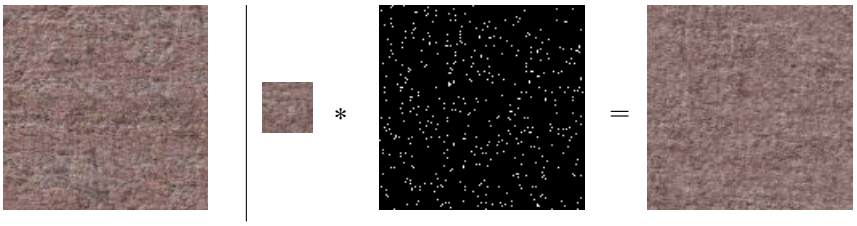


Figure 3.11: **Spot noise synthesis at low intensity.** The synthesized texture on the right was obtained by the convolution of a synthesis-oriented texton with a sparse Poisson process. The exemplar texture is shown on the left.

impacts per pixels, the DSN associated with the SOT produces visually satisfying results, and is thus more competitive than the spectral simulation algorithm. In Subsection 3.3.3, we present a failure case of SOT computation which allows us to discuss one drawback of the optimal transport distance between Gaussian models: its definition does not take into account our more important sensitivity to high frequencies. Finally, in Subsection 3.3.4, we propose a modified algorithm which computes a SOT whose associated ADSN restitutes the high-frequency components of the exemplar texture in a more faithful way.

A part of our work on the synthesis-oriented texton has been published in the conference paper [Galerne *et al.* 2014].

Let us mention that for technical convenience, the SOT computation will be explained in the framework of circular Gaussian models.

3.3.1 Alternating Projections

Let us assume that we observe an exemplar texture $u : \Omega \longrightarrow \mathbb{R}^d$ on a domain Ω of size $M \times N$. We will assume that the convolution operations are performed with periodic boundary conditions. Let us introduce a small subset $S \subset \Omega$ (which can be understood as a memory budget). For the sake of clarity, we will assume that S is symmetrical with respect to zero: $S = -S$. The goal of this subsection is to compute a kernel $t : \Omega \longrightarrow \mathbb{R}^d$ with support $S_t \subset S$, whose associated DSN realizes a visually satisfying synthesis of u even for a low intensity λ.

In order to measure the distance between two circular random fields on Ω, we use the optimal transport distance defined in Subsection 2.1.6. Thus we would like that, in terms of this optimal transport distance, the DSN associated to t realizes a good approximation of the Gaussian model $\mathcal{N}(0, c_u)$, which is the ADSN associated to the normalized spot

$$t_u(\mathbf{x}) = \frac{1}{\sqrt{\Omega}}(u(\mathbf{x}) - \bar{u}) \ .$$

Unfortunately, as we have said in that subsection, we do not know how to compute the optimal transport distance between a finite intensity DSN and an ADSN. Therefore, we will rather require that the circular ADSN associated to t realizes a good approximation of $\mathcal{N}(0, c_u)$ and that the circular DSN associated to t quickly converges to its Gaussian limit.

The approximation of the Gaussian models with a prescribed small support is already not a trivial problem. A perfect solution would be a kernel t with support $S_t \subset S$ and such that $c_u = t * \tilde{t}$. Taking the Fourier transform, this equality rewrites

$$\widehat{c_u} = \hat{t}\hat{t}^*$$

so that this problem is an analog of the phase retrieval problem [Hayes 1982], [Fienup 1982], [Levi & Stark 1984], [Bauschke *et al.* 2002]. It may lead to multiple solutions (if t is a solution, so are $-t$, $\tilde{t}$ or $-\tilde{t}$, and in particular cases there may be other solutions, see [Hayes 1982]). Here, because of the constraint on S_t, there is no exact solution in general, but we can look for an approximate solution by trying to solve

$$\underset{\substack{t:\Omega \longrightarrow \mathbb{R}^d \\ S_t \subset S}}{\text{Argmin}} \quad d_{OT}(\mu_t, \mu_{t_u}) \tag{3.14}$$

where μ_t is the circular ADSN on Ω associated to the kernel t. Let us recall that $t_u = \frac{1}{\sqrt{|\Omega|}}(u - \bar{u})$ where $\bar{u}$ is the empirical mean of u, and that the optimal transport distance in the last formula is given by

$$d_{\mathrm{OT}}(\mu_t, \mu_{t_u})^2 = \frac{1}{|\Omega|} \sum_{\boldsymbol{\xi} \in \Omega} \|\hat{t}(\boldsymbol{\xi})\|^2 + \|\widehat{t_u}(\boldsymbol{\xi})\|^2 - 2|\,\hat{t}(\boldsymbol{\xi})^* \widehat{t_u}(\boldsymbol{\xi})| \ .$$

Since it is not a convex function of t, the problem (3.14) is not a convex optimization problem. However, one can tackle this problem by alternating between the projection on the support constraint

$$q_S(t) = t\,\mathbf{1}_S \tag{3.15}$$

(which is actually the orthogonal projection on a convex set), and the projection on the model constraint which has been shown in Subsection 2.1.6 to be given in Fourier domain by

$$\widehat{p_{t_u}(h)} = \frac{\hat{t}_u \hat{t}_u^* \hat{t}}{|\hat{t}_u^* \hat{t}|} \mathbf{1}_{\hat{t}_u^* \hat{t} \neq 0} + \hat{t}_u \mathbf{1}_{\hat{t}_u^* \hat{t} = 0} \ . \tag{3.16}$$

This alternating projection algorithm was introduced and called "error reduction algorithm" by the author of [Fienup 1982].

Now, a difficult task is to incorporate in this optimization problem the constraint that the DSN associated to t realizes a good approximation of the corresponding ADSN even for low intensity λ. Actually, we do not even dispose of a perceptually-compliant criterion to assess the convergence speed of the DSN towards the ADSN. In the first place, one can think that this convergence speed can be analyzed through the marginal distributions. As was done in [Galerne 2010, Theorem 3.4], one can show (using a Berry-Esseen theorem on Poisson random sums) that the Kolmogorov distance between the marginal distributions of the DSN with intensity λ and the ADSN (i.e. the uniform distance between the cumulative distribution functions) is less than

$$\frac{\Gamma}{\sqrt{\lambda}}\, \sigma(t)^3 \ ,$$

where Γ is a universal constant and where

$$\sigma(t) = \frac{\|t\|_3}{\|t\|_2} = \left(\sum_{\mathbf{x}\in S} |t(\mathbf{x})|^3\right)^{1/3} \left(\sum_{\mathbf{x}\in S} t(\mathbf{x})^2\right)^{-1/2}$$

can be thought of a sparsity measure of t. Even if the exact convergence speed is not known, this indicates that the convergence tends to be faster if the kernel t exploits all the available support S.

However, the analysis of the marginal distributions is clearly not sufficient because the human visual system is more sensitive to local gradients than marginal distributions. More intuitively, an observer will not be able to distinguish between realizations of the DSN and the ADSN if the geometrical details of the kernel t (in particular, the limitation of its support) are not easily perceptible in the DSN with low intensity. This is why we will seek for a kernel t which has as less salient features as possible; in other words, the kernel t must have a minimally structured phase information.

Therefore, one possible way to incorporate the requirement of fast Gaussian convergence in the algorithm is to use a uniform random phase initialization. This choice is less satisfactory than the incorporation of a term that would soundly reflect the visual convergence speed of the DSN in the objective function (3.14). But we will see that it already leads to interesting textons (and in particular more efficient for synthesis than the ones that we obtained by minimizing $d_{OT}(\mu_t, \mu_{t_u}) + \lambda\sigma(t)$ for a certain constant λ). We can now summarize the computation of the synthesis-oriented texton associated to the texture u.

Algorithm: SOT computation

- Initialization: $\widehat{t} \leftarrow \widehat{t}_u e^{i\psi}$ where ψ is a uniform random phase function, and $t_u = \frac{1}{\sqrt{|\Omega|}}(u - \bar{u})$.
- Repeat (n times) $t \leftarrow q_S(p_{t_u}(t))$.

Let us remark that the alternating projection algorithm with zero-phase initialization was already proposed in [Cadzow *et al.* 1993] to compute a kernel associated to a moving-average model. But the zero-phase initialization does not lead to kernels that can be used for fast DSN synthesis.

The questions of the convergence and the influence of the random initialization were raised by [Hayes 1982]. We will see that both these issues are negligible in terms of the resulting Gaussian texture. Even if we can show that the value of the objective function decreases along the iterations (which explains the name "error reduction"), the convergence of the iterates is not ensured (recall that the objective function is not convex). Indeed, we analyzed the behavior of the algorithm by considering the (squared) relative model error

$$\mathrm{RME}(t, t_u)^2 = \frac{\sum_{\boldsymbol{\xi}} \|\widehat{t}_u(\boldsymbol{\xi})\|^2 + \|\widehat{t}(\boldsymbol{\xi})\|^2 - |\widehat{t}_u(\boldsymbol{\xi})^* \widehat{t}(\boldsymbol{\xi})|}{\sum_{\boldsymbol{\xi}} \|\widehat{t}_u(\boldsymbol{\xi})\|^2} \,. \tag{3.17}$$

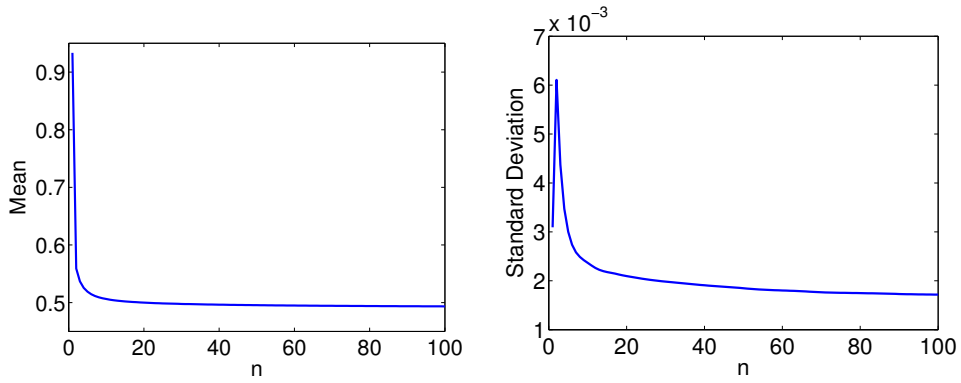


Figure 3.12: **Iterates of the alternating projections algorithm.** Evolution of the empirical mean (left) and standard deviation (right) of the RME computed after n iterations of the alternating projection algorithm (estimated over 1000 samples). Observe that the mean RME quickly decreases, which shows that most of the Gaussian model approximation is done in the first iterations. Notice also that the standard deviation does not tend to zero; this reflects that the algorithm does not have a unique convergence point.

The numerator is (up to the constant $\frac{1}{|\Omega|}$) the optimal transport distance between the circular ADSN $t_u * W$ and $t * W$, and the denominator is (up to the same constant) the marginal variance of $t_u * W$.

A direct observation of the iterates shows that for each random initialization, they seem to stabilize after a small number of iterations, as already mentioned in [Hayes 1982]. To be more precise, we computed the empirical mean $\bar{d}_n$ and variance $\bar{\sigma}_n^2$ of the random variable $D_n = \mathrm{RME}(t_u, T_n)$, where T_n is the SOT obtained after n iterations of the algorithm with random initialization. As one can see in Fig. 3.12, $\bar{d}_n$ and $\bar{\sigma}_n$ do not change much for $n \geqslant 50$, reflecting again the quick stabilization of the iterates.

We also investigated the idea of running the algorithm several times with different random initializations and selecting the output with the smallest RME, but numerical simulations showed that the improvement in RME (for a fixed computation time) was not significant (below 1%).

The SOTs presented in the next subsection were obtained by applying the alternating projections algorithm with only one random initialization and 100 iterations.

3.3.2 Results

In each example of DSN synthesis, we precise the number of impacts per pixel

$$N_{imp} = \lambda |S| \ ,$$

which represents the expected number of points of the Poisson process that will be involved in the computation of one value of the spot noise. It is the relevant constant for computational comparison of DSN synthesis with different kernels.

Let us first show in Fig. 3.13 some examples of SOTs associated to synthetic Gaussian textures. We can see that the DSN synthesis with the SOT is generally

satisfactory in terms of frequency content, even for a low number of impacts per pixel. Using the SOT, the direct summation method thus becomes a competitive way of synthesizing Gaussian textures, with an expected number of operations per pixel below 100.

A comparative diagram is shown in Fig. 3.14. They first confirm that the distinction between the DSN and the ADSN associated to the SOT is difficult, even for a very low mean number of impacts per pixel (thirty). Next, it indicates that the synthesis results are as good as Gabor noise by example [Galerne *et al.* 2012] which requires at least ten times more operations per pixel. Remember though that the Gabor noise was designed to perform procedural synthesis, which justifies its heavier computational cost. Another remark that can be drawn from Fig. 3.14 is that the SOT outperforms the canonical/luminance texton of [Desolneux *et al.* 2012] in terms of Gaussian model approximation. This confirms that the cropped canonical/luminance texton does not realize the minimal model error for a prescribed compact support, even if it solves a similar optimization problem with another concentration criterion [Desolneux *et al.* 2015]. In terms of DSN synthesis, one can see that the luminance texton is clearly not appropriate: since it presents a strong spike located at zero, the Gaussian convergence of the DSN is very slow; this defect can be observed through the simplistic analysis of marginal distributions as we will see later.

Let us comment the number of operations per pixel. Again, the expected number of operations per pixel for DSN synthesis equals the expected number of impacts per pixel, that is $N_{imp} = \lambda|S|$. Considering the Poisson process to be previously drawn, the mean complexity of the DSN synthesis on a domain Ω is thus $\lambda|S||\Omega|$. As we have just seen, the SOT allows to obtain satisfying DSN synthesis results with a mean number of impacts per pixel equal to 30; with this parameter, the mean complexity of DSN synthesis on Ω is thus $30|\Omega|$. In comparison, the complexity of the circular ADSN synthesis using the spectral method is the same than the complexity of the FFT on Ω, which is at most $4|\Omega|\log(|\Omega|)$ operations (and even $2|\Omega|\log(|\Omega|)$ when $|\Omega|$ is a power of two). Therefore, the DSN synthesis with the SOT will be more efficient for very big images with a domain $|\Omega|$ satisfying

$$\log_2(|\Omega|) > 30/4 = 7.5$$

or even $> 30/2 = 15$ when $|\Omega|$ is a power of two. But the real interest of the DSN synthesis over spectral simulation is its flexibility. Indeed, using a coherent evaluation procedure for the Poisson point process (explained in [Lagae *et al.* 2009] or in Fig. 2.6), the DSN synthesis can be parallelized; in contrast, the parallelization of the spectral simulation scheme is more difficult and certainly not adapted to the requirements of GPU programming (in particular, very low memory storage).

As can be observed in Fig 3.13, one drawback of the raw SOT is that it does not preserve the marginal color distribution of the exemplar. A similar observation was drawn in [Desolneux *et al.* 2015] for the cropped luminance texton and we suggest to apply the same color correction that is explained in the rest of the paragraph. Let us only analyze what happens in the Gaussian case where the distribution is characterized by first and second order moments. There is no problem on the mean

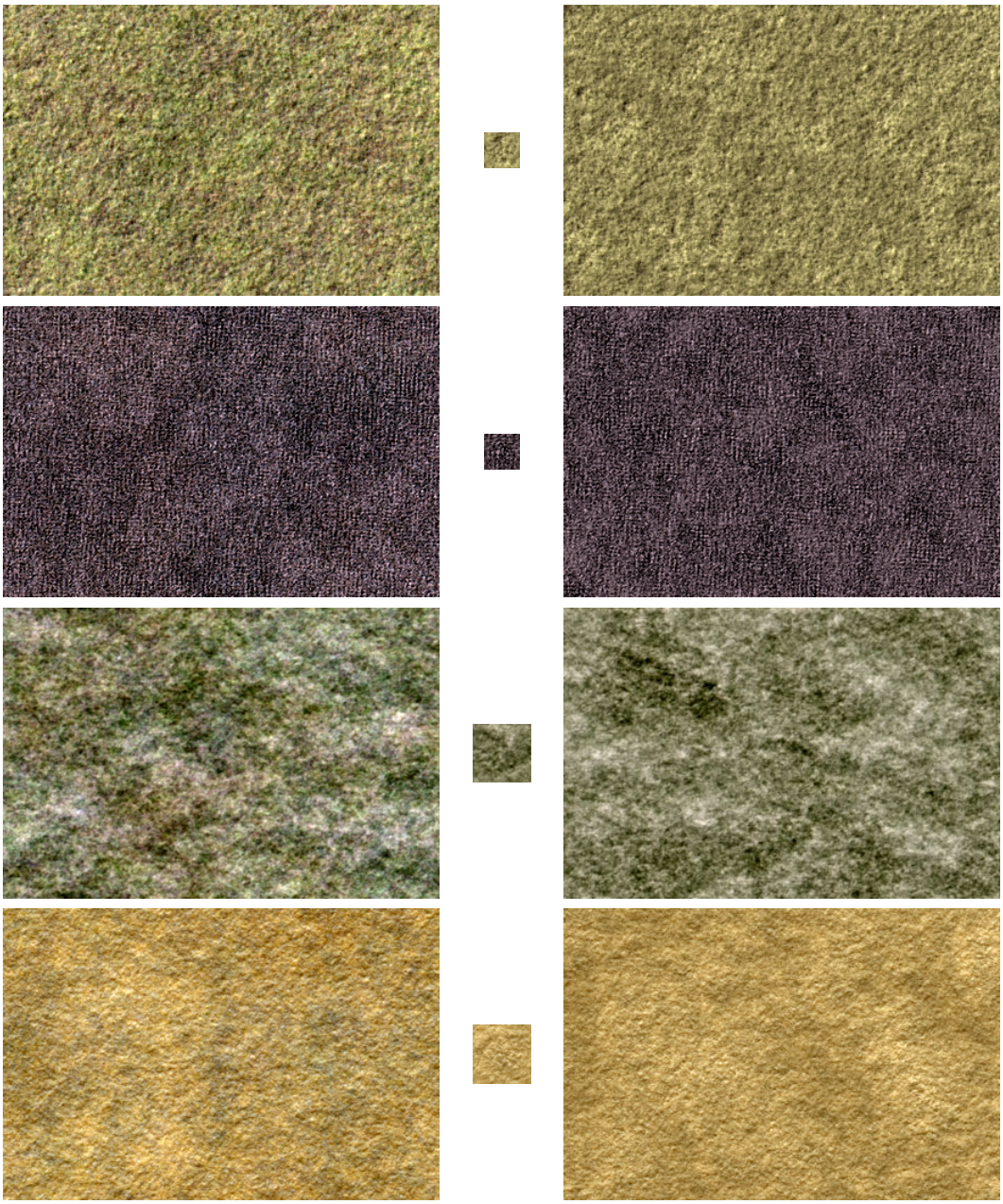


Figure 3.13: **Examples of DSN Synthesis with the SOT**. One can see on the left some synthetic Gaussian textures (of size 384×256), and on the right the results of DSN synthesis (with 50 impacts per pixel) using the SOTs shown in the middle. The SOT has square support of size 31×31 in the two first examples and 51×51 in the two last examples. Notice that the texture grain is well preserved which means that the SOT realizes a good approximation of the Gaussian model. However, one can notice also a slight loss in color diversity.

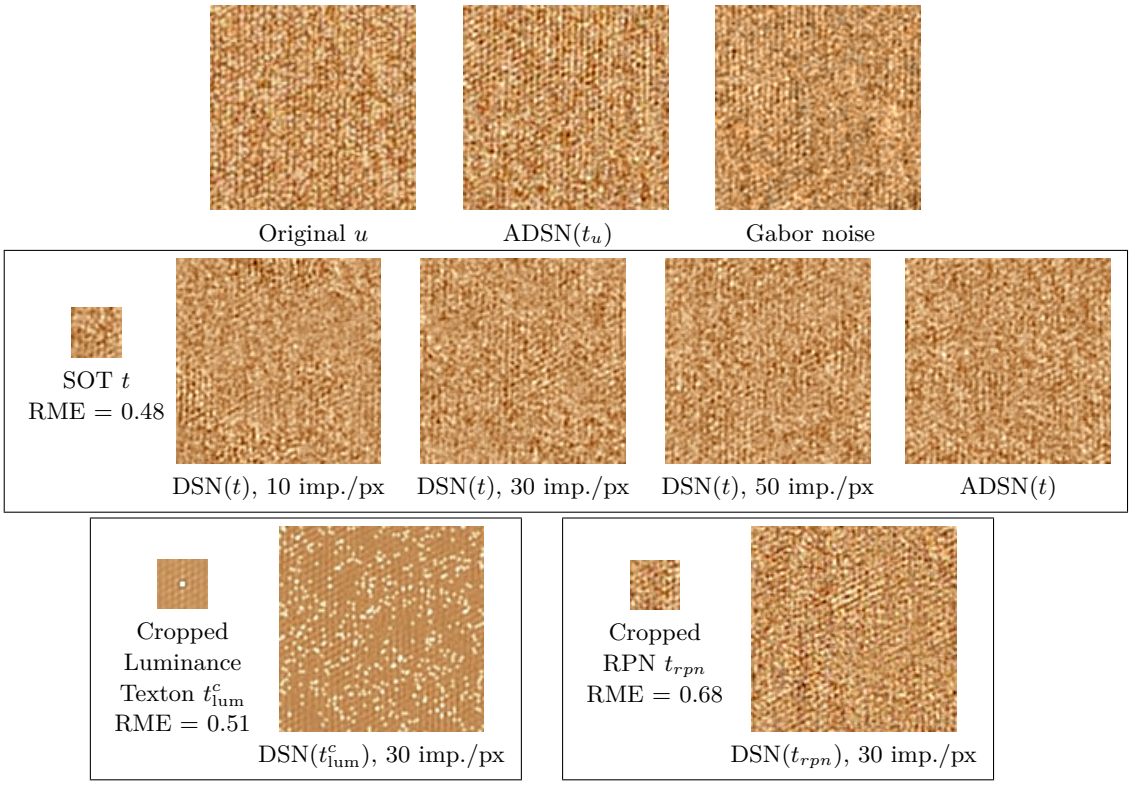


Figure 3.14: **DSN synthesis of a natural color texture, comparison.** First row: original texture (u), ADSN synthesis with the kernel t_u, and result of Gabor noise synthesis [Galerne *et al.* 2012]. Second row: DSN and ADSN synthesis results obtained with a 31×31 SOT t. The DSN intensities were set in order to match a given average number of impacts per pixel. Bottom row: DSN obtained with the cropped luminance texton t^c_{lum} [Desolneux *et al.* 2012], DSN obtained with a texton t_{rpn} that was cropped from a RPN realization of u. Each DSN model is displayed with its corresponding kernel. Contrary to other DSN models, the proposed SOT achieves a good visual proximity with the reference model ADSN(t_u) as the number of impacts per pixel attains 30. It also defines the most accurate asymptotic model (smallest RME).

value because our synthesis scheme imposes the expectation to be exactly equal to the empirical mean of the exemplar texture. But for the second-order moments, after the alternating projections algorithm, nothing ensures that

$$C_{t*W}(0) = t * t^T(0) = \sum_{\mathbf{x}\in S} t(\mathbf{x})t(\mathbf{x})^T$$

is equal to the marginal color covariance $c_u(0)$ of the exemplar texture.

However, one can apply a color transformation to the SOT in order to correct the marginal color distribution of the resulting DSN. Indeed, the matrices $A = t * t^T(0)$ and $B = c_u(0)$ are symmetrical and non-negative and thus admits a unique non-negative square root (which can be easily computed with an eigenvector basis); then, as soon as A is invertible (which is always the case in practice), one can see that for any $P \in \mathcal{O}_3(\mathbb{R})$, the kernel with color correction

$$t_P = B^{1/2} P A^{-1/2} t \tag{3.18}$$

satisfies

$$t_P * t_P^T(0) = B^{1/2} P A^{-1/2} t * t^T(0) (A^{-1/2})^T P^T (B^{1/2})^T = B = c_u(0)$$

so that the DSN associated to t_P has the correct marginal color distribution. The choice of P will not be thoroughly questioned here (as opposed to [Desolneux *et al.* 2015]); the SOTs with color correction that we present in the following experiments are obtained with $P = \mathrm{Id}$; we thus use the notation $t_{cc} = t_{\mathrm{Id}}$. Of course, correcting the marginal color covariance may increase the relative model error measured by the optimal transport distance: we may have

$$\mathrm{RME}(t_u, t_{cc}) > \mathrm{RME}(t_u, t) \ .$$

In the experiments presented in the following figures, we always applied the color correction of (3.18) with $P = I_3$. As one can see in Fig. 3.15, the color correction in general increases the perceived quality of DSN synthesis, even if it slightly increases the relative model error. This indicates that the optimal transport distance (on which relies the model error) may not reflect faithfully our texture perception; this fact will be confirmed by other practical observations in Subsection 3.3.3.

Let us also mention that in a fews cases, we observed that this color correction does not suffice to restitute the perceived colors of the original texture. On the counter-example of Fig. 3.16, we see that the color correction leads to a synthesized texture with perceived colors that are actually farer from the colors of the original texture. The fact that such a counter-example can be found in the very specific Gaussian case may be quite surprising at first sight; but it is easily explained by the fact that our color perception does not depend only on the first order distribution of the pixel values, but also on the spatial mixing of the pixel colors. As was kindly suggested by J. Delon, a very elegant and simple way to confirm that is to shuffle the pixel values in the images that we have to compare: the results can also be observed in Fig. 3.16. Notice that for such a counter-example, the RME is greatly

increased by the color correction. Notice also that this problem does not appear anymore with a larger support for the SOT. This means that, with a too small support, the SOT computation leads to an error (in terms of texture perception) that cannot be rectified through the adjustment of the color distribution.

One can find in Fig. 3.17 several examples of DSN synthesis results with the color corrected SOT associated to natural microtextures. These results confirm that many microtextures can be synthesized using a DSN associated to a small kernel (31×31) and with a very low number of impacts per pixel (30).

Let us now discuss the random phase initialization of the alternating projection algorithm. We have said that this initialization is important to have a fast convergence of the DSN associated to the SOT towards its Gaussian limit. Here we will compare to the case where the alternating projections algorithm is initialized with a zero-phase image. A detailed comparison would require a precise perceptual study of the visual convergence; here we only analyze the visual convergence through the visual evaluation of the results and an analysis of the marginal distributions. In Fig. 3.18, one can compare the results obtained with the zero-phase and the random phase initialization. Let us denote the corresponding textons t_{zp} and t_{rp}. As we have said, the Gaussian convergence of the DSN is much slower with t_{zp} than with t_{rp}. Indeed, one can see in the upper part of Fig. 3.18 that the convergence of the marginal distributions (in terms of the Kolmogorov-Smirnov distance, that is, the L^∞ distance between the cumulative distribution functions) is much slower with t_{zp}: it only needs around 20 impacts per pixel with t_{rp} for the error on marginal distributions to be of order 0.01, whereas for a similar intensity, the error is ten times higher with t_{zp}. This is confirmed by the visual inspection of the DSN with low intensity: at 30 impacts per pixel, the central spike of t_{zp} is still clearly visible in the DSN; this "white dots" effect also explains the error made on marginal distributions. Notice that, in terms of the model error, the zero-phase initialization leads to a slightly better value (which is not reflected by a perceptible difference in the corresponding ADSN). This confirms that the quality of a synthesis-oriented texton must not be measured only through the model error.

Let us end this subsection by discussing the texton support. In the previous experiments, we have only shown SOTs computed with a square support. We used this simple choice because it complies well with the DSN computation based on the parallel sampling of the Poisson point process illustrated in Fig. 2.6. However, any shape is *a priori* possible; in particular, in terms of image, it would be more natural to use a circular support because it does not favor any direction. In the next subsection, we will see that horizontal and vertical artifacts may be sometimes encountered on the synthesis results with a square SOT.

Once the shape has been chosen, one can also question the support size. The size of the support constrains the dependency range in the corresponding ADSN because the covariance of the ADSN associated to the SOT is supported by $S - S$. Therefore, in a sense, replacing a Gaussian texture by the ADSN associated to the SOT amounts to approximate the covariance function by a compactly-supported one. Another way to perform such a compact approximation of the covariance function is to replace the Gaussian texture by the ADSN associated to the cropped

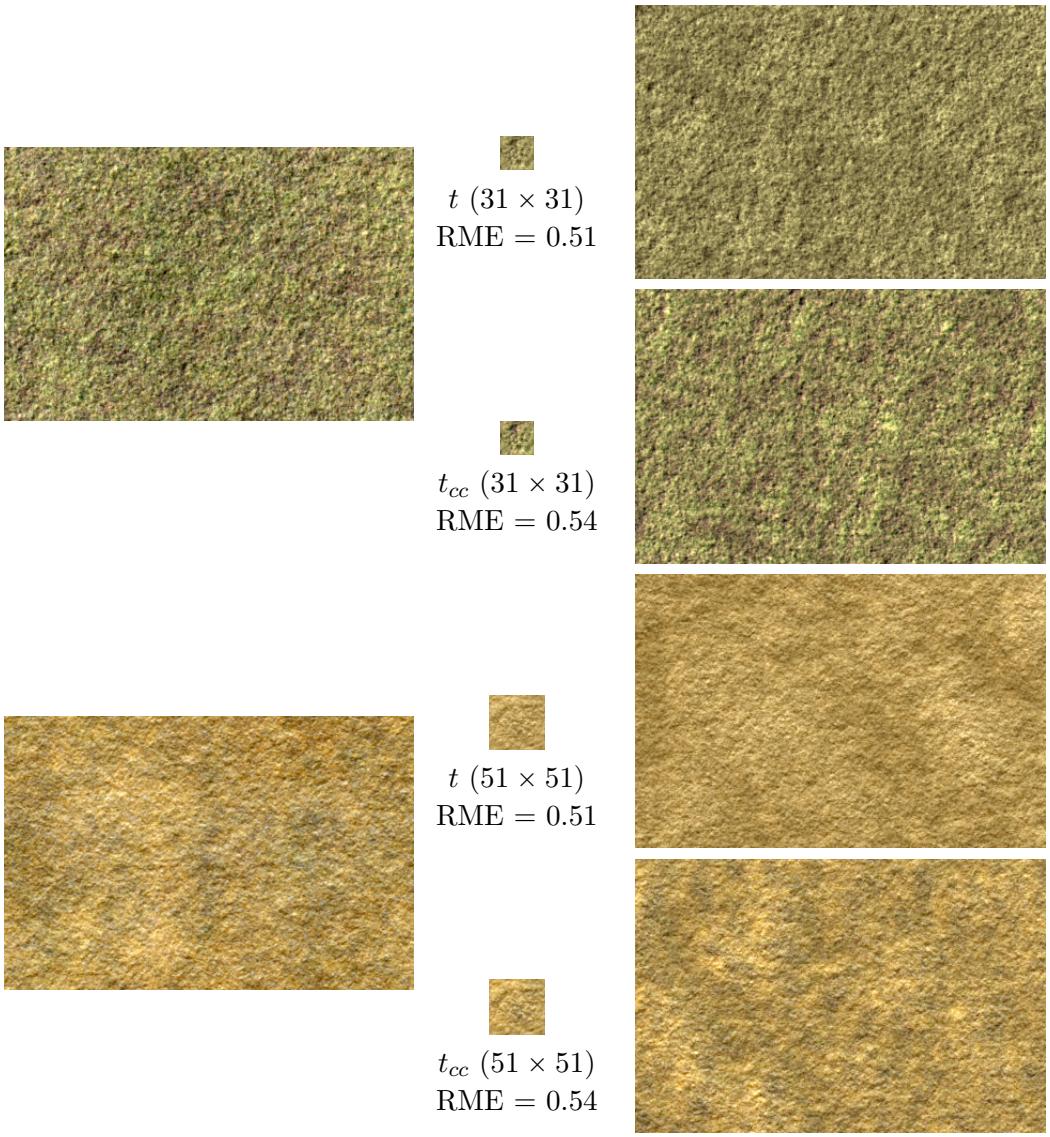


Figure 3.15: **Color correction of the SOT**. The upper and lower part of the figure is composed as follows. Left : Original texture. Middle : SOT t, and SOT t_{cc} with color correction. Right : DSN synthesis results with t (top) and t_{cc} (bottom) with 50 impacts per pixel. Even if the RME increases after color correction, the texture grain is still well preserved, with a more faithful color distribution.

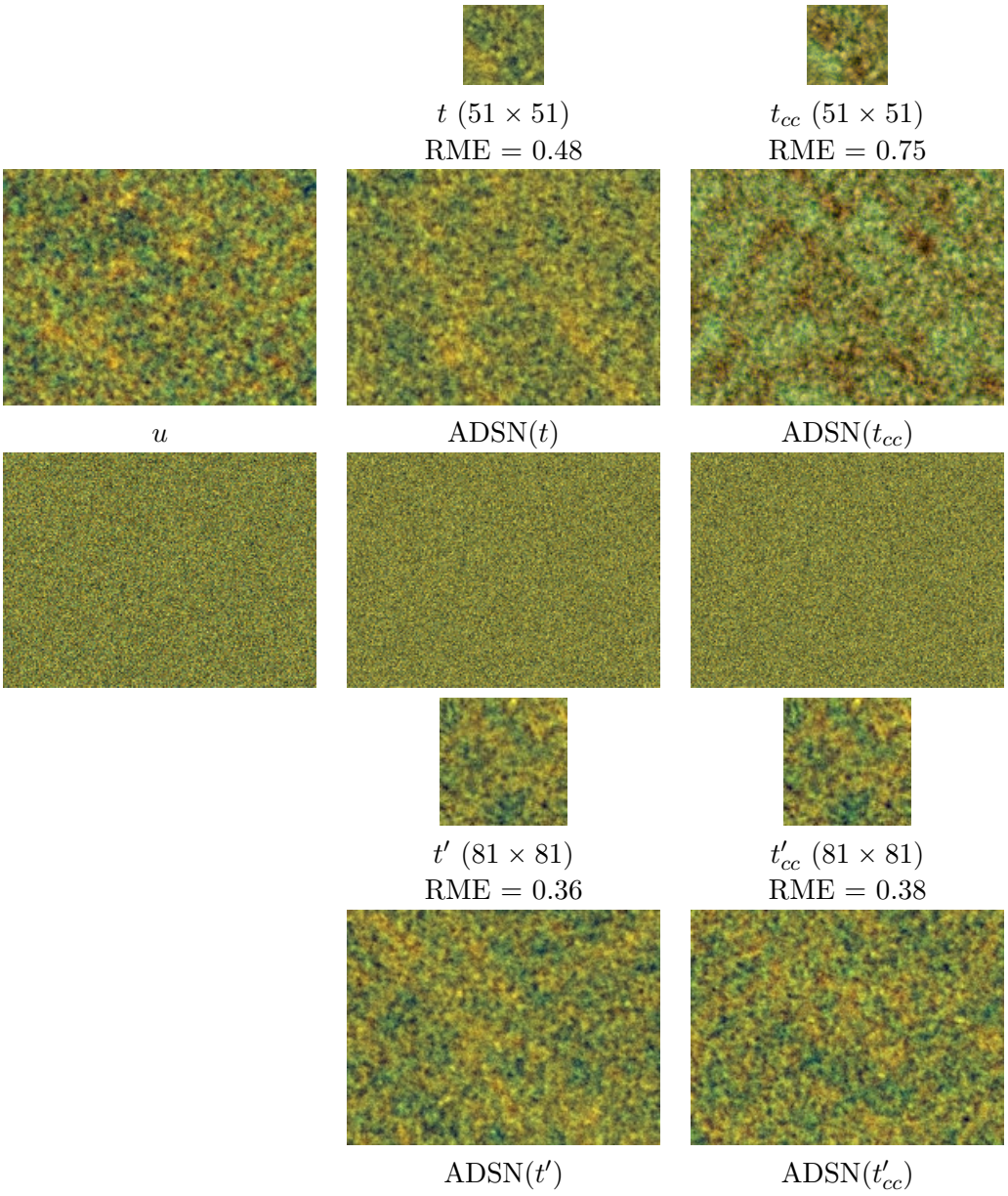


Figure 3.16: **Counter-example of color correction**. In the second row, one can see an original texture u and two ADSN synthesis results ADSN(t) and ADSN(t_{cc}) obtained with the SOT t and the color corrected SOT t_{cc} shown in the first row. The images of the third row are shuffled versions of the ones of the second row (meaning that we applied on these three image the same random permutation of the pixels). In the fourth row, we also show a larger-support SOT with the color corrected version, and in the last row, we show the corresponding ADSN realizations. On this example, the color correction (explained in Equation (3.18)) applied to the SOT does not suffice to get back the perceived colors of the original texture. Even in the Gaussian case, the equality of the marginal color distributions of u and ADSN(t_{cc}) is not sufficient to ensure that a human will perceive the same colors. One can also observe that this problem of the color correction does not appear anymore with a larger SOT support.

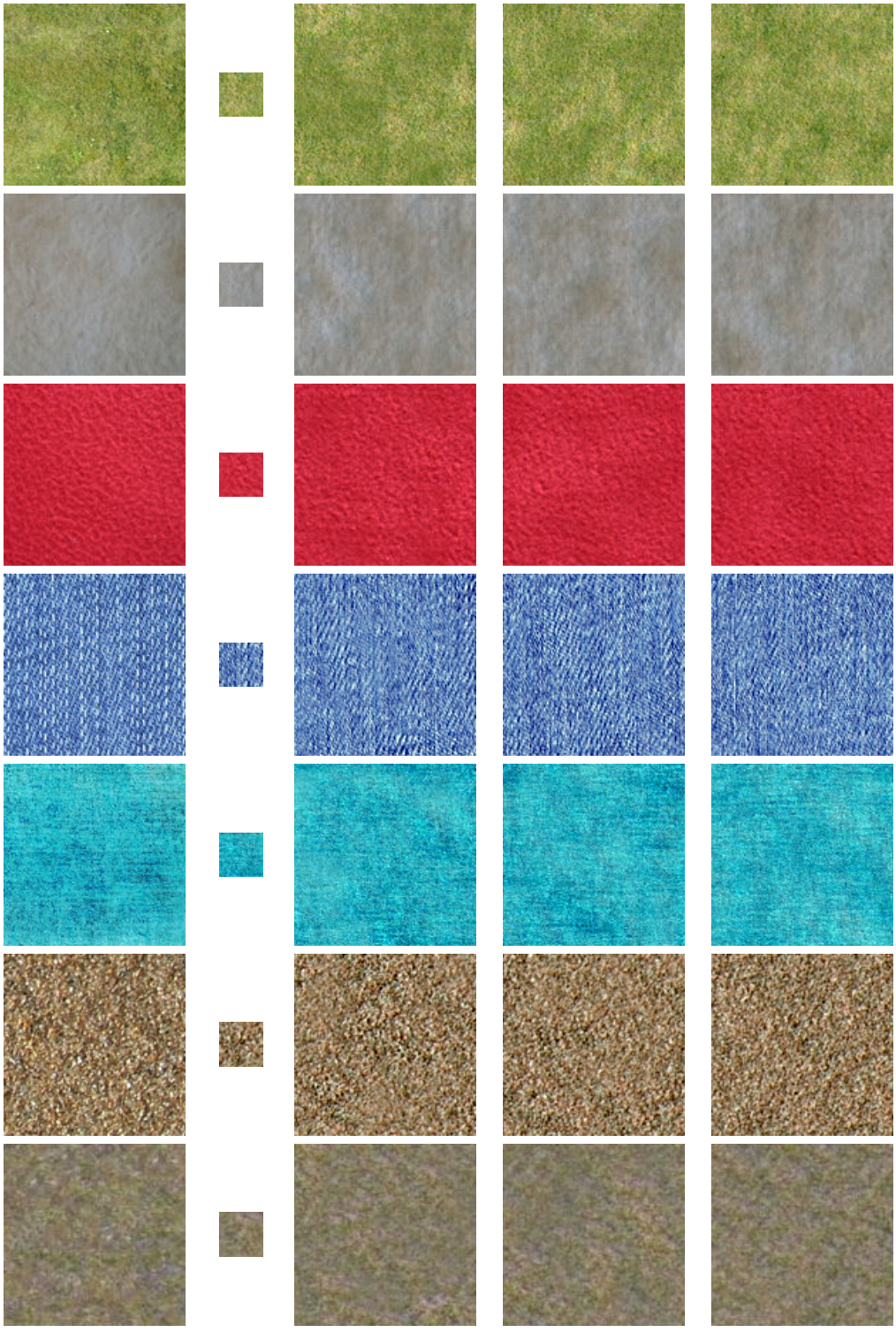


Figure 3.17: **Examples of DSN texture synthesis with the SOT.** For each row and from left to right : natural texture u, SOT with square support of size 31×31, results of DSN synthesis using t with respectively 10 and 30 impacts per pixel, and ADSN synthesis with t. The SOTs presented here were computed with the color correction. As one can see, many natural microtextures can be efficiently and faithfully reproduced with a DSN synthesis associated to a small SOT and with a very low number of impacts per pixel.

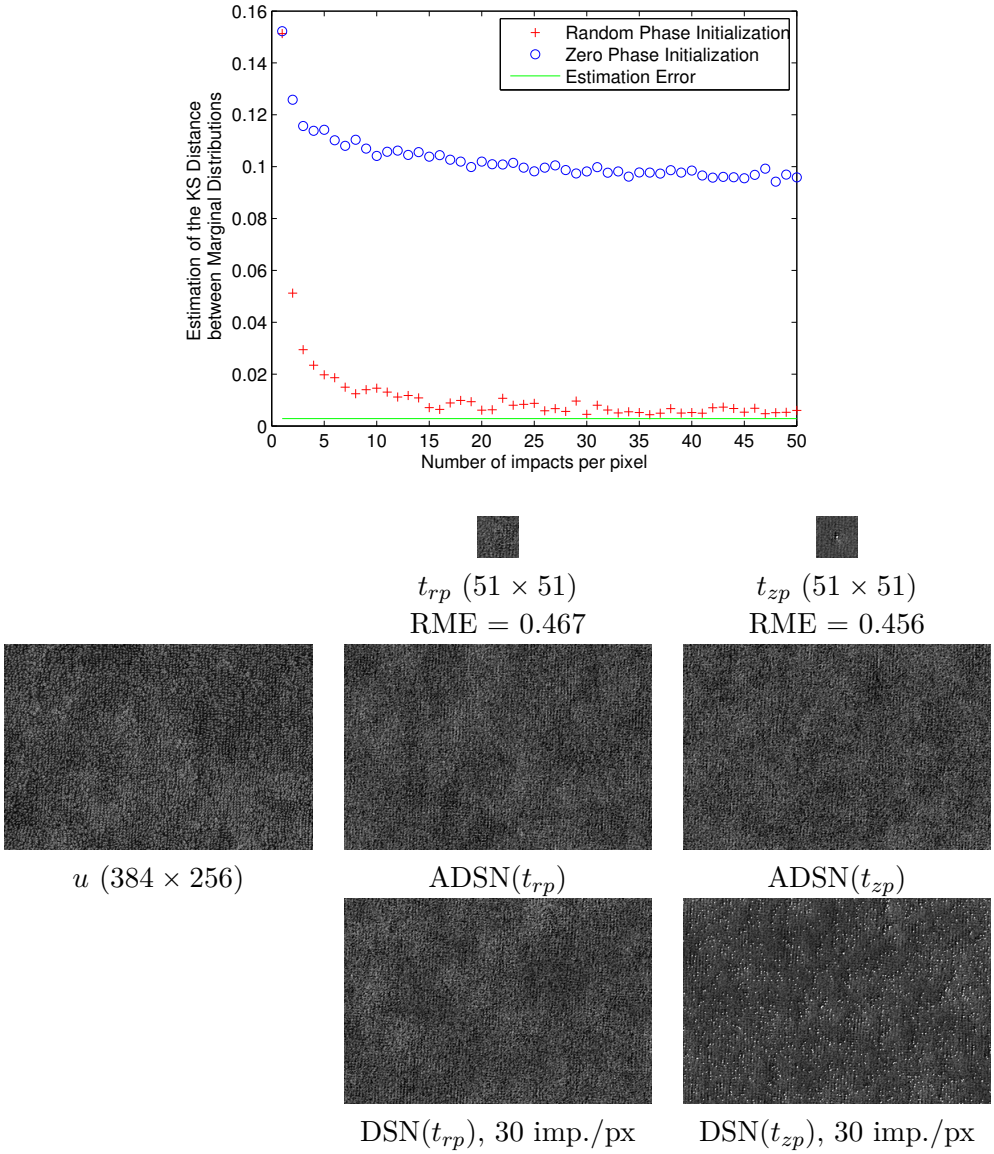


Figure 3.18: **Importance of the random-phase initialization.** This figure allows to compare the random-phase and zero-phase initializations of the alternating projections algorithm; the corresponding output textons are denoted by t_{rp} and t_{zp}. The upper part of the figure contains a diagram plotting the Kolmogorov-Smirnov (KS) distance between the marginal distribution of the DSN with intensity λ and the ADSN associated to t_{rp} and t_{zp} (the x-axis refers to the number of impacts per pixel, which is proportional to the intensity). Precisely, since this KS distance is not explicitely computable, we estimated it by the KS distance between the empirical cumulative disribution function of 10^5 marginal samples of the DSN and its Gaussian limit. The minimal estimation error (that is, the KS distance between a normal distribution and its empirical counterpart computed with 10^5 samples) is plotted as a green line on the diagram. On the second part of the figure, one can see the original texture u and DSN and ADSN synthesis results obtained with t_{rp} and t_{zp}. See the text for additional comments.

canonical/luminance texton. In Fig. 3.19 one can see the resulting ADSN obtained with several sizes of SOTs; this figure also shows that in terms of the resulting textures, the ADSN obtained with the SOT of support S or the canonical/luminance texton cropped on S are comparable. In order to find the support size that is suited to an exemplar texture, we suggest to compute the SOT with different sizes and to chose the smallest one that allows to retrieve in the synthesis the perceptucal characteristics of the original texture. It may be possible to set the support size automatically using a threshold on the RME, or using a more precise analysis of the dependency range in the exemplar texture (see [Costantini *et al.* 2004] for such a measure of spatial dependency).

3.3.3 Comments on the optimal transport distance

In the last subsection, we have seen that after the color correction of the SOT, the relative model error (measured by the optimal transport distance) increases while the visual approximation of the original texture by the ADSN gets better. This illustrates that the optimal transport distance does not suffice to measure precisely the visual proximity between two texture models. In this subsection, we show that this fact is supported by another argument: the equal contribution of all Fourier frequencies in the raw optimal transport distance does not comply well with our texture perception.

For that, let us comment the peculiar example of SOT presented in Fig. 3.20. The exemplar texture is a synthetic "sky" texture with an isotropic aspect. We computed the SOTs associated to this texture for rectangular and circular supports of different sizes. In order to assess the quality of the Gaussian model approximation realized by each SOT, one has to compare the corresponding ADSN synthesis result with the initial texture.

Since the original texture of Fig. 3.20 has a slowly-decreasing spatial covariance, the SOT must be relatively large in order to get a satisfying synthesis. However, for a square SOT of size 61×61 (that is, for $r = 30$), the synthesis seems satisfactory, but on closer inspection, one can observe horizontal and vertical artifacts. These artifacts are due to the discontinuity of the SOT along the boundary of its support. Indeed, since the original texture has more long-range variations than small scale details (grain), the alternating projection algorithm fills the available support with a cloudy pattern which has *a priori* no reason to go to zero at the support boundary. These horizontal and vertical artifacts are in fact comparable to the patterns encountered in the ADSN associated to the indicator function of the square support, which confirms the previous analysis. Notice that these artifacts are less visible when r is further increased.

An interesting comment can be drawn from the observation of the corresponding relative model errors. Notice that in Fig. 3.20, the RME obtained with $t_{30}^{\square}$ is less than the one obtained with t_{30}°. This can be easily explained in terms of optimization problem: these two textons minimize the same objective function, but $t_{30}^{\square}$ is less constrained than t_{30}° because the circular support of radius r is included in the square support of size $(2r + 1) \times (2r + 1)$. But this becomes counter-intuitive

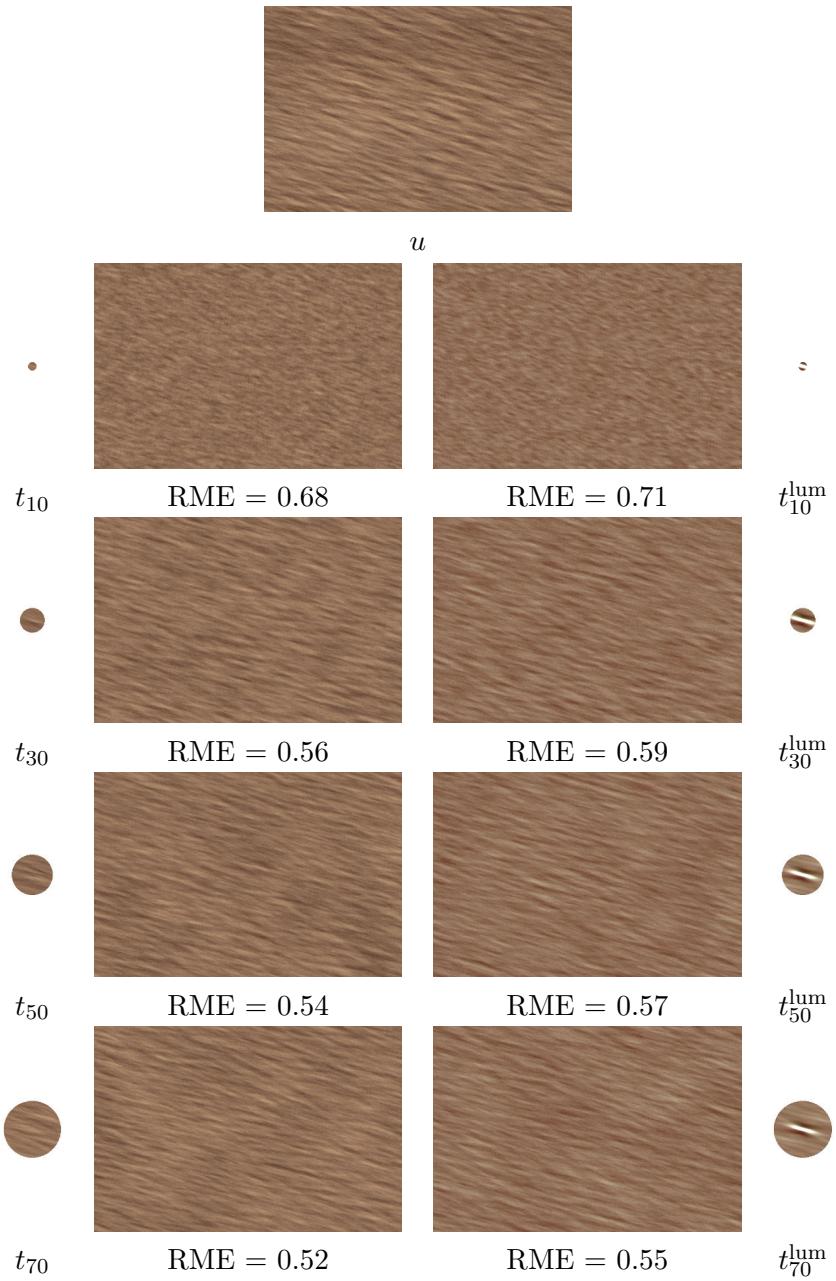


Figure 3.19: **Influence of the support size.** A Gaussian texture u of size 768×512 is displayed in the first row. The following rows correspond to the values $r = 10, 30, 50, 70$ and are organized as follows: the left column contains the SOT t_r computed with a circular support of radius r, the right column contains the luminance texton t_r^{lum} cropped with the same support, and the two middle columns contain samples of the ADSN models associated to the adjacent textons. As expected, increasing r allows larger range dependencies in the resulting ADSN field. Notice that the resulting textures associated to the SOT or the cropped luminance texton with same support are comparable.

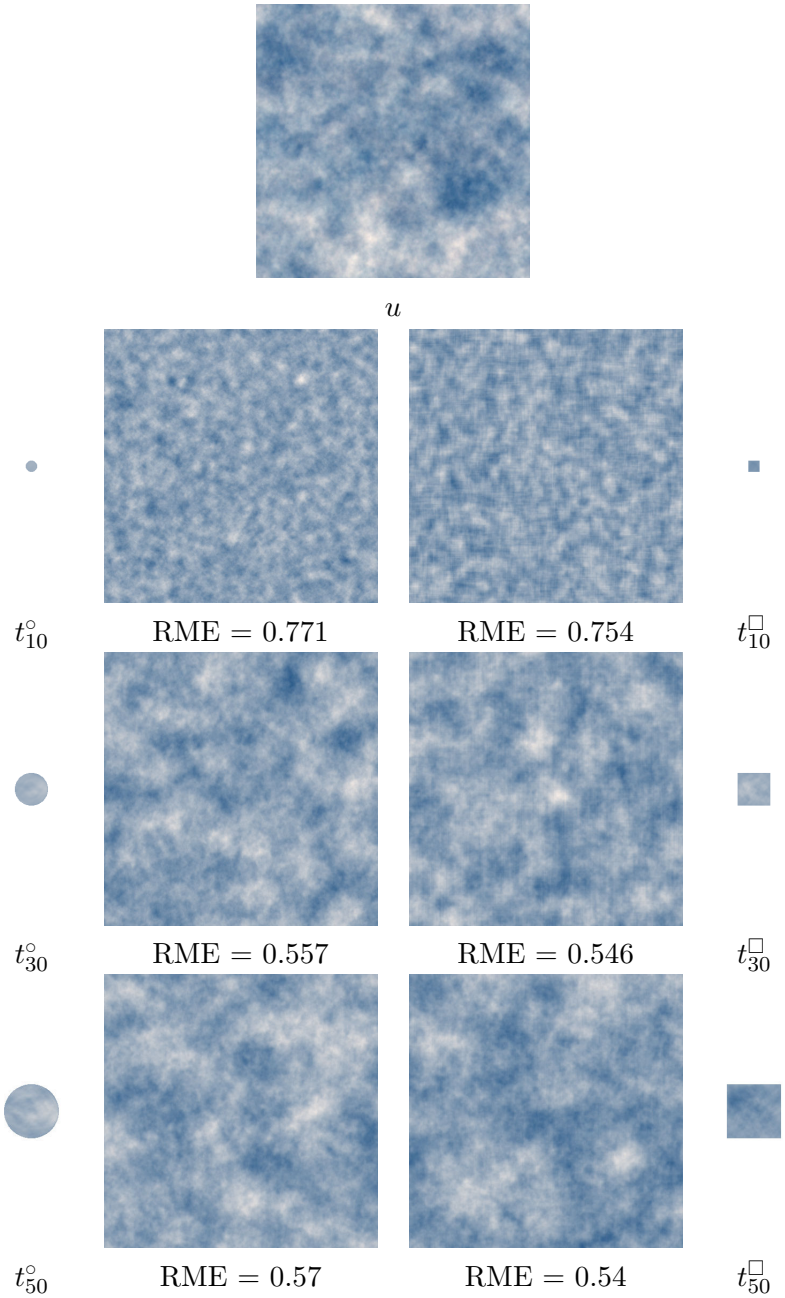


Figure 3.20: **Possible high-frequency artifacts with the raw SOT.** A Gaussian texture u of size 512×512 is displayed in the first row. The following rows correspond respectively to the values $r = 10$, 30, 50 and are organized as follows: the left column contains the SOT t°_r computed with a circular support of radius r, the right column contains the SOT $t^{\square}_r$ computed with a square support of size $(2r+1) \times (2r+1)$, and the two middle columns contain samples of the ADSN models associated to the adjacent textons. Notice that in the case of the square SOT $t^{\square}_{30}$, horizontal and vertical artifacts appear in the resulting ADSN which make the synthesis not satisfactory. In contrast, these artifacts do not appear in the circular SOT with same radius, even if the corresponding RME is lower.

when observing the corresponding ADSN realizations: the ADSN associated to t_{30}° indeed realizes a better approximation of the original texture (because the ADSN associated to $t_{30}^{\square}$ suffers from the uppermentioned horizontal and vertical artifacts).

This example shows that the model error (and thus the optimal transport distance on which it is based) does not suffice to faithfully measure the visual proximity between two Gaussian textures. A major drawback of this optimal transport distance is that its computation does not take into account the fact that the human textural perception is more sensitive to high frequencies than low frequencies. Indeed, in the expression (2.11) of the optimal transport distance between circular ADSN fields, all the frequencies appear with the same weight. This explains why the optimal transport distance is not able to discriminate the directional artifacts that can be encountered with the SOT of square support. In Subsection 3.3.4, we propose a variant of the optimal transport distance that includes frequency weights.

To confirm that the optimal transport distance is not a perfect reflection of the perceptual similarity between two textures, we suggest to study the variation of the relative model error between a fixed Gaussian texture u, and a realization of the ADSN associated to u. More precisely, let us fix a gray-level Gaussian texture $u : \Omega \longrightarrow \mathbb{R}$ of size $M \times N$ and recall the notation $t_u = \frac{1}{\sqrt{MN}}(u - \bar{u})$. Let us consider the circular Gaussian random field $U = \bar{u} + t_u * W$ associated to u (W is a normalized Gaussian white noise on Ω) and its normalized version $T = \frac{1}{\sqrt{MN}}(U - \bar{u})$ (notice that $\bar{U} = \bar{u}$ because t_u has zero mean). Then, we propose to study the random variable

$$\mathrm{RME}(T, t_u)^2 = \frac{\sum_{\boldsymbol{\xi}} \left(|\widehat{t}_u(\boldsymbol{\xi})| - |\widehat{T}(\boldsymbol{\xi})| \right)^2}{\sum_{\boldsymbol{\xi}} |\widehat{t}_u(\boldsymbol{\xi})|^2} \, . \tag{3.19}$$

It can be rewritten

$$\mathrm{RME}(T, t_u)^2 = \frac{\sum_{\boldsymbol{\xi} \neq 0} |\widehat{t}_u(\boldsymbol{\xi})|^2 \left(1 - \frac{|\widehat{W}(\boldsymbol{\xi})|}{\sqrt{MN}} \right)^2}{\sum_{\boldsymbol{\xi} \neq 0} |\widehat{t}_u(\boldsymbol{\xi})|^2} \, .$$

Since the $\frac{|\widehat{W}(\boldsymbol{\xi})|^2}{MN}$ follow the exponential distribution of parameter 1, one has

$$\mathbb{E}\Big(\mathrm{RME}(T, t_u)^2 \Big) = 2 - \sqrt{\pi} \approx 0.23 \, .$$

Therefore, in some sense, this value corresponds to the best we can hope when approximating the Gaussian model associated to t_u. It has to be compared with the RME^2 values obtained when approximating a Gaussian texture by the ADSN associated to the SOT. For example, in Fig. 3.20, the approximation of the sky texture by the ADSN associated to $t_{30}^{\square}$ leads to a RME^2 value of ≈ 0.3; and in Fig. 3.16 the approximation leads to a RME^2 value of ≈ 0.26. These values are just slightly over $2 - \sqrt{\pi}$ even if the approximated textures are not perfectly similar to the original in both these cases.

In conclusion, the relative model error (3.17) (which is based on the optimal transport distance (2.11)) does not suffice to precisely assess the perceptual similarity between two texture samples.

3.3.4 Optimal transport distance with frequency weights

In this subsection, we propose to incorporate in the optimal transport distance given by (2.11) a frequency weight that will give more importance to high frequencies, thus reflecting the human textural sensitivity in a more faithful manner. We validate this approach by showing that, at the cost of extra parameters, this variant leads to a more precise SOT, and in particular avoids the horizontal and vertical synthesis artifacts presented in Subsection 3.3.3.

Definition 3.3.1. Let $h_0, h_1 : \Theta \longrightarrow \mathbb{R}^d$ be two kernel functions defined on a circular domain Θ and let μ_0, μ_1 be the distributions of the circular ADSN $h_0 * W$ and $h_1 * W$ where W is a Gaussian white noise on Θ of variance 1. Let also $w : \Theta \longrightarrow]0, +\infty[$ be a weighting function defined in the frequency domain. The weighted L^2-optimal transport distance between μ_0 and μ_1 is defined by

$$d^w_{\mathrm{OT}}(\mu_0, \mu_1)^2 = \frac{1}{|\Theta|} \sum_{\boldsymbol{\xi} \in \Theta} w(\boldsymbol{\xi})^2 \Big(\|\hat{h}_0(\boldsymbol{\xi})\|^2 + \|\hat{h}_1(\boldsymbol{\xi})\|^2 - 2|\hat{h}_0(\boldsymbol{\xi})^* \hat{h}_1(\boldsymbol{\xi})| \Big) . \tag{3.20}$$

The corresponding weighted relative model error is then defined by

$$\mathrm{RME}^w(\mu_1, \mu_0)^2 = \frac{\sum_{\boldsymbol{\xi}} w(\boldsymbol{\xi})^2 \Big(\|\hat{h}_0(\boldsymbol{\xi})\|^2 + \|\hat{h}_1(\boldsymbol{\xi})\|^2 - |\hat{h}_0(\boldsymbol{\xi})^* \hat{h}_1(\boldsymbol{\xi})| \Big)}{\sum_{\boldsymbol{\xi}} w(\boldsymbol{\xi})^2 \|\hat{h}_0(\boldsymbol{\xi})\|^2} . \tag{3.21}$$

If k is the inverse DFT of w, computing d^w_{OT} amounts to compute the usual optimal transport distance between the ADSN associated to $k * h_0$ and $k * h_1$; thus the weighted optimal transport distance can be understood as an optimal transport distance on a filtered version of the random fields.

Of course, one has to make a choice for the function w that is compliant with the human texture perception. We suggest to take a weight that is proportional to the power of the normalized frequency

$$w(\xi_1, \xi_2) = \beta \left(2\Big(\frac{\xi_1}{M}\Big)^2 + 2\Big(\frac{\xi_2}{N}\Big)^2 \right)^{\alpha/2} ,$$

where $\alpha, \beta > 0$. The parameter β has no importance for now, and it is only useful to normalize the values of w between 0 and 1. The parameter α must be set in order to reflect the human sensitivity to high-frequency; thus, a very precise setting of α would require a thorough perceptual study of our frequency-dependent perception of textures. Here, to keep things simple, we suggest to take $\alpha = 0.9$ in order to counterbalance the power-law encountered in the spectrum of natural images [Ruderman 1994].

Now that this frequency-weighted optimal transport distance is defined, the question is: "how can it be embedded" in the SOT computation to make it more

precise. This question is not trivial because the integration of the frequency weight does not change the model projection. Indeed, whatever be the weighting function w, one solution of

$$\operatorname*{Argmin}_{t:\Omega \longrightarrow \mathbb{R}^d} \quad d_{OT}^w(\mu_t, \mu_{t_u}) \tag{3.22}$$

(where μ_t is the circular ADSN on Ω associated to the kernel t) is still given in Fourier domain by

$$\widehat{p_{t_u}(h)} = \frac{\hat{t}_u \hat{t}_u^* \hat{t}}{|\hat{t}_u^* \hat{t}|} \mathbf{1}_{\hat{t}_u^* \hat{t} \neq 0} + \hat{t}_u \mathbf{1}_{\hat{t}_u^* \hat{t} = 0} \, . \tag{3.23}$$

The integration of the frequency weights thus does not change the alternating projections algorithm.

One possibility to take account of the weights in the algorithm is to replace the alternating projections by a gradient descent on the functional

$$F(t) = F_1(t) + F_2(t) = \frac{1}{2} d_{OT}^w(\mu_t, \mu_{t_u})^2 + \frac{1}{2} \sum_{\mathbf{x} \notin S} \|t(\mathbf{x})\|^2 \, .$$

Notice that we do not include a Lagrange multiplier because a multiplicative constant β is already included in the function w. One can see that F is differentiable almost everywhere. For almost every t (precisely, for each t such that (3.22) admits a unique solution), the gradient of the first part is given in Fourier domain by

$$\widehat{\nabla F_1(t)} = w \left(\hat{t} - \widehat{p_{t_u}(t)} \right) \, . \tag{3.24}$$

The gradient of the second part is simply given by

$$\nabla F_2(t) = t - t \mathbf{1}_S \, . \tag{3.25}$$

Notice that for the ℓ^2-norm, ∇F_2 is always 1-Lipschitz, and that ∇F_1 is 1-Lipschitz as soon as β is chosen to have $\|w\|_\infty \leqslant 1$.

The gradient descent on F leads to the computation of a texton depending on the frequency weight w, and denoted by SOT^w. The corresponding algorithm is summarized below. In practice, we oberve the convergence of the algorithm as soon as the gradient step ν is less than 1. In constrast with the alternating projection algorithm, the weighting function w appears in this new algorithm through the gradient of F_1.

Algorithm: SOT^w computation

- Initialization: $\widehat{t} \leftarrow \widehat{t}_u e^{i\psi}$ where ψ is a uniform random phase function, and $t_u = \frac{1}{\sqrt{|\Omega|}}(u - \bar{u})$.
- Repeat (n times) $t \leftarrow t - \nu(\nabla F_1(t) + \nabla F_2(t))$ (where ∇F_1 and ∇F_2 are given by (3.24) and (3.25)) .

In Fig. 3.21, we give an example of frequency-weighted SOT for the sky texture presented in Subsection 3.3.3. In this figure, one can see that the SOT^w is more precise than the raw SOT because the corresponding ADSN does not suffer from the directional artifacts encountered in Fig. 3.20. However, the approximation of the marginal distribution becomes worse with SOT^w, and thus it is very recommended to apply to SOT^w the color correction step explained in Subsection 3.3.2. Notice in particular that, in contrast to the original RME, the frequency-weighted RME penalizes the bad approximation of the high frequencies, even after the color correction step. In Fig. 3.22 we show several other examples of frequency-weighted SOT. These results confirm that the frequency-weighted SOT is in general better, even for more complex microtextures.

An interesting perspective to pursue this work would be to integrate in the objective function a distance between the color distribution of the DSN and the estimated color distribution of the exemplar. The setting of Lagrange parameters in this new optimization problem may be interesting because they would reflect the balance realized by human texture perception between sensitivity to the frequency content and sensitivity to the color distributions.

3.4 Conditional Simulation and Inpainting

In this section, we address the inpainting problem for images composed of one homogeneous microtexture using conditional sampling of an ADSN model.

The inpainting problem consists in filling unknown areas of an image based on the surrounding content. The general problem is very difficult and obviously ill-posed. Still, it has been addressed in the literature by a wide variety of methods. Here, we will not give an exhaustive overview on this subject, but let us still mention [Masnou & Morel 1998], [Bertalmio *et al.* 2000] (which both perform inpainting by level line completion), [Criminisi *et al.* 2004] (whose algorithm is based on a fastened version of the texture synthesis scheme of [Efros & Leung 1999]) and [Mairal *et al.* 2008] (which exploits patch sparsity in a learned dictionary). However, as observed in [Criminisi *et al.* 2004] and [Newson 2014, §6.6.2], many existing inpainting techniques are not able to restore the textural content in a faithful way because the filled content is already regularized in some way.

The inpainting problem becomes less difficult if one has prior knowledge on the type of content that has to be filled. For example, the authors of [Bertalmio *et al.* 2003] propose to separate the structural content and the textural content and to inpaint them with two different techniques. One can further simplify this problem by trying to inpaint an image that is composed only of a single homogeneous texture. This is still a chalenging problem which can be elegantly formulated in terms of conditional simulation. Indeed, since we are dealing with a texture image, it is natural to assume that the image is modelled by a stationary random field U on Ω with probability distribution P. Then, if one wants to inpaint U on the domain Ω_m (of missing values) from the known values $u_{|\Omega_k}$ on $\Omega_k = \Omega \backslash \Omega_m$, then one can draw $U_{|\Omega_m}$ by sampling from the probability distribution

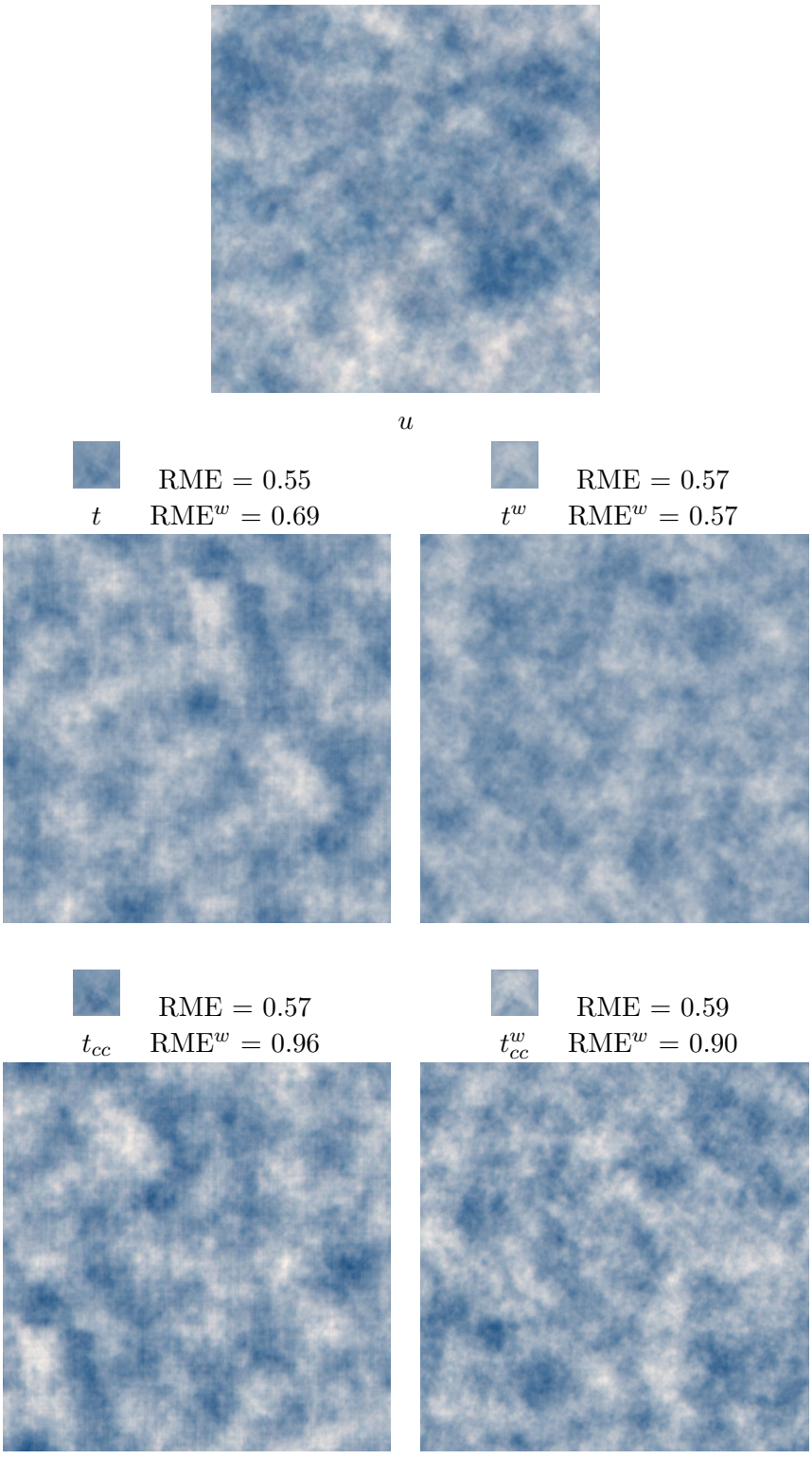


Figure 3.21: **SOT^w computed with frequency weight** w. An original texture u is shown in the upper part of the figure. The rest is divided in four parts showing a texton and a realization of the corresponding ADSN model together with the values of RME and RME^w. We first show the results with t (original SOT), t^w (SOT computed with frequency weights) and below, we show the results with the textons t_{cc}, t^w_{cc} obtained after the color correction explained in Subsection 3.3.2. See the text for further comments.

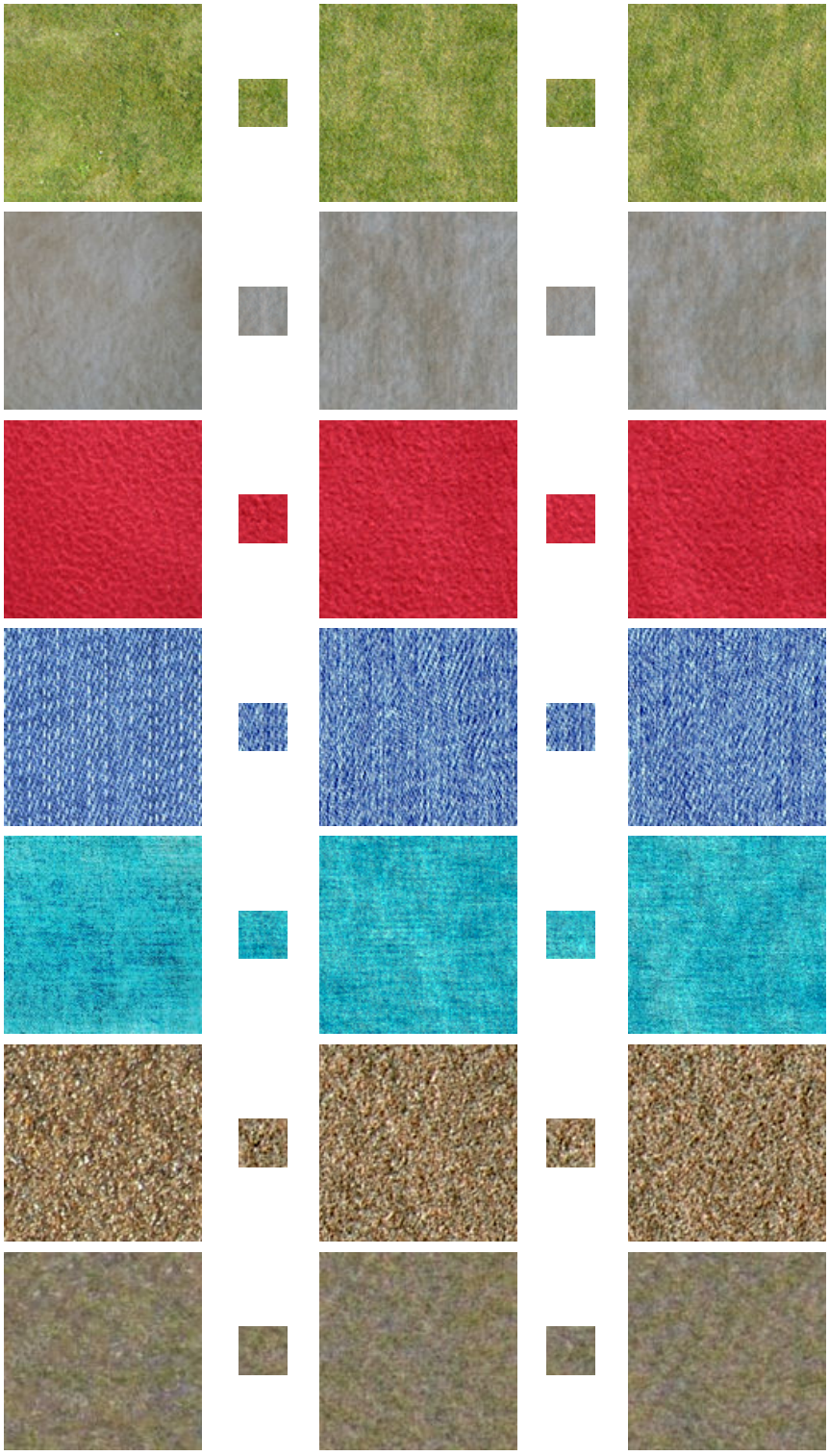


Figure 3.22: **Examples of weighted SOT.** In each row from left to right: original texture u, SOT obtained with frequency weights and color correction t_{cc}^{w}, the corresponding ADSN, SOT with color correction t_{cc} and the corresponding ADSN.

of U conditionally to $U_{|\Omega_k} = u_{|\Omega_k}$. Therefore, the global simulation scheme is consistent because, if $U_{|\Omega_k}$ is already drawn from the marginal distribution, then the inpainted image U will indeed follow the distribution P, and thus be a valid sample of the texture model.

So we are led to the problem of finding interesting texture models for which the conditional simulation can be performed in practice. Let us first notice that the random fields that are *a priori* more adapted to the conditional simulation are Markov random fields. Indeed, the distribution of a stationary Markov random field is constrained by the "local specification", which is the distribution of a pixel conditionned by the neighbooring values. For such a Markov random field, the distribution of $U_{|\Omega_m}$ conditionally to the rest only depends on the values $U_{|\partial\Omega_m}$ on the pixels that are neighboors of pixels of Ω. However, the conditional distribution

$$\mathcal{L}(U_{|\Omega_m} \mid U_{|\partial\Omega_m}) \ ,$$

may not be easy to explicit or sample. Still, one major contribution of the authors of [Efros & Leung 1999] was to show that the conditional simulation for one pixel could be performed with a patch-based sampling method (that is, sampling one pixel in the set of the exemplar values that have a similar neighborhood). Using this method, they obtained a textural inpainting result presented under the name "constrained texture synthesis" [Efros & Leung 1999, Fig.4]. In their example, one can notice that the inpainted region is quite small; indeed, the patch-based approximation of the conditional simulation is less precise on a wider domain (because the progressive filling algorithm will struggle to comply with the long-range constraints). To sum up, we mention that, even with a Markov random field model, it is not clear that the conditional simulation of a whole region can be done exactly and efficiently.

In this section, we present a conditional simulation scheme for Gaussian random fields, which can be used to address inpainting of Gaussian textures. As in [Lantuéjoul 2002], we present the Gaussian conditional simulation by relying on kriging estimates of the unknown values (which are the conditional expectations of the unknown values based on the unmasked pixels). This sampling algorithm is a perfect conditional simulation scheme and we demonstrate here that it is able to fill large holes in a Gaussian texture. When the set of conditioning points or the set of masked pixels becomes too large, the computational cost becomes prohibitive. However, a further assumption on the Gaussian field (respectively a Markov property or covariance compactness) allows to cope with this problem.

Let us also mention that the article [Jassim 2013] also proposes a texture inpainting technique based on kriging interpolation but there is no conditional sampling in their algorithm because the inpainted values are exactly given by the kriging estimates (which are deterministic functions of the unmasked values). This is why their algorithm only works for very thin masked regions of the exemplar, in contrast of ours that is able to inpaint very holes.

3.4.1 Kriging Estimates and Conditional Simulation

For the sake of simplicity, let us explain the conditional simulation for the case of gray-level images as in [Lantuéjoul 2002].

Let $U : \mathbb{Z}^2 \longrightarrow \mathbb{R}$ be a stationary Gaussian random field with mean m. We can assume that $m = 0$. We introduce a finite domain $\Omega \subset \mathbb{Z}^2$ and also the subset $\mathcal{C} \subset \Omega$ of conditioning points (i.e. the pixels whose values are initially available). We will denote by Γ the restriction to $\Omega \times \Omega$ of the covariance:

$$\forall \mathbf{x}, \mathbf{y} \in \Omega, \quad \Gamma(x, y) = \mathbb{E}(U(\mathbf{x})U(\mathbf{y})) \ .$$

We will derive the distribution of U conditionally to $U_{|\mathcal{C}}$. Since U is globally Gaussian, the conditional expectation

$$U^*(\mathbf{x}) = \mathbb{E}(\ U(\mathbf{x}) \mid U(\mathbf{c}) \ , \ \mathbf{c} \in \mathcal{C} \ )$$

is the L^2-orthogonal projection of $U(\mathbf{x})$ on the subspace spanned by the random variables $U(\mathbf{c})$, $\mathbf{c} \in \mathcal{C}$. So, for each $\mathbf{x} \in \Omega$, there exists $(\lambda_{\mathbf{c}}(\mathbf{x}))_{\mathbf{c}\in\mathcal{C}} \in \mathbb{R}^{\mathcal{C}}$ such that

$$U^*(\mathbf{x}) = \sum_{\mathbf{c}\in\mathcal{C}} \lambda_c(\mathbf{x})U(\mathbf{c}) \ . \tag{3.26}$$

This value $U^*(\mathbf{x})$ is sometimes called the kriging estimate of $U(\mathbf{x})$, and the values $(\lambda_c(\mathbf{x}))_{\mathbf{c}\in\mathcal{C}}$ are called the kriging coefficients.

Let us fix $\mathbf{x} \in \Omega$. The coefficients $(\lambda_{\mathbf{c}}(\mathbf{x}))_{\mathbf{c}\in\mathcal{C}}$ can be computed by solving a linear system. Indeed, by definition of the conditional expectation, we have

$$\forall \mathbf{d} \in \mathcal{C}, \quad \mathbb{E}(U^*(\mathbf{x})U(\mathbf{d})) = \mathbb{E}(U(\mathbf{x})U(\mathbf{d})) \ ,$$

and thus

$$\forall \mathbf{d} \in \mathcal{C}, \quad \sum_{\mathbf{c}\in\mathcal{C}} \lambda_{\mathbf{c}}(\mathbf{x})\Gamma(\mathbf{c}, \mathbf{d}) = \Gamma(\mathbf{x}, \mathbf{d}) \ . \tag{3.27}$$

Notice that this is a linear system of $|\mathcal{C}|$ equations and with $|\mathcal{C}|$ unknown variables whose matrix is exactly $(\Gamma(\mathbf{c}, \mathbf{d}))_{(\mathbf{c},\mathbf{d})\in\mathcal{C}\times\mathcal{C}}$. We have just shown that this system has indeed a solution given by $(\lambda_{\mathbf{c}}(\mathbf{x}))_{\mathbf{c}\in\mathcal{C}}$. Let us remark that its matrix is symmetrical and nonnegative (because Γ is a covariance). In particular cases (for example, if there exists $\mathbf{c}_1, \mathbf{c}_2 \in \mathcal{C}$ such that $U(\mathbf{c}_1) = U(\mathbf{c}_2)$, which may be the case for perfectly periodic patterns), this matrix may be noninvertible, and in that case, there may be more than one solution.

The next proposition explains why the conditional simulation is very simple with Gaussian random fields.

Proposition 3.4.1 ([Lantuéjoul 2002])**.** *The random vectors U^* and $U - U^*$ are independent.*

Proof. Since the random vector $(U, U-U^*)$ is Gaussian, it is enough to show that U and $U - U^*$ are uncorrelated, that is for every $\mathbf{x}, \mathbf{y} \in \Omega$,

$$\mathbb{E}\Big(U^*(\mathbf{x})(U(\mathbf{y}) - U^*(\mathbf{y}))\Big) = 0 \ .$$

But this derives from the definition of the conditional expectation: $U^*(\mathbf{x})$ belongs to the subspace

$$\mathrm{Span}(\{\ U(\mathbf{c})\ ,\ \mathbf{c} \in \mathcal{C}\ \})$$

and $U(\mathbf{y}) - U^*(\mathbf{y})$ is orthogonal to that same subspace. □

Let us emphasize that U^* only depends on the values $U(\mathbf{c}), \mathbf{c} \in C$, and that $U - U^*$ vanishes on every point of $\mathcal{C}$. Therefore, if V follows the same distribution as U and is independent of U, then

$$U^* + V - V^*$$

follows the same distribution as U and takes the same values than U on $\mathcal{C}$. Another way to put this is to say that if we have prescribed values $u_{|\mathcal{C}}$ on $\mathcal{C}$, then conditionally to $U_{|\mathcal{C}} = u_{|\mathcal{C}}$, U follows the same distribution as

$$u^* + V - V^*$$

where

$$u^*(\mathbf{x}) = \sum_{\mathbf{c}\in\mathcal{C}} \lambda_c(\mathbf{x}) u(\mathbf{c})\ , \tag{3.28}$$

and where V follows the same distribution as U. This gives a straightforward simulation scheme for U with conditioning points in $\mathcal{C}$, once the coefficients $(\lambda_{\mathbf{c}}(\mathbf{x}))_{\mathbf{x}\in\mathcal{C}}$ have been computed.

Let us give a brief comment about the computation of the kriging coefficients. Notice that the matrix $M = (\Gamma(\mathbf{c}, \mathbf{d}))_{(\mathbf{c},\mathbf{d})\in\mathcal{C}\times\mathcal{C}}$ associated to the linear system (3.27) does not depend on $\mathbf{x}$. Therefore, it is interesting to invert the matrix M once and for all. The inversion of the matrix M of size $|\mathcal{C}| \times |\mathcal{C}|$ can be done with standard numerical techniques in $\mathcal{O}(|\mathcal{C}|^3)$. Then, in order to obtain the kriging coefficients for all $\mathbf{x}$ in a subset $\omega \subset \Omega$, one must compute $|\omega|$ products of M with a column vector, so the complexity of this step is $\mathcal{O}(|\omega||\mathcal{C}|^2)$. For that reason, the number of conditioning points must stay low for the conditional simulation to be feasible.

Notice also that when the covariance function of U has compact support K, then we will have $\Gamma(\mathbf{x}, \mathbf{d}) = 0$ as soon as $\mathbf{x} - \mathbf{d} \notin K$. Therefore, for $\mathbf{x} \notin K + \mathcal{C}$, the right-hand side of the linear system (3.27) is zero and thus we can choose $\lambda_{\mathbf{c}}(\mathbf{x}) = 0$ for all $\mathbf{c} \in \mathcal{C}$. In other words, for such a point $\mathbf{x}$, the kriging component is zero and the pixel value can be drawn independently (which agrees with the fact that $U(\mathbf{x})$ and $U(\mathbf{d})$ are independent). This case happens in particular when the Gaussian model derives from a compactly-supported texton (like the synthesis-oriented texton presented in Section 3.3). Therefore, if the Gaussian model has a covariance function with compact support K, then the kriging coefficients $(\lambda_{\mathbf{c}}(\mathbf{x}))$ must only be computed for the neighborhood $K + \mathcal{C}$ of the conditioning points $\mathcal{C}$.

3.4.2 Inpainting Results

In this subsection, we provide some experimental results which confirm that Gaussian conditional sampling can be used to perform microtexture inpainting.

As a proof of concept, we first suggest to inpaint a Gaussian texture with a previously learnt Gaussian model. Let us recall the notation Ω_m (resp. $\Omega_k = \Omega \backslash \Omega_m$) which refers to the masked pixels (resp. to the initially known pixels). For a Gaussian model μ on the domain Ω learnt on an exemplar texture, we propose to draw a realization u of μ on the domain Ω, to lose the values of the pixels of Ω_m, and to complete the lost values by conditional simulation leading to a new image v which coincides with u on Ω_k (actually, the process is equivalent to resynthesize the texture on Ω_m). In order to get a perfect global simulation scheme, we would have to perform the simulation on Ω_m with all the conditional points in Ω_k. But as we have said, the method becomes computationnally infeasible when there are too many conditioning points. Therefore, we suppose that Ω_m is a rectangular subdomain of Ω, and that the set $\mathcal{C}$ of conditioning points consists only of the bordering pixels of Ω_m (in the experiments, the thickness of the border was set to 5 pixels). Notice that if the random field has the Markov property, then the simulation is still perfect even if we restrict the conditioning points to be on the border of Ω_m.

The results of this validation experiment are shown in Fig. 3.23 and Fig. 3.24. One can see on both these examples that the inpainted version is a texture that is as plausible as the original one and that the border of the masked region is difficult to distinguish after the inpainting process.

The previous experiments show that the conditional simulation can be used to inpaint Gaussian textures. But we still have to explain how to learn a Gaussian model on a masked texture. The estimation of the mean value is easily restricted to the unmasked values. However, it is less trivial to adapt the covariance estimator. One possible way to do it is to modify the covariance estimator (3.1) so that it only takes account of the values of u on Ω_k. At first, one can thus think of the estimator

$$\mathbf{v} \mapsto \frac{1}{|\Omega_k \cap (\Omega_k - \mathbf{v})|} \sum_{\mathbf{x} \in \Omega_k \cap (\Omega_k - \mathbf{v})} (u(\mathbf{x}+\mathbf{v}) - \bar{u})(u(\mathbf{x}) - \bar{u}) \ ,$$

but it has the same drawback than the renormalized nonperiodic estimator of the covariance: its DFT is not nonnegative and thus it is not the true covariance function of a Gaussian model (see the discussion at the end of Subsection 3.1.3). Instead, we propose to consider the very simple

$$c_{u,\Omega_k} = \frac{1}{|\Omega_k|^2} \sum_{\mathbf{x} \in \Omega_k \cap (\Omega_k - \mathbf{v})} (u(\mathbf{x}+\mathbf{v}) - \bar{u})(u(\mathbf{x}) - \bar{u}) \ . \tag{3.29}$$

The normalizing factor is adjusted so that $c_{u,\Omega_k}(0)$ is the natural estimator of the marginal variance. This adaptation appears quite brutal at first, but we oberved in practice that for simple masks Ω_m, it defines a Gaussian model that will not be too far from the Gaussian model associated to the whole texture. It would be interesting to seek some conditions on the domain Ω_k that would entail a statistical validation in some sense.

We can now illustrate the Gaussian texture inpainting by using a covariance estimation based only on $u_{|\Omega_k}$. The results are shown in Fig. 3.25 and Fig. 3.26. On both these results we have inpainted a texture image of size 256×256 with

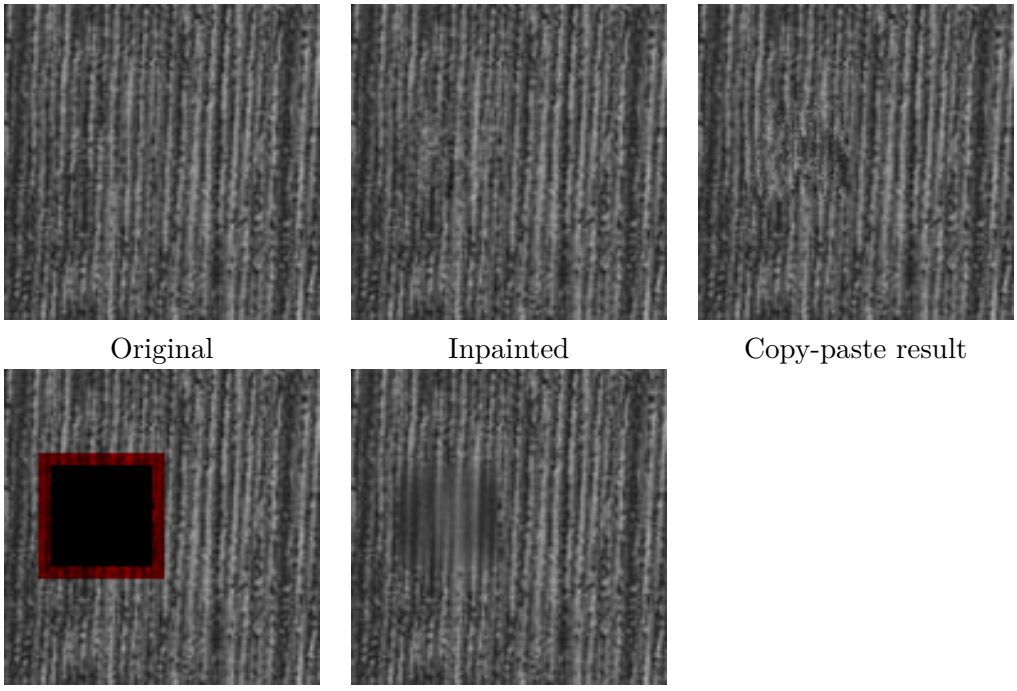


Figure 3.23: **Conditional simulation**. First column: original texture u (of size 128×128) and the masked version with the conditioning points colored in red (the mask has size 41×41). Second column: inpainted texture v with conditional simulation (obtained in 0.2 second), and kriging component u^*. Third column: inpainted texture obtained with a copy-paste method inspired by [Wei & Levoy 2000] (with a nearest-neighbor search in an ADSN realization of u). Notice that the algorithm is able to restore the linear structures that must be continued through the masked region. Actually, these structures are already visible in the kriging component; the conditional sampling adds the texture grain in a way that respects the texture covariance (which is not the case with the copy-paste method).

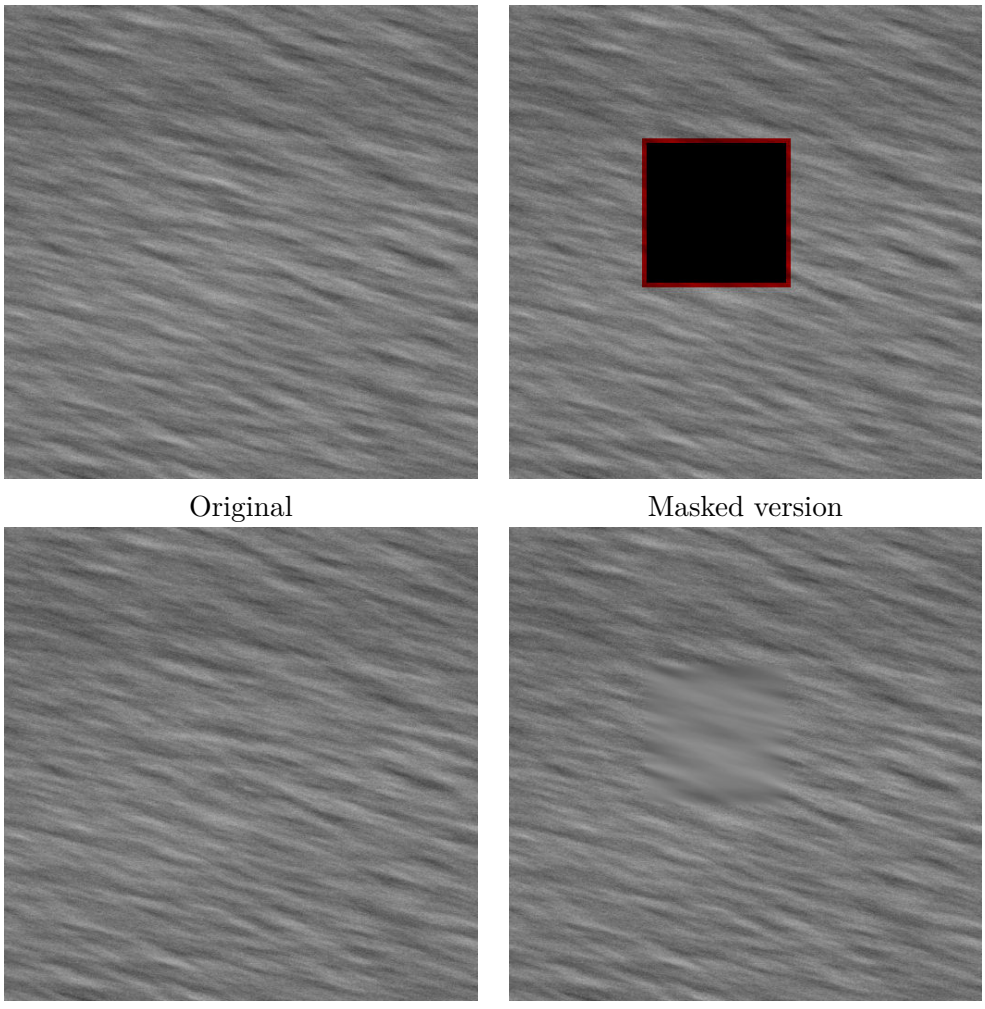


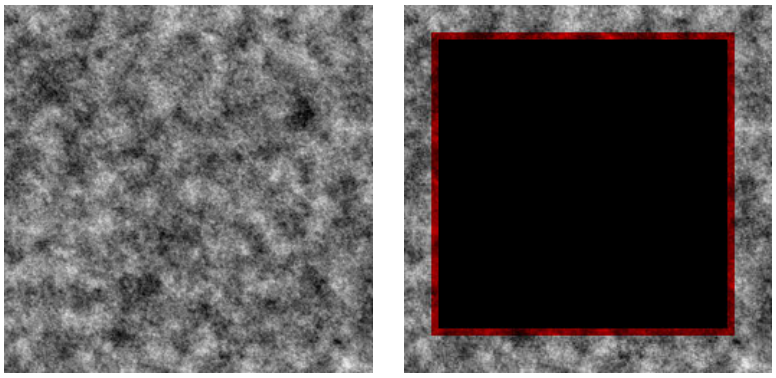
Figure 3.24: **Conditional simulation**. First row: original texture u (of size 512×512) and the masked version with the conditioning points colored in red (the mask has size 151×151). Second row: inpainted texture v (obtained in 18 seconds), and kriging component u^*.

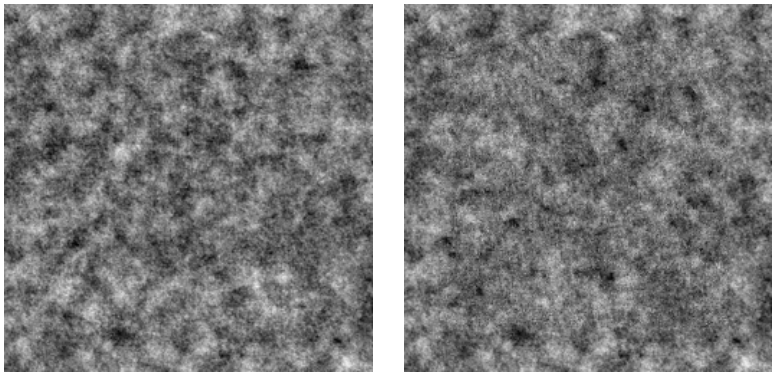
a square hole of size 201×201. Notice that in Fig. 3.25, the inpainting result is satisfactory: one can see on the ADSN samples (associated to the usual covariance estimator, and to the one obtained on the masked version of u) that the estimation procedure produces a satisfying result, even if the mask is very large compared to the whole domain. However, in Fig. 3.26, the estimation procedure fails because the available pixels are located on a border of the domain that is too thin to represent the relevant features of this texture.

In Fig. 3.23 and Fig. 3.25, we propose to compare the conditional simulation with a copy-paste method inspired by [Efros & Leung 1999] and [Wei & Levoy 2000]. Precisely, in a raster scan order, we sample each unknown value with a nearest-neighbor search in an ADSN realization of the non-masked original texture u. We copy the value of the pixel that has the best patch similarity based on the available neighboring pixels. One can see that the copy-paste result is not satisfactory for the example of Fig. 3.23 since one can clearly distinguish the inpainted zone. For the example of Fig. 3.25, the result is quite convincing but on closer inspection, one can see that the texture grain is not well preserved (because it is noisier). Since the copy-paste method is very dependent on the pixel scan order, it may not be able to reconstruct linear strutures crossing the masked zone, unless it is applied in a multi-pass manner (as suggested by [Wei & Levoy 2000, Fig.12]). Actually, these copy-paste methods are certainly more adapted to texture extrapolation than to texture inpainting.

Of course, the success of the patch-based methods for texture synthesis cannot be denied, and in particular they are more relevant than random phase models for a wide variety of structured textures. Still, the inpainting algorithm obtained by Gaussian conditional simulation illustrates once again that the mathematical flexibility of the Gaussian model has undeniable practical applications.

In conclusion, we have seen in this section that Gaussian conditional simulation can be used to inpaint large holes in image regions which are composed of one single microtexture. Since this algorithm inherently respects the textural content, the inpainted area does not suffer from over-regularization. Beyond the limitations of the stationary Gaussian model, the main limitation of this algorithm is that the computation time grows quickly with the number of conditioning points. One possible way to cope with this problem would be to restrict the texture model to Gaussian Markov random fields, for which the kriging estimation can be done more efficiently, as shown in [Hartman & Hössjer 2008].





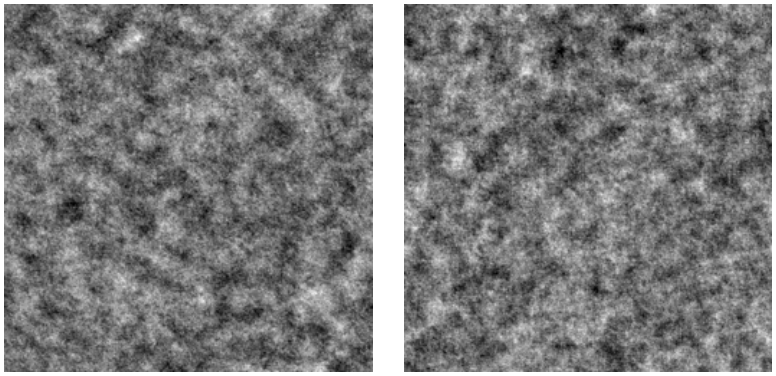


Figure 3.25: **Textural Inpainting**. First row: original texture u (of size 256×256) and the masked version with the conditioning points colored in red (the mask has size 201×201). Second row: inpainted texture v (obtained in 55 seconds), and result of the copy-paste method inspired by [Wei & Levoy 2000] (with a nearest-neighbor search in an ADSN realization of u). Third row: ADSN samples obtained with the usual covariance estimator (left), and with the covariance estimator (3.29) adapted for the masked version of u (right). Notice that in this example, the Gaussian model estimated on the masked version of u is still a convincing approximation of the texture u even if only less than 40% of the pixels are available.

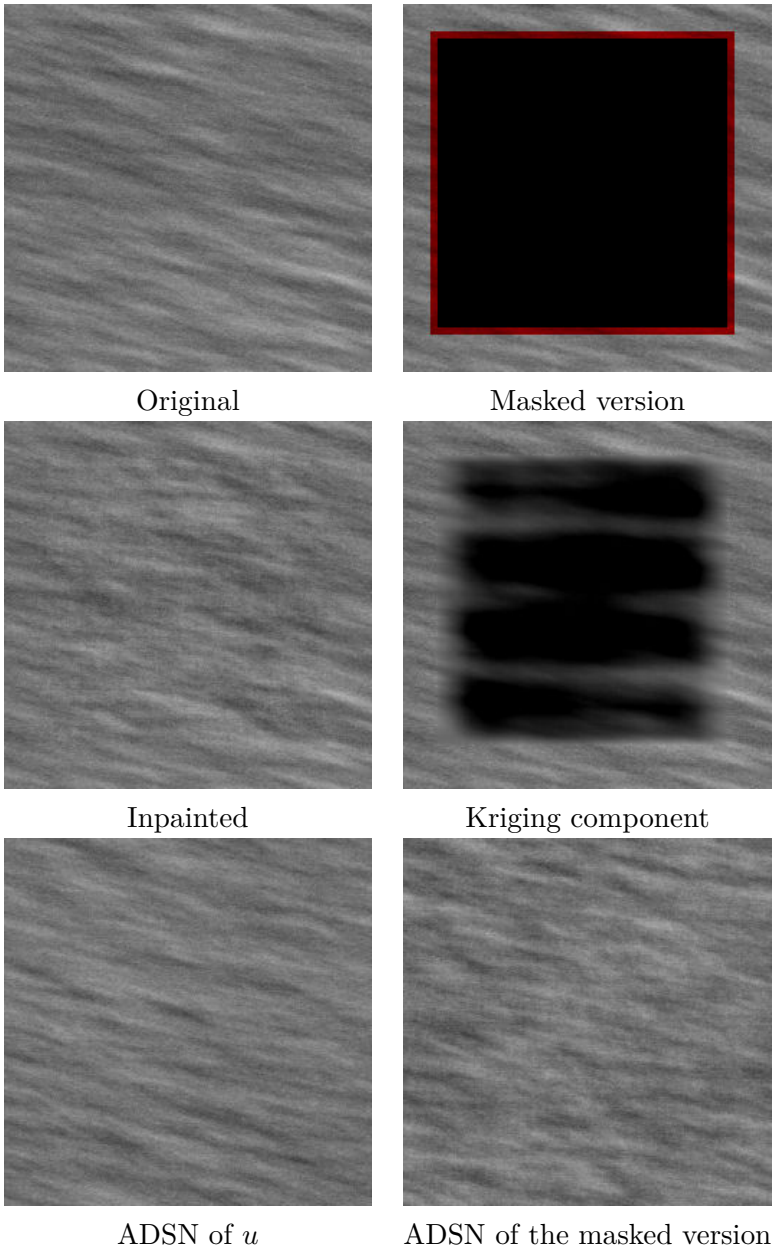


Figure 3.26: **Textural Inpainting**. First row: original texture u (of size 512×512) and the masked version with the conditioning points colored in red (the mask has size 151×151). Second row: inpainted texture v (obtained in 20 seconds), and kriging component u^*. Third row : ADSN samples with the usual covariance estimator (left), and with the covariance estimator (3.29) adapted for the masked version of u (right). In this example, the set of available pixels is to thin to allow for a correct estimation of the Gaussian model, thus making the inpainting result not satisfactory.

Chapter 4

No-reference image quality assessment and blind deblurring with sharpness metrics exploiting Fourier phase information

Contents

4.1 Introduction

In the two previous chapters, we studied random fields with uniform phase and saw how they could be used to perform by-example texture synthesis. Across some examples of texture synthesis, we were able to highlight the limits of the Gaussian model and in particular we saw that the uniform random phase models are not adapted to the case of images with sharp edges or salient features. Let us recall again that the authors of [Oppenheim & Lim 1981] already showed that the loss of the phase information of an image entails the destruction of the image geometry. This suggests that the precision of the image geometry (and thus, in some sense, the image quality) could be assessed through the coherence of the Fourier phase information.

Quality indices divide into three categories : full-reference, reduced-reference, and no-reference, depending on whether a supposedly ideal version of the image is assumed to be fully or partially known. As concerns the no-reference case (which is the one we are interested in), the introduction of Chapter 4 of [Wang & Bovik 2006] points out the difficulty to design generic image quality measures, concluding (in 2006) that *"the design of application-specific no-reference quality assessment systems appears to be much more approachable than the general, assumption-free no-reference image quality assessment problem."* Nevertheless, several interesting no-reference quality measures have been proposed in the literature (see the recent review [Chandler 2013]). Some of them try to assess the quality through the direct analysis of edges [Marziliano *et al.* 2004] or through the gradient singular values [Zhu & Milanfar 2010]. Others use a perceptual analysis of certain image features, like in [Ferzli & Karam 2009]. The concept of local phase coherence, originally introduced and developed in [Morrone & Burr 1988, Kovesi 2000, Kovesi 1999] for edge detection purposes, was later linked to the perception of blur by Wang and Simoncelli [Wang & Simoncelli 2004], which ultimately led to the definition of a no-reference image quality index [Hassen *et al.* 2010]. Closer to our work lies the index [Vu & Chandler 2009] which combines some spectral and spatial characteristics.

In 2008, a notion of global phase coherence was proposed [Blanchet *et al.* 2008], and related to image sharpness. The idea was to use a kind of *a contrario framework*[1] [Desolneux *et al.* 2008] to quantize how much the regularity of the image (more precisely, its total variation) was affected by the destruction of the phase information. This led to the definition of three phase coherence measures,

[1]The principle of *a contrario* methods is to detect structures as the cause of measurements that could not be observed in random data.

namely the Global Phase Coherence [Blanchet *et al.* 2008], the Sharpness Index [Blanchet & Moisan 2012], and the index S [Leclaire & Moisan 2013a]. It has been shown that these measures could be interpreted as quality indices because of the relation existing between the image geometry and the phase information (for example, degradation by blur or noise tends to reduce the coherence of the Fourier phase). The present chapter gives a more detailed and merged discussion about these global phase coherence indices. Starting from their construction in Section 4.2, we establish some of their mathematical properties in Section 4.3. Section 4.4 discusses several practical aspects of these indices, including their validation as no-reference quality measures, and finally Section 4.5 describes a way to use these indices to address the blind deblurring problem.

The content of this chapter is to appear in the *Journal of Mathematical Imaging and Vision*.

4.2 Three Phase Coherence Indices

This section presents the detailed construction of the phase coherence indices introduced in [Blanchet *et al.* 2008, Blanchet & Moisan 2012, Leclaire & Moisan 2013a].

4.2.1 Main notations

Let

$$\Omega = \mathbb{Z}^2 \cap \left(\left[-\frac{M}{2}, \frac{M}{2} \right) \times \left[-\frac{N}{2}, \frac{N}{2} \right) \right)$$

be a rectangular discrete domain of size $M \times N$. Let $u : \Omega \to \mathbb{R}$ be a discrete image, the real number $u(\mathbf{x})$ referring to the gray level at pixel $\mathbf{x}$. The Ω-periodization of u is the image $\dot{u} : \mathbb{Z}^2 \to \mathbb{R}$ defined by

$$\forall (k,l) \in \mathbb{Z}^2,\ \forall (x,y) \in \Omega,\ \dot{u}(x + kM, y + lN) = u(x,y)\ .$$

In the following, we will use a gradient scheme computed with periodic boundary conditions,

$$\nabla u(x,y) = \begin{pmatrix} \partial_x \dot{u}(x,y) \\ \partial_y \dot{u}(x,y) \end{pmatrix} = \begin{pmatrix} \dot{u}(x+1,y) - \dot{u}(x,y) \\ \dot{u}(x,y+1) - \dot{u}(x,y) \end{pmatrix}\ ,$$

and the corresponding (periodic) Total Variation (TV) of u

$$\mathrm{TV}(u) = \sum_{\mathbf{x} \in \Omega} |\partial_x \dot{u}(\mathbf{x})| + |\partial_y \dot{u}(\mathbf{x})|\ ,$$

which measures in some sense how much the function $\dot{u}$ oscillates. Precisely, the $TV(u)$ is the l^1-norm of the gradient of u, and thus it assigns small values (relatively to the l^2-norm) to images whose gradient is sparse (in particular cartoon images). Algorithms based on TV minimization have been used for a long time to address image processing tasks, for example, denoising [Rudin *et al.* 1992, Chambolle 2004].

Let us recall that the DFT is defined by (2.1), that the function $|\hat{u}|$ will be called modulus of u and that a phase function for u is any function $\varphi : \mathbb{Z}^2 \to \mathbb{R}$ such that

(a) *House* (b) *Lena* phase of (a) with modulus of (b)

Figure 4.1: **Phase and perceived geometric content.** When an image is built (in Fourier domain) with the phase of an image (a) and the modulus of an image (b), the perceived geometry is that of (a). This famous experiment of Oppenheim and Lim [Oppenheim & Lim 1981] shows that the geometry of an image is mostly encoded in the phase component.

for all $\boldsymbol{\xi} \in \mathbb{Z}^2$, $\hat{u}(\boldsymbol{\xi}) = |\hat{u}(\boldsymbol{\xi})|e^{i\varphi(\boldsymbol{\xi})}$. If $\hat{u}(\boldsymbol{\xi}) \neq 0$, the phase coefficient $\varphi(\boldsymbol{\xi})$ is uniquely defined modulo 2π while any arbitrary value can be chosen if $\hat{u}(\boldsymbol{\xi}) = 0$. The term "random phase function" was defined in Definition 2.1.5.

We shall also need the (non-necessarily integer) Nyquist frequencies denoted by $\boldsymbol{\eta}_x = (-\frac{M}{2}, 0)$, $\boldsymbol{\eta}_y = (0, -\frac{N}{2})$, $\boldsymbol{\eta}_{xy} = (-\frac{M}{2}, -\frac{N}{2})$. When integer, these are (with zero) the only points $\boldsymbol{\xi} \in \Omega$ which are equal to $-\boldsymbol{\xi}$ modulo (M, N).

Finally, we will also use the Gaussian tail distribution defined by

$$\forall t \in \mathbb{R}, \quad \Phi(t) = \frac{1}{\sqrt{2\pi}} \int_t^{+\infty} e^{-s^2/2} ds \; . \tag{4.1}$$

4.2.2 Global Phase Coherence

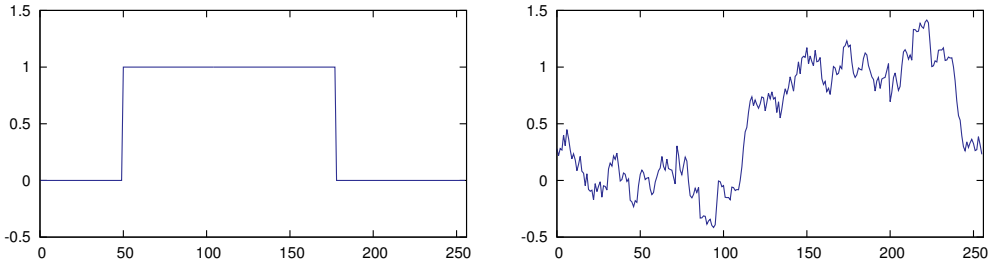
As noticed in [Oppenheim & Lim 1981], most of the geometry of an image is encoded in its phase coefficients. In Fig. 4.1, we reproduce the experiment which consists in exchanging the moduli of two images: as can be seen, the geometry of the image whose phase was kept persists. From there, in an *a contrario* framework, the authors of [Blanchet *et al.* 2008] define the global phase coherence (GPC) by measuring how much the geometry is affected when the phase information is lost.

More precisely, given u and a random phase function ψ (in the sense of Definition 2.1.5), one can define a random real-valued image u_ψ by

$$\forall \boldsymbol{\xi} \in \Omega, \quad \widehat{u_\psi}(\boldsymbol{\xi}) = |\hat{u}(\boldsymbol{\xi})|e^{i\psi(\boldsymbol{\xi})} \; .$$

or equivalently, using the reconstruction formula, by

$$\forall \mathbf{x} \in \Omega, \quad u_\psi(\mathbf{x}) = \frac{1}{MN} \sum_{\boldsymbol{\xi} \in \Omega} |\hat{u}(\boldsymbol{\xi})|e^{i\langle \mathbf{x}, \boldsymbol{\xi} \rangle + i\psi(\boldsymbol{\xi})} \; . \tag{4.2}$$



Step function: TV = 2 After phase randomization: TV = 17.4

Figure 4.2: **Phase randomization of a step function.** Notice the large increase of TV caused by phase randomization.

Up to the mean value, the random image u_ψ is the random-phase noise (RPN) associated with u [Van Wijk 1991, Galerne *et al.* 2011b], and using the notation of Definition 2.1.9, we have $u_\psi = R_{\frac{1}{\sqrt{MN}}} u$. Notice that $\frac{1}{\sqrt{MN}} u$ and the normalized spot $\frac{1}{\sqrt{MN}}(u - \bar{u})$ only differs from a constant image but this has no importance regarding the following phase coherence indices because they only involve the derivatives of u_ψ. Equation (4.2) can also be written with cosine functions only. For example, if M and N are odd integers (to get rid of Nyquist frequencies), one has

$$\forall \mathbf{x} \in \Omega, \quad u_\psi(\mathbf{x}) = |\hat{u}(0)|(-1)^{\varepsilon_0} + \frac{1}{MN} \sum_{\boldsymbol{\xi} \in \Omega_+} 2|\hat{u}(\boldsymbol{\xi})| \cos(\psi(\boldsymbol{\xi}) + \langle \mathbf{x}, \boldsymbol{\xi} \rangle) \,,$$

where $\varepsilon_0 = \mathbf{1}_{\psi(0)=\pi}$, and Ω_+ is a subset of $\Omega \backslash \{0\}$ that contains one point from each pair of symmetrical points of Ω, so that $\Omega = \{0\} \cup \Omega_+ \cup (-\Omega_+)$ is a partition of Ω. This formula shows that the phase randomization shifts the placement of the cosine components of the signal so that some oscillations will appear in the regions where the original image was flat. Thus, it becomes natural to expect the TV to increase greatly after phase randomization. This effect is striking on the one-dimensional example given in Fig. 4.2. The authors of [Blanchet *et al.* 2008] derive from this observation the following

Definition 4.2.1 (Global Phase Coherence [Blanchet *et al.* 2008])**.**
The global phase coherence of an image u is the number

$$\mathrm{GPC}(u) = -\log_{10} \mathbb{P}(\mathrm{TV}(u_\psi) \leqslant \mathrm{TV}(u)) \,. \tag{4.3}$$

In other words, the higher the GPC, the smaller the probability for TV to decrease by phase randomization. Notice that this probability can be very small (10^{-1000} and even less), and thus out of reach of most computer representations of floating point numbers (arithmetic underflow). This is why the $\log_{10}$ function is introduced in the definition (another reason is the nice interpretation of (minus) the logarithm of a probability in Information Theory).

Experimentally, it has been observed that corrupting an image with blur or noise tend to decrease its GPC. Intuitively, when an image u is blurred, its high-frequency components are attenuated, so that the oscillations of the RPN realizations are smoother; therefore, the TV increase entailed by the phase randomization

is expected to be less dramatic than in the sharp case. Now, in a noisy image, the flat regions are corrupted (by the noise) with high frequency variations leading to a TV value which is already high, so that the TV increase produced by the phase randomization is smaller than in a clean image. For now, we have no theoretical justification that goes beyond these heuristic remarks, but they will be confirmed by a practical study in Subsection 4.4.4.

The major drawback of Definition 4.2.1 is that no closed-form formula is available to compute $\mathrm{GPC}(u)$ as an explicit function of u, so that one has to use a computationally heavy Monte-Carlo procedure to estimate it. Assuming the distribution of $\mathrm{TV}(u_\psi)$ to be approximately Gaussian, the authors of [Blanchet *et al.* 2008] suggested to approximate $\mathrm{GPC}(u)$ by ("ga" stands for Gaussian approximation)

$$\mathrm{GPC}_{ga}(u) = -\log_{10} \Phi\left(\frac{\mu_0 - \mathrm{TV}(u)}{\sigma_0}\right) , \tag{4.4}$$

$$\text{where} \quad \mu_0 = \mathbb{E}(\mathrm{TV}(u_\psi)) , \quad \sigma_0^2 = \mathrm{Var}(\mathrm{TV}(u_\psi)) , \tag{4.5}$$

$$\text{and} \quad \Phi(t) = \frac{1}{\sqrt{2\pi}} \int_t^{+\infty} e^{-s^2/2} ds \tag{4.6}$$

is the Gaussian tail function ("ga" stands for Gaussian approximation). The values of μ_0 and σ_0 can be estimated through N Monte-Carlo samples

$$\mathrm{TV}\left(u_\psi^{(1)}\right), \mathrm{TV}\left(u_\psi^{(2)}\right), \ldots, \mathrm{TV}\left(u_\psi^{(N)}\right)$$

of the r.v. $\mathrm{TV}(u_\psi)$, which leads to a numerical approximation $\mathrm{GPC}_N(u)$ of $\mathrm{GPC}(u)$. Unfortunately, due to the fact that each Monte Carlo sample requires the computation of a Fourier transform, the resulting algorithm is quite slow (even with a good C implementation, it takes about one minute to obtain a merely decent estimate of the GPC of a 512×512 image on a standard 3Ghz laptop). Let us mention that the Gaussian approximation of $\mathrm{TV}(u_\psi)$ is analyzed theoretically in Appendix 4.A and Appendix 4.B. From a numerical point of view, the quality of the Gaussian approximation can be evaluated by a Monte-Carlo approach. Using N samples of u_ψ, one can compute F_N, the empirical estimate of the tail distribution of $\mathrm{TV}(u_\psi)$, and compare it to its Gaussian counterpart Φ. We checked for $N = 10,000$ and several different images that $\|F_N - \Phi\|_\infty < 0.01$.

4.2.3 Sharpness Index

In a later work [Blanchet & Moisan 2012], a new measure of phase coherence was introduced. It was noticed that when replacing the random model u_ψ by $u * W$, that is, the convolution of u with a Gaussian white noise W, the expectation and variance of $\mathrm{TV}(u * W)$ could be computed explicitly as a function of u. Thus, with the same framework as above, one can define

$$\mathcal{SI}(u) = -\log_{10} \mathbb{P}(\mathrm{TV}(u * W) \leqslant \mathrm{TV}(u)) \tag{4.7}$$

and, assuming as in [Blanchet & Moisan 2012] that the r.v. $\mathrm{TV}(u * W)$ is approximately Gaussian,

Definition 4.2.2 (Sharpness Index [Blanchet & Moisan 2012])**.**
The Sharpness Index (SI) of an image u is

$$\mathrm{SI}(u) = -\log_{10}\Phi\left(\frac{\mu - \mathrm{TV}(u)}{\sigma}\right) \tag{4.8}$$

where Φ is defined by (4.6),

$$\mu = \mathbb{E}(\mathrm{TV}(u * W)) \;, \quad \sigma^2 = \mathrm{Var}(\mathrm{TV}(u * W)) \;, \tag{4.9}$$

and W is a Gaussian white noise with standard deviation $|\Omega|^{-1/2}$ (i.e. the r.v. $W(\mathbf{x}), \mathbf{x} \in \Omega$ are independent with distribution $\mathcal{N}(0, |\Omega|^{-1})$).

There are several reasons to expect GPC and SI to behave in the same way. First, the corresponding random image models (RPN for GPC, Gaussian for SI) are known to be close, both mathematically (they only differ by a Rayleigh noise on the Fourier modulus) and perceptually (see [Galerne *et al.* 2011b]). Second, it has been noticed experimentally in [Blanchet & Moisan 2012] that the values of μ_0 (Equation (4.5)) and μ (Equation (4.9)) were very close in general (a relative error below 1%). In Appendix 4.A, we confirm this experimental observation by a precise asymptotic result (Theorem 4.A.2) based on Berry-Esseen theorem.

The fact that $\mathrm{TV}(u * W)$ is nearly Gaussian (which is used without formal justification in [Blanchet & Moisan 2012]) can again be confirmed by a Monte-Carlo estimation of the distribution of $\mathrm{TV}(u * W)$. We also give an asymptotic proof in Appendix 4.B using a particular central limit theorem devoted to sums of non-independent random variables controlled by a dependency graph.

The great interest of SI over GPC is that it can be computed with explicit formulae instead of a costly Monte-Carlo simulation, as shown in

Theorem 4.2.1 ([Blanchet & Moisan 2012])**.** *Let $u : \Omega \to \mathbb{R}$ be an image, and let $W : \Omega \to \mathbb{R}$ be a Gaussian white noise with mean 0 and standard deviation $|\Omega|^{-1/2}$. Then*

$$\mu = \mathbb{E}(\mathrm{TV}(u * W)) = (\alpha_x + \alpha_y)\sqrt{\frac{2}{\pi}}\sqrt{|\Omega|} \;, \tag{4.10}$$

$$\begin{aligned}\sigma^2 = \mathrm{Var}(\mathrm{TV}(u * W)) &= \frac{2}{\pi}\sum_{\mathbf{z}\in\Omega} \alpha_x^2 \cdot \omega\left(\frac{\Gamma_{xx}(\mathbf{z})}{\alpha_x^2}\right) \\ &+ 2\alpha_x\alpha_y \cdot \omega\left(\frac{\Gamma_{xy}(\mathbf{z})}{\alpha_x\alpha_y}\right) + \alpha_y^2 \cdot \omega\left(\frac{\Gamma_{yy}(\mathbf{z})}{\alpha_y^2}\right) \;,\end{aligned} \tag{4.11}$$

where

$$\begin{aligned}\alpha_x^2 &= \|\partial_x \dot{u}\|_2^2 = \sum_{(x,y)\in\Omega} |\dot{u}(x+1,y) - \dot{u}(x,y)|^2 \;, \\ \alpha_y^2 &= \|\partial_y \dot{u}\|_2^2 = \sum_{(x,y)\in\Omega} |\dot{u}(x,y+1) - \dot{u}(x,y)|^2 \;, \\ \forall t \in [-1,1], \quad \omega(t) &= t\arcsin(t) + \sqrt{1-t^2} - 1 \;, \\ \Gamma(\mathbf{z}) = \begin{pmatrix}\Gamma_{xx}(\mathbf{z}) & \Gamma_{xy}(\mathbf{z}) \\ \Gamma_{yx}(\mathbf{z}) & \Gamma_{yy}(\mathbf{z})\end{pmatrix} &= \sum_{\mathbf{y}\in\Omega} \nabla\dot{u}(\mathbf{y}) \cdot \nabla\dot{u}(\mathbf{y}+\mathbf{z})^T \;.\end{aligned} \tag{4.12}$$

Proof. A short proof was given in [Blanchet & Moisan 2012]. In order not to break the discussion about the different definitions of phase coherence, we postpone the complete proof to Appendix 4.C. □

Remark: What happens if we replace the TV (l^1-norm of gradient) by the H^1-norm (l^2-norm of gradient) in the definition of SI? With Parseval's formula, one can see that the H^1-norm only depends on the Fourier modulus, so that it is not affected by the phase randomization. Hence, the corresponding indices obtained with the H^1-norm are trivial. Considering another $W^{1,p}$-norm (that is, the l^p-norm of gradient) could be interesting, but it is likely that the easiest calculations are obtained with TV ($p = 1$).

4.2.4 A Simplified Version of SI

In [Leclaire & Moisan 2013a], we suggested to approximate the denominator of the fraction appearing in (4.8), which led us to a new index (written S) that is analytically close to SI but can be computed much faster. We will see empirically later in Section 4.3 and Section 4.4 that S also behaves like SI with respect to basic image transformations.

4.2.4.1 Definition of S

Lemme 4.2.1. *The function ω defined by (4.12) satisfies*

$$\forall t \in [-1,1], \quad \frac{1}{2}t^2 \leqslant \omega(t) \leqslant \frac{1}{2}t^2 + ct^4 \,, \tag{4.13}$$

where $c = \frac{\pi-3}{2} \approx 0.0708$ is the optimal (that is, minimal) constant in (4.13).

Proof. One has for all $t \in [-1,1]$,

$$\omega'(t) = \arcsin(t) = \sum_{n\geqslant 0} \frac{(2n)!}{2^{2n}(n!)^2}\left(\frac{1}{2n+1}\right) t^{2n+1} \,,$$

(notice that the series is absolutely convergent for $|t| = 1$ thanks to Stirling's formula). After term-by-term integration, one can write

$$\omega(t) = \sum_{n\geqslant 0} \frac{(2n)!}{2^{2n}(n!)^2}\left(\frac{1}{2n+1}\right)\left(\frac{1}{2n+2}\right) t^{2n+2} \,.$$

Noticing that $t \mapsto \frac{1}{t^4}(\omega(t) - \frac{1}{2}t^2)$ is an even function which is increasing on $[0,1]$, the result follows by taking

$$\begin{aligned} c &= \sum_{n\geqslant 1} \frac{(2n)!}{2^{2n}(n!)^2}\left(\frac{1}{2n+1}\right)\left(\frac{1}{2n+2}\right) \\ &= \lim_{t\to 1} \frac{\omega(t) - \frac{1}{2}t^2}{t^4} = \omega(1) - \frac{1}{2} = \frac{\pi-3}{2} \,. \end{aligned}$$

□

The term (4.11) can thus be approximated by replacing $\omega(t)$ by $\frac{t^2}{2}$. This leads to

$$\sigma_a^2 = \frac{1}{\pi}\left(\frac{\|\Gamma_{xx}\|_2^2}{\alpha_x^2} + 2\cdot\frac{\|\Gamma_{xy}\|_2^2}{\alpha_x\alpha_y} + \frac{\|\Gamma_{yy}\|_2^2}{\alpha_y^2}\right), \tag{4.14}$$

and to

Definition 4.2.3 (S index [Leclaire & Moisan 2013a])**.** The simplified sharpness index associated to an image u is

$$S(u) = -\log_{10}\Phi\left(\frac{\mu - \mathrm{TV}(u)}{\sigma_a}\right),$$

where σ_a is given by (4.14), Φ by (4.6), and μ by (4.10).

4.2.4.2 Fast calculation

Since the last formula is now free of ω, the index S is, compared to SI, simpler to understand (it only depends on the autocorrelation gradient matrix through its energy) and faster to compute. In Algorithm 1, we can notice that the most costly step is the FFT computation (2.a): once $\hat{u}$ is computed, the FFTs of the two derivatives follow immediately (step 2.b), and the FFTs of the cross-correlation of the derivatives (step 2.c) follow from, e.g.,

$$\Gamma_{xx} = \partial_x\dot{u} * \widetilde{\partial_x\dot{u}} \quad\Rightarrow\quad |\widehat{\Gamma_{xx}}| = |\widehat{\partial_x\dot{u}}|^2, \tag{4.15}$$

with the convention that $\widetilde{w}(\mathbf{x}) = w(-\mathbf{x})$. In the end, the computation of $S(u)$ requires only 1 FFT, whereas 3 more FFTs are required for $\mathrm{SI}(u)$. In both cases, however, the complexity is the same, $\mathcal{O}(MN\log MN)$.

4.2.4.3 Theoretical comparison with SI

We here investigate the quality of the approximation of SI by S, showing that the fraction

$$\mathrm{v}_a(u) = \frac{\mu - \mathrm{TV}(u)}{\sigma_a}$$

is a good approximation of

$$\mathrm{v}(u) = \frac{\mu - \mathrm{TV}(u)}{\sigma}.$$

Proposition 4.2.2. *We have*

$$0 \leqslant \frac{\mathrm{v}_a(u) - \mathrm{v}(u)}{\mathrm{v}_a(u)} \leqslant 1 - \frac{1}{\sqrt{\pi - 2}} \approx 0.064. \tag{4.16}$$

Proof. We first show that

$$0 \leqslant \frac{\sigma^2 - \sigma_a^2}{\sigma_a^2} \leqslant 2c = \pi - 3 \approx 0.142. \tag{4.17}$$

Algorithm 1 : Computation of $S(u)$

1. Compute the derivatives $\partial_x \dot{u}$, $\partial_y \dot{u}$ and deduce their l^1 and l^2 norms

$$\mathrm{TV}(u)\ , \quad \alpha_x = \|\partial_x \dot{u}\|_2\ , \quad \alpha_y = \|\partial_y \dot{u}\|_2\ .$$

2. Compute (in Fourier domain) the components of the autocorrelation gradient matrix Γ:

   2.a Compute the FFT $\hat{u}$ of u.

   2.b Deduce the FFTs of the derivatives using

$$\left|\widehat{\partial_x \dot{u}}(\boldsymbol{\xi})\right|^2 = 4\sin^2\left(\frac{\pi\xi_1}{M}\right)|\hat{u}(\boldsymbol{\xi})|^2\ ,$$

$$\left|\widehat{\partial_y \dot{u}}(\boldsymbol{\xi})\right|^2 = 4\sin^2\left(\frac{\pi\xi_2}{N}\right)|\hat{u}(\boldsymbol{\xi})|^2\ .$$

   2.c Compute the moduli of the FFTs of Γ_{xx}, Γ_{xy} and Γ_{yy} using

$$|\widehat{\Gamma_{xx}}| = |\widehat{\partial_x \dot{u}}|^2,\ |\widehat{\Gamma_{xy}}| = |\widehat{\partial_x \dot{u}}||\widehat{\partial_y \dot{u}}|,\ |\widehat{\Gamma_{yy}}| = |\widehat{\partial_y \dot{u}}|^2.$$

3. Compute μ and σ_a with

$$\mu = (\alpha_x + \alpha_y)\sqrt{\frac{2}{\pi}}\sqrt{MN} \quad \text{and}$$

$$\sigma_a^2 = \frac{1}{\pi MN}\left(\frac{\|\widehat{\Gamma_{xx}}\|_2^2}{\alpha_x^2} + 2\cdot\frac{\|\widehat{\Gamma_{xy}}\|_2^2}{\alpha_x\alpha_y} + \frac{\|\widehat{\Gamma_{yy}}\|_2^2}{\alpha_y^2}\right)\ .$$

4. Finally compute

$$S(u) = -\log_{10}\Phi\left(\frac{\mu - \mathrm{TV}(u)}{\sigma_a}\right)$$

   using, if required, the logerf function detailed in [Louchet & Moisan 2014, Algorithm 1].

With the expressions of σ and σ_a, one can write

$$\begin{aligned}\sigma^2 - \sigma_a^2 = \frac{2}{\pi}\sum_{\mathbf{x}\in\Omega}\alpha_x^2\left[\omega\left(\frac{\Gamma_{xx}(\mathbf{x})}{\alpha_x^2}\right) - \frac{1}{2}\left(\frac{\Gamma_{xx}(\mathbf{x})}{\alpha_x^2}\right)^2\right] \\ +2\alpha_x\alpha_y\left[\omega\left(\frac{\Gamma_{xy}(\mathbf{x})}{\alpha_x\alpha_y}\right) - \frac{1}{2}\left(\frac{\Gamma_{xy}(\mathbf{x})}{\alpha_x\alpha_y}\right)^2\right] \\ +\alpha_y^2\left[\omega\left(\frac{\Gamma_{yy}(\mathbf{x})}{\alpha_y^2}\right) - \frac{1}{2}\left(\frac{\Gamma_{yy}(\mathbf{x})}{\alpha_y^2}\right)^2\right] .\end{aligned}$$

Using Lemma 4.2.1, we thus obtain

$$\forall t \in [-1,1]\,, \quad 0 \leqslant \omega(t) - \frac{1}{2}t^2 \leqslant ct^4 \leqslant ct^2\,,$$

which implies

$$\begin{aligned}0 &\leqslant \sigma^2 - \sigma_a^2 \\ &\leqslant \frac{2c}{\pi}\sum_{\mathbf{x}\in\Omega}\alpha_x^2\cdot\left(\frac{\Gamma_{xx}(\mathbf{x})}{\alpha_x^2}\right)^2 + 2\alpha_x\alpha_y\cdot\left(\frac{\Gamma_{xy}(\mathbf{x})}{\alpha_x\alpha_y}\right)^2 + \alpha_y^2\cdot\left(\frac{\Gamma_{yy}(\mathbf{x})}{\alpha_y^2}\right)^2\,,\end{aligned}$$

and the right-hand term equals $2c\sigma_a^2$, which proves (4.17). Now, since

$$\frac{\mathrm{v}(u)}{\mathrm{v}_a(u)} = \left(1 + \frac{\sigma^2 - \sigma_a^2}{\sigma_a^2}\right)^{-1/2}\,,$$

we get (4.16) as expected. □

Notice that (4.16) provides a simple universal bound on the relative error $\frac{\mathrm{v}_a(u)-\mathrm{v}(u)}{\mathrm{v}_a(u)}$. Using the same technique, it could be possible to derive a sharper bound depending on u.

To end this section, let us recall the definitions of $\mathcal{SI}$, SI, and S.

$$\begin{aligned}\mathcal{SI}(u) &= -\log_{10}\mathbb{P}(\mathrm{TV}(u*W) \leqslant \mathrm{TV}(u)) \\ \mathrm{SI}(u) &= -\log_{10}\Phi\left(\frac{\mu - \mathrm{TV}(u)}{\sigma}\right) \\ S(u) &= -\log_{10}\Phi\left(\frac{\mu - \mathrm{TV}(u)}{\sigma_a}\right)\end{aligned}$$

where Φ is given by (4.6), μ by (4.10), σ by (4.11), and σ_a by (4.14).

4.3 Mathematical Properties

4.3.1 First properties

Proposition 4.3.1. *The functions* GPC, $\mathcal{SI}$, SI, S *are non-negative and invariant with respect to affine contrast changes, that is, for* $f \in \{\mathrm{GPC}, \mathcal{SI}, \mathrm{SI}, S\}$, *one has*

$$\forall a, b \in \mathbb{R}, a \neq 0, \quad f(a\cdot u + b) = f(u)\,.$$

Proof. These properties directly result from the definitions. □

Let us now explore the Fourier representation of the random field $u * W$. Its DFT is $\hat{u}\hat{W}$. Since W is a Gaussian white noise, $\hat{W}$ is a complex Gaussian white noise. In particular, one can write

$$\hat{W}(\boldsymbol{\xi}) = |\hat{W}(\boldsymbol{\xi})|e^{i\psi(\boldsymbol{\xi})}$$

where ψ is a random phase function in the sense of Definition 2.1.5. Denoting by T the random image such that $\hat{T} = |\hat{W}|$, one has

$$u * W = u_{\varphi+\psi} * T$$

where $\varphi+\psi$ is also a random phase. Therefore, in comparison to the phase randomization model, the operation $u \mapsto u * W$ also includes a convolution by an image T whose Fourier transform is $|\hat{W}|$. Following [Desolneux *et al.* 2012], we can say that T is the white noise texton. Proposition 1 of [Desolneux *et al.* 2012] shows that, statistically, T looks like a Dirac mass in zero (up to a factor $\sqrt{\pi}/2$). Hence, one can expect that this convolution will not drastically modify the statistical properties of the model, and, subsequently, that SI(u) behaves like GPC(u). Incidentally, the discussion above brings an interesting remark, formulated by the following

Proposition 4.3.2. GPC(u), $\mathcal{SI}(u)$, SI(u), *and* $S(u)$ *only depend on the modulus and the TV of* u.

Proof. For GPC(u) and $\mathcal{SI}(u)$, this is because the distributions of u_ψ and $u * W$ only depend on $|\hat{u}|$. For SI(u) and $S(u)$ this is because the gradient autocorrelation and energy only depend on $|\hat{u}|$. □

Thus, all these indices measure the global phase coherence of an image u only by its impact on the TV, in a way (a "scale") that is determined by the modulus of u. As we shall see later in Section 4.4, when an image is filtered by a symmetrical kernel that has a positive Fourier Transform (e.g., a Gaussian kernel), its phase is not changed but the indices above tend to decrease (with the exception of the Dirac image that will be discussed in Subsection 4.4.5).

Notice also that since we are using a periodic scheme for TV, these indices take the same values on u and on the periodic translation $\tau_{(a,b)}u$ defined by

$$\forall (x,y) \in \mathbb{Z}^2, \quad \tau_{(a,b)}u(x,y) = \dot{u}(x-a, y-b) \ .$$

4.3.2 Regularity, Analytical Difficulties

The expression for SI(u) in Theorem 4.2.1 is not defined when u is a constant image. In that case, Equation (4.7) implies that $\mathcal{SI}(u)$ is zero. It is not a big issue because natural images are never really constant. Apart from these singular points, one can state the following

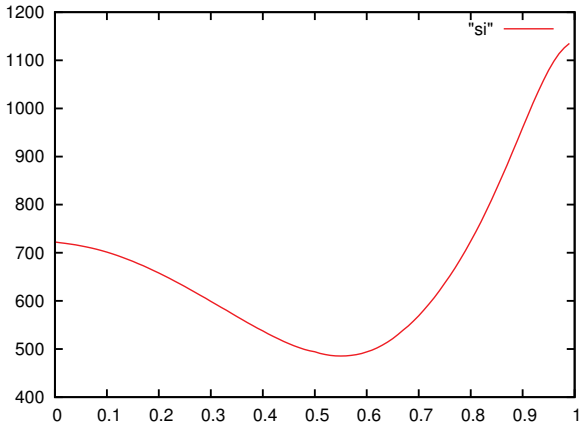


Figure 4.3: **A one-dimensional profile of SI**. This graph of the function $\lambda \mapsto \mathrm{SI}(\lambda u_1 + (1-\lambda)u_2)$ (where u_1 and u_2 refer to the images *Lena* and *Barbara* respectively) shows that SI is neither convex nor concave.

Proposition 4.3.3. *Let us introduce*

$$D = \{u \in \mathbb{R}^\Omega,\ \|\partial_x \dot{u}\|_2 \neq 0 \text{ and } \|\partial_y \dot{u}\|_2 \neq 0\} \text{ and}$$

$$D' = \{u \in \mathbb{R}^\Omega,\ \forall \mathbf{x} \in \Omega,\ \partial_x \dot{u}(\mathbf{x}) \neq 0 \text{ and } \partial_y \dot{u}(\mathbf{x}) \neq 0\}\ .$$

The functions SI *and* S *are defined and continuous on* D *and infinitely differentiable on* D'*.*

Proof. Let us consider an image $u \in D$. Thanks to (4.15) we have $\|\Gamma_{xx}\|_2 \neq 0$, and similarly $\|\Gamma_{xy}\|_2 \neq 0$ and $\|\Gamma_{yy}\|_2 \neq 0$. Consequently, σ and σ_a are non-zero, and $\mathrm{SI}(u)$ and $S(u)$ are well-defined. Moreover, the continuity of SI and S follows from the one of α_x, α_y, Γ and TV. For the second part, we simply notice that the functions α_x, α_y, σ and σ_a are smooth on D, so the singular points of SI and S are those of TV, that is, the images that do not belong to D'. □

The fact that SI have some singular points would not be very embarrassing in an optimization perspective. Indeed, several techniques are available to optimize non-smooth quantities, in particular for convex functions [Ekeland & Témam 1999]. Unfortunately, the function SI is neither convex nor concave, as shown in Fig. 4.3. For those reasons, applying classical optimization techniques (like gradient or sub-gradient descent schemes) on SI may not be efficient. We will overcome this difficulty in Section 4.5 by considering simple generic algorithms relying on stochastic optimization.

4.3.3 Distribution of GPC on a random phase field

We continue with an explicit statement that generalizes a property mentioned (without proof) in [Blanchet *et al.* 2008].

Proposition 4.3.4. *If U is a random image such that its phase is uniform (in the sense of Definition 2.1.5) and independent of its modulus, then*

$$\forall t > 0, \quad \mathbb{P}(\mathrm{GPC}(U) \geqslant t) \leqslant 10^{-t} . \tag{4.18}$$

Furthermore, if conditionally on $|\widehat{U}|$, the r.v. $\mathrm{TV}(U)$ admits a probability density function, then

$$\forall t > 0, \quad \mathbb{P}(\mathrm{GPC}(U) \geqslant t) = 10^{-t} , \tag{4.19}$$

that is, $10^{-\mathrm{GPC}(U)}$ is uniform on $[0, 1]$.

A consequence of (4.18) is that a texture obtained as the realization of a RPN model or a stationary Gaussian model is expected to have a small GPC value (that is, below 3 or 4 in general), which is in accordance with the fact that such texture models do not carry any phase information. As concerns the hypothesis required for the second part of Proposition 4.3.4, it may be satisfied as soon as U is not constant almost surely, but we did not find the proof of such a statement yet.

Proposition 4.3.4 can be obtained from the following two Lemmas by considering conditional distributions given $|\widehat{U}|$. Lemma 4.3.5 is a general result about cumulative distribution functions that is the key of the proof of Lemma 4.3.6.

Lemme 4.3.5. *If Y is a r.v. and $F(v) = \mathbb{P}(Y \leqslant v)$, then*

$$\forall s \in [0, 1], \quad \mathbb{P}(F(Y) \leqslant s) \leqslant s ,$$

and the equality holds for all s as soon as Y admits a probability density function.

Proof. This is a reformulation of Lemma 1 of [Grosjean & Moisan 2009]. □

Lemme 4.3.6. *If u is an image and if ψ is a random phase function (in the sense of Definition 2.1.5), then*

$$\forall t > 0, \quad \mathbb{P}(\mathrm{GPC}(u_\psi) \geqslant t) \leqslant 10^{-t} .$$

Furthermore, if the r.v. $\mathrm{TV}(u_\psi)$ admits a probability density function, then

$$\forall t > 0, \quad \mathbb{P}(\mathrm{GPC}(u_\psi) \geqslant t) = 10^{-t} .$$

Proof. Let us denote by F_u the cumulative distribution function of the r.v. $\mathrm{TV}(u_\psi)$, defined by

$$\forall t \in \mathbb{R}, \quad F_u(t) = \mathbb{P}(\mathrm{TV}(u_\psi) \leqslant t) .$$

The definition of GPC implies that for any image u,

$$\mathrm{GPC}(u) = -\log_{10} F_u(\mathrm{TV}(u)) .$$

Since the distribution of $\mathrm{TV}(u_\psi)$ only depends on the modulus of u, we have $F_u = F_{u_\chi}$ for any phase function χ. In particular, if ψ is a random phase function, one can write

$$\mathrm{GPC}(u_\psi) = -\log_{10} F_{u_\psi}(\mathrm{TV}(u_\psi)) = -\log_{10} F_u(\mathrm{TV}(u_\psi))$$

so that for all $t > 0$,

$$\mathbb{P}(\mathrm{GPC}(u_\psi) \geqslant t) = \mathbb{P}\left(F_u(\mathrm{TV}(u_\psi)) \leqslant 10^{-t}\right) .$$

Because F_u is the cumulative distribution function of $\mathrm{TV}(u_\psi)$, Lemma 4.3.5 allows us to conclude that this probability is smaller than 10^{-t}. The equality case is obtained similarly from the equality case of Lemma 4.3.5. □

Now we provide a similar result for the approximation of GPC defined in (4.4).

Proposition 4.3.7. *Let u be an image and ψ a random phase function (in the sense of Definition 2.1.5). Write $\mu_0 = \mathbb{E}(\mathrm{TV}(u_\psi))$, $\sigma_0^2 = \mathrm{Var}(\mathrm{TV}(u_\psi))$, and denote by $\widetilde{F}_u$ the tail distribution of the normalized r.v.*

$$T = \frac{\mu_0 - \mathrm{TV}(u_\psi)}{\sigma_0} ,$$

and by G_u the cumulative distribution function of the r.v. $10^{-\mathrm{GPC}_{ga}(u_\psi)}$. If $\mathrm{TV}(u_\psi)$ admits a probability density function then

$$\sup_{s\in[0,1]} |G_u(s) - s| \leqslant \sup_{t\in\mathbb{R}} |\widetilde{F}_u(t) - \Phi(t)| \tag{4.20}$$

Proposition 4.3.7 shows that, in terms of the L^∞ distance between the cumulative distribution functions, the approximation of $10^{-\mathrm{GPC}_{ga}(u_\psi)}$ by the uniform distribution on $[0,1]$ is at least as good as the Gaussian approximation of $\mathrm{TV}(u_\psi)$.

Proof. One can remark that

$$\begin{aligned} 10^{-\mathrm{GPC}(u)} &= \mathbb{P}\left(\frac{\mu_0 - \mathrm{TV}(u_\psi)}{\sigma_0} \geqslant \frac{\mu_0 - \mathrm{TV}(u)}{\sigma_0}\right) \\ &= \widetilde{F}_u\left(\frac{\mu_0 - \mathrm{TV}(u)}{\sigma_0}\right) . \end{aligned}$$

Moreover, we have by definition

$$10^{-\mathrm{GPC}_{ga}(u)} = \Phi\left(\frac{\mu_0 - \mathrm{TV}(u)}{\sigma_0}\right) .$$

Since $\widetilde{F}_u$, μ_0 and σ_0 depend on u only through its modulus, we also have

$$\begin{aligned} 10^{-\mathrm{GPC}(u_\psi)} &= \widetilde{F}_u\left(\frac{\mu_0 - \mathrm{TV}(u_\psi)}{\sigma_0}\right) \\ \text{and} \quad 10^{-\mathrm{GPC}_{ga}(u_\psi)} &= \Phi\left(\frac{\mu_0 - \mathrm{TV}(u_\psi)}{\sigma_0}\right) . \end{aligned}$$

In particular,

$$\left|10^{-\mathrm{GPC}(u_\psi)} - 10^{-\mathrm{GPC}_{ga}(u_\psi)}\right| \leqslant \varepsilon ,$$

where $\varepsilon = \sup_{t\in\mathbb{R}} |\widetilde{F}_u(t) - \Phi(t)|$. Since we assumed that $\mathrm{TV}(u_\psi)$ admits a probability density function, Lemma 4.3.6 ensures that the r.v. $X = 10^{-\mathrm{GPC}(u_\psi)}$ follows the uniform distribution on $(0,1)$. So we have almost surely

$$\left|X - 10^{-\mathrm{GPC}_{ga}(u_\psi)}\right| \leqslant \varepsilon ,$$

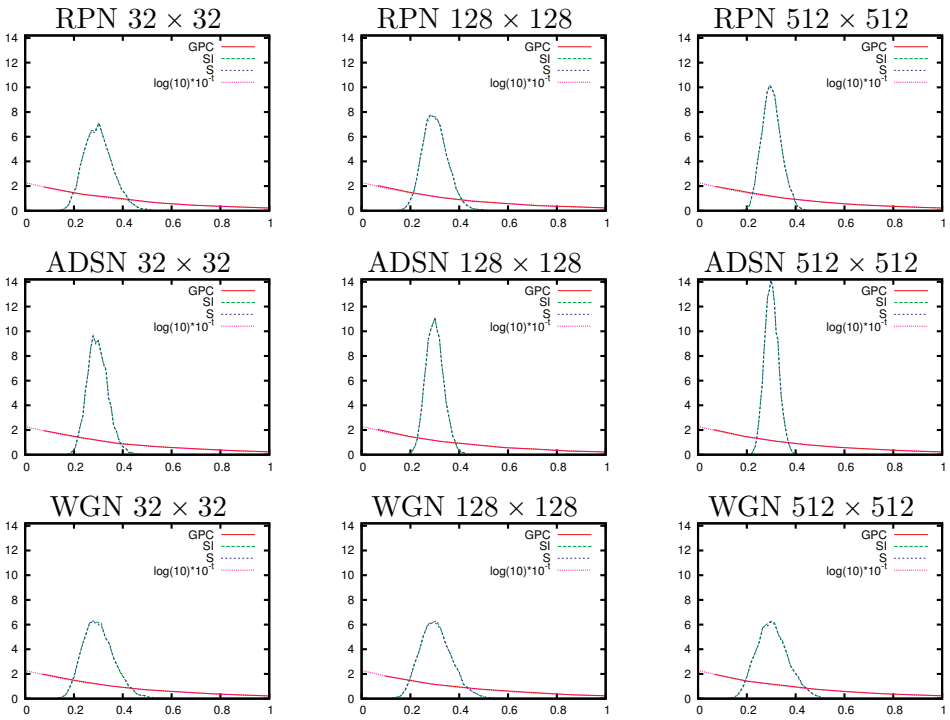


Figure 4.4: **Phase coherence indices of random phase fields.** Each graph represents the estimated distributions (using the same 10,000 samples) of the r.v. GPC(U), SI(U) and $S(U)$. The size of the random image U is, respectively, 32×32 for the left column, 128×128 for the middle column and 512×512 for the right column. For the first line, U is the random phase noise (RPN) associated to the image *Lena*. For the second line, U is the asymptotic discrete spot noise (ADSN) $u * W$ where u is again the image *Lena*. And for the third line, U is simply a white Gaussian noise (WGN). First, we observe as predicted by Proposition 4.3.4 that the distribution of GPC(U) has density $t \mapsto \log(10)10^{-t}\mathbf{1}_{t>0}$. Furthermore, we can also observe that the distributions of SI(U) and $S(U)$ appear similar but that they do not coincide with the one of GPC(U). Last, we can see that on the RPN and ADSN models, the distributions of SI and S depend on the size of the random field, whereas they apparently do not for the WGN model. However, the mean values of SI(U) and $S(U)$ remain close to 0.3.

where X is uniform on $(0,1)$, which implies the inequality (4.20) for the cumulative distribution functions. □

Notice that the result of Proposition 4.3.4 does not extend to $\mathcal{SI}$. Actually, one can see empirically in Fig. 4.4 that it neither extends to SI or S. Let us try to understand this by considering the distribution of

$$\mathrm{SI}(u_\psi) = -\log_{10}\Phi\left(\frac{\mu - \mathrm{TV}(u_\psi)}{\sigma}\right)$$

where $\mu = \mathbb{E}(\mathrm{TV}(u * W))$ and $\sigma = \mathrm{Var}(\mathrm{TV}(u * W))$. Once more, one can assume that $\mathrm{TV}(u_\psi)$ is nearly Gaussian. Concerning the first moment, it has been observed numerically in [Blanchet & Moisan 2012] that $\mathbb{E}(\mathrm{TV}(u_\psi)) \approx \mathbb{E}(\mathrm{TV}(u * W))$ (this approximation is mathematically investigated in Appendix A). Concerning the variance of $\mathrm{TV}(u_\psi)$, however, numerical simulations indicate that it significantly differs (by a factor 7-8 in general [Blanchet & Moisan 2012]) from that of $\mathrm{TV}(u * W)$. A consequence is that the r.v. $G = \frac{\mu - \mathrm{TV}(u_\psi)}{\sigma}$ has a distribution close to $\mathcal{N}(0, s^2)$ for some s that is *not* close to 1. Therefore, one cannot expect the distribution of $\Phi(G) = 10^{-\mathrm{SI}(u_\psi)}$ to be close to the uniform distribution on $(0,1)$. However, one can see in Fig. 4.4 that the sharpness values of random phase fields is in general concentrated around 0.3.

To end this subsection, we mention (without proof) another result concerning the RPN model.

Proposition 4.3.8. *If u is an image and ψ a discrete random phase field (in the sense of Definition 2.1.5), then*

$$\begin{aligned}\mathbb{P}(\mathrm{SI}(u_\psi) \geqslant \mathrm{SI}(u)) &= \mathbb{P}(\mathcal{SI}(u_\psi) \geqslant \mathcal{SI}(u)) \\ &= \mathbb{P}(\mathrm{TV}(u_\psi) \leqslant \mathrm{TV}(u)) = 10^{-\mathrm{GPC}(u)}\,. \end{aligned} \tag{4.21}$$

4.4 Phase Coherence Indices and No-Reference Quality Assessment

This section is devoted to the practical study of the phase coherence indices. Since the computation of S is the fastest of all, we led the experiments on it, but the major part of what follows extends to GPC and SI.

4.4.1 Periodization

The index S deals more with the periodized image $\dot{u}$ than with u itself. Actually, since a periodic translation of u has no effect on $S(u)$, a discontinuity of u on the boundary has the same effect as if it were positioned in the middle of the image. So the index S is affected, and actually biased, by the discontinuities that generally occur between two opposite boundaries of an image. In [Blanchet *et al.* 2008], the authors suggest to compute the phase coherence index not on u, but on its

periodic component [Moisan 2011]. This operation subtracts from the original image a smooth component that cancels border-to-border discontinuities, see Subsection 3.1.2 for more details.

Let us also mention that it is possible to replace in Equation (4.2.1) the gradient, the TV, and the autocorrelation by their non-periodic counterparts. It leads to a "Local Sharpness Index" [Leclaire & Moisan 2013b] which is a little slower to compute but naturally insensitive to border effects.

4.4.2 Quantization

Another classical operation that can bias the phase coherence is quantization. The gray levels of 8-bits natural images are generally quantized on $\{0, 1, \dots, 255\}$, and this quantization process creates artificially flat regions.

The contribution of those regions to the TV is exactly zero, whereas it should be a small (but non-zero) number. To avoid that undesirable effect of quantization, as suggested in [Blanchet *et al.* 2008], before computing these indices, one can apply a sub-pixel translation of vector $(1/2, 1/2)$, with the following definition for the sub-pixel translation of vector (α, β),

$$\forall \boldsymbol{\xi} \in \Omega, \quad \widehat{\tau_{(\alpha,\beta)} u}(\boldsymbol{\xi}) = e^{-2i\pi\left(\frac{\alpha\xi_1}{M} + \frac{\beta\xi_2}{N}\right)} \hat{u}(\boldsymbol{\xi}) \; . \tag{4.22}$$

More generally, one could consider the sub-pixel-translation-invariant sharpness index

$$\inf_{(\alpha,\beta)\in\mathbb{R}^2} S(\tau_{(\alpha,\beta)} u) \; . \tag{4.23}$$

Since $\tau_{(a,b)} u$ and u have the same modulus, the vector (α, β) corresponding to the minimum value of $S(\tau_{(\alpha,\beta)} u)$ is actually the one that realizes the maximum value of $\mathrm{TV}(\tau_{(\alpha,\beta)} u)$. In practice, one can observe that, for most natural images, this vector is usually near $(1/2, 1/2)$, which justifies the use of $\tau = \tau_{(1/2,1/2)}$ alone.

Another way to avoid the quantization bias on the sharpness indices would be to consider

$$\min_{\|v-u\|_\infty \leqslant q/2} S(v) \; , \tag{4.24}$$

where q is the quantization step ($q = 1$ for integer-valued images). Unfortunately, S may have a lot of local minima in the neighborhood $\{\|v - u\|_\infty \leqslant q/2\}$ of u, and it seems difficult to solve (4.24) by standard optimization techniques.

To end this subsection, we would like to mention that it makes sense to penalize the quantization through the aliasing it produces in the image. The ideal solution to that would be to replace in our construction the simple discrete TV by another TV operator which is invariant by sub-pixel translation. Integrating such an operator (for example, the one suggested in [Moisan 2007]) in our model would be an interesting development. Considering (4.23) gives an alternative solution which, if u is a natural image, can be approximated by $S(\tau(u))$. Ultimately, the $(1/2, 1/2)$-sub-pixel translation is a precise and efficient solution to avoid the quantization bias.

In the experiments that are presented in the following sections, before computing the indices SI and S, we extracted the periodic component [Moisan 2011] of the image and applied to it a sub-pixel translation of vector $(1/2, 1/2)$. Since the DFT of the periodic component of u can be computed with one FFT (see [Moisan 2011]), including these two preprocessing steps yields an overall computation cost of 5 FFTs for SI and 2 FFTs for S.

4.4.3 Variations of S on natural images

Before we explore the links between the S index and the perceived sharpness of an image, we give in Fig. 4.5 some examples of the values obtained for typical natural images. Several observations can be made from these examples, which are confirmed on larger image sets:

- the S index attains higher values for images that present sharp edges and smooth regions at the same time; conversely, out-of-focus images tend to produce relatively low values of S;
- spectrally concentrated textures (in particular, periodic patterns like stripes) lead to surprisingly low values of S, even if the texture patterns are made of sharp edges;
- in general, S rapidly increases with the size of images, but since it is very content-dependent, counterexamples (image parts whose S-value is greater than the S-value of the whole image) can be found.

4.4.4 Influence of blur and noise

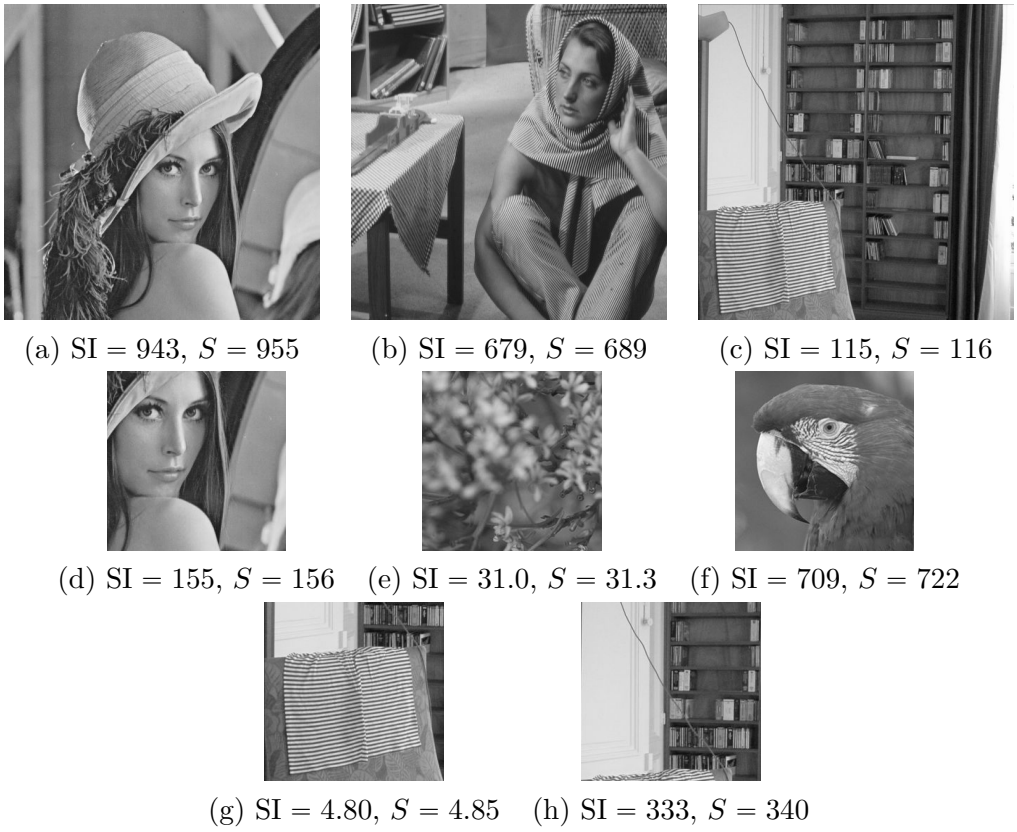
In [Blanchet *et al.* 2008] and [Blanchet & Moisan 2012], experiments show that even if the values assigned to an image by GPC and SI can be quite different, both indices decrease when an image is degraded with noise and/or blur. We here check that the same property holds for the S index. Given an initial image u, we computed $S(\kappa_\rho * u + \sigma n)$ for several values of ρ (the level of blur) and σ (the level of noise), where the Gaussian blur kernel κ_ρ is defined in Fourier domain by

$$\forall \boldsymbol{\xi} \in \Omega, \quad \widehat{\kappa_\rho}(\boldsymbol{\xi}) = \exp\left(-2\pi^2\rho^2\Big(\frac{\xi_1^2}{M^2} + \frac{\xi_2^2}{N^2}\Big)\right) , \tag{4.25}$$

and n is a realization of a white Gaussian noise with unit variance. The obtained values were then averaged over 10 noise realizations, yielding an estimate of the expectation map

$$(\sigma, \rho) \mapsto \mathbb{E}\left(S(\kappa_\rho * u + \sigma n)\right).$$

The resulting blur-noise diagrams are displayed in Fig. 4.6 for the images *Barbara* and *Lighthouse* using a representation by level curves (isovalues of S). We can observe that the S index, like GPC and SI, smoothly decreases with blur and noise, These diagrams are also interesting because they show that S induces an



(a) SI = 943, $S = 955$ (b) SI = 679, $S = 689$ (c) SI = 115, $S = 116$

(d) SI = 155, $S = 156$ (e) SI = 31.0, $S = 31.3$ (f) SI = 709, $S = 722$

(g) SI = 4.80, $S = 4.85$ (h) SI = 333, $S = 340$

Figure 4.5: **Examples of values of** SI **and** S **for some natural images**. One can observe that the values of SI and S are very close, and tend to be small for out-of-focus images like (e) and in the case of a strong high-frequency spectral component (g). Also, the order of magnitude of SI and S grows with the image size (compare the values for the 512×512 images of the first row to those of the 256×256 images of the second row), but it may happen that a sub-part of an image has a larger value of S (or SI) than the whole image, as in (c) and (h).

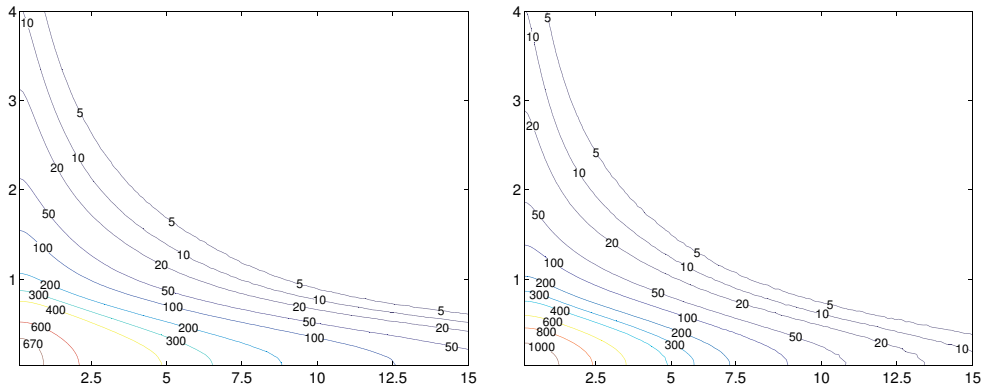


Figure 4.6: **Blur-noise diagrams.** Each diagram displays the isolevel curves of S obtained when a given image (*Barbara* on the left, *Lighthouse* on the right) is degraded with a certain amount of blur (vertical coordinate) and noise (horizontal coordinate). As expected, the largest value of S is obtained in each case at the origin (no blur, no noise), and decreases smoothly (in a rather similar way) as the levels of blur and noise increase.

(image-dependent) equivalence between blur and noise. In the case of *Barbara* for example, we can see that a Gaussian blur of 1.5 pixel is, according to S, equivalent to a Gaussian noise with standard deviation 12.6.

4.4.5 The Dirac paradox

Although it seems that for all natural images the value of S decreases when the image is blurred, we found an exceptional case where the opposite phenomenon happens for a very small level of blur. Indeed, if we consider a Dirac image (a single bright pixel on a constant background) and examine the evolution of S when it is blurred by a Gaussian kernel with parameter ρ (as defined in Equation (4.25)), it happens that S first increases as ρ departs from 0, then decreases when ρ increases further (Fig. 4.7). So far, we have not found a theoretical explanation of this phenomenon. We can remark, however, that it is not really incompatible with the idea that S is linked to image quality and our perception of sharpness: since a Dirac image is aliased, one could consider that a slightly smoother (and hence less aliased) version is sharper (in the sense: more geometrically accurate).

This kind of paradox raises interesting questions linked to the aliasing-ringing-blur trade-off that must face any image reduction (zoom out) algorithm [Blanchet *et al.* 2005]. What is, among the images that represent a single light source (in a sense to be defined), the one that maximizes the value of S? (the experiment reported in Fig. 4.7 proves that this is not a Dirac image). What is the unimodal (increasing, then decreasing) one-dimensional signal that maximizes the value of S? Notice that these questions may be addressed numerically by using the stochastic optimization framework that we describe in Section 4.5.

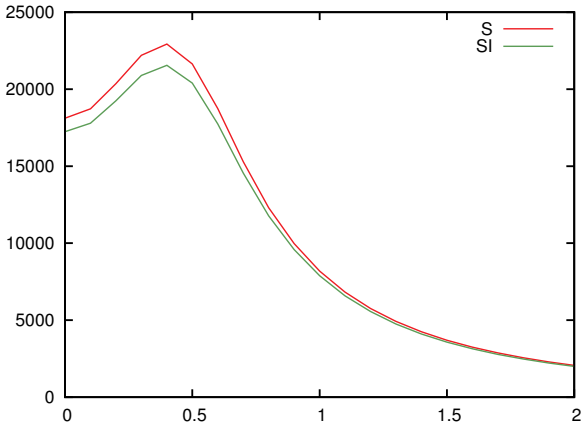


Figure 4.7: **The Dirac paradox.** Evolution of S and SI (vertical axis) for a discrete Dirac image convolved with a Gaussian kernel of width ρ (horizontal axis). Surprisingly, S and SI only decrease after a certain critical value of ρ, which shows that the Gaussian kernel that reaches the maximum value of S is not the Dirac, but a slightly blurrier kernel ($\rho \approx 0.4$ pixels).

4.4.6 Sensitivity to ringing, parametric deconvolution

Suppose that we observe a blurry image v that is the result of the convolution of a clean (unobserved) image u_0 with a Gaussian kernel (4.25), plus some (unknown) noise. We can try to invert this blurring process by using the special case of the Wiener filter obtained with H^1 regularization in a variational setting. Indeed, there is a unique image $u_{\lambda,\rho}$ that minimizes the convex energy

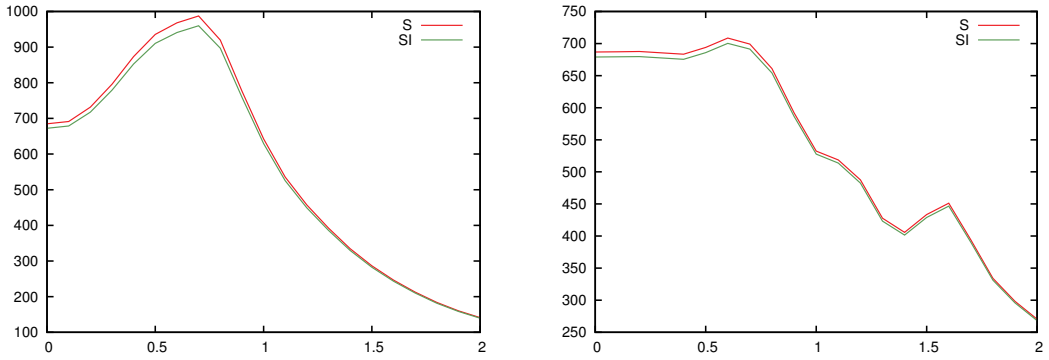
$$\|\kappa_\rho * u - v\|_2^2 + \lambda \|u\|_{H^1}^2 \ , \tag{4.26}$$

and it is explicitly given (thanks to Parseval's formula) in Fourier domain by

$$\forall \boldsymbol{\xi} \in \Omega, \quad \widehat{u_{\lambda,\rho}}(\boldsymbol{\xi}) = \widehat{v}(\boldsymbol{\xi}) \cdot \frac{\widehat{\kappa_\rho}^*(\boldsymbol{\xi})}{|\widehat{\kappa_\rho}|^2(\boldsymbol{\xi}) + \lambda 4\pi^2 (\frac{\xi_1^2}{M^2} + \frac{\xi_2^2}{N^2})} \ . \tag{4.27}$$

This deconvolution method has two parameters λ and ρ. The first one λ, sets the importance of the regularization term $\|u\|_{H^1}^2$ of (4.26) in comparison to the fidelity term $\|\kappa_\rho * u - v\|_2^2$, so that if λ increases, the image is more regularized. The balance between fidelity and regularization is an interesting problem which is encountered in several image processing tasks, but we will not discuss it here. We decided to set $\lambda = 0.01$ which, in our simulations, always gave satisfying results.

The second parameter ρ, however, is critical. If ρ is underestimated, some blur remains; if it is overestimated, spurious oscillations (called *ringing*) appear. As we can see in Fig. 4.8, SI and S can be used in a very simple way to design an automatic procedure that selects an optimal value of ρ (in the sense of the quality of the deconvolved image), because $\mathrm{SI}(u_{\lambda,\rho})$ and $S(u_{\lambda,\rho})$ are maximal for a value of ρ that corresponds very well to the transition between blur and ringing (see Fig. 4.9). This is quite a remarkable property, for classical image quality indices (including



Yale *Barbara*

Figure 4.8: **Blur-ringing trade-offs.** These diagrams plot the values of SI (in green) and S (in red) of the H^1 regularization $u_{\lambda,\rho}$ defined by (4.27) with $\lambda = 0.01$, as functions of the parameter ρ (in pixels) for images *Yale* (left) and *Barbara* (right). SI and S attain their maximum value for a very similar value of ρ, which corresponds in each case to a good trade-off between blur and ringing for these images (see Fig. 4.9).

the metric Q presented below) are not sensitive to ringing artifacts in general (see [Moreno & Calderero 2013]).

4.4.7 Comparison with Zhu and Milanfar's Q metric

In [Zhu & Milanfar 2010], Zhu and Milanfar proposed a sharpness metric Q based on the singular values of the local gradient field of the image. Given a patch p of the image, they consider the two eigenvalues $s_1 \geqslant s_2 \geqslant 0$ of the gradient covariance matrix[2] of p, and define from it the coherence $R(p) = \frac{s_1 - s_2}{s_1 + s_2}$ (linked to the anisotropy of the patch p) and the image content metric $Q(p) = s_1 R(p)$ (which represents the energy in the local dominant orientation). Then, from a set of nonoverlapping patches, a subset $\mathcal{P}$ of anisotropic patches is extracted by thresholding the coherence R, and the metric Q of the whole image is defined as the mean value of $Q(p)$ for $p \in \mathcal{P}$. Notice that when comparing the values of Q on different (possibly noisy, blurred or restored) versions of a particular image, the same set of anisotropic patches must be used. Since $\mathcal{P}$ is extracted from a set of nonoverlapping patches, the computation time for Q is $\mathcal{O}(MN)$.

In particular, Zhu and Milanfar used Q to select an optimal number of iterations in the steering kernel regression (SKR) denoising algorithm of Takeda et al. [Takeda *et al.* 2007]. We reproduced the same experiment and compared the effects of the Q and the S indices in Fig. 4.10. Interestingly enough, the global behavior of both indices is the same: as the level of denoising (that is, the number of iterations in [Takeda *et al.* 2007]) increases, both indices grow, attain a maximal value, then decrease. However, it can be observed that the S index attains its maximum value

[2]The gradient covariance matrix of an image u is the value at $\mathbf{z} = 0$ of the gradient autocorrelation matrix Γ defined in Theorem 4.2.1.

Original | Wiener deconvolution ($\rho = 0.7$) | Wiener deconvolution ($\rho = 1$)

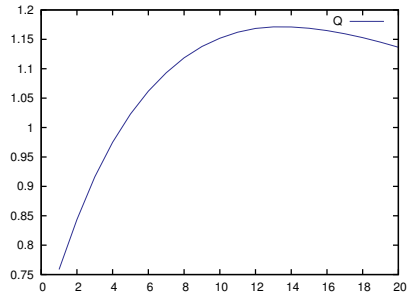
Figure 4.9: **Parametric blind deconvolution using sharpness indices.** On the first row, we can see the original Yale image (left), and two Wiener-H^1 deconvolution results obtained with a kernel width of $\rho = 0.7$ (middle) and $\rho = 1$ (right). Close-up views of these three images are shown on the second row. The value $\rho = 0.7$, which maximizes the sharpness indices SI and S (see Fig. 4.8), corresponds surprisingly well to the desired critical value that rules the transition between blur and ringing.

for a smaller number of iterations (8, versus 14 for Q). This effect is confirmed on other experiments (not displayed here): the S index seems to consider that at some point, the denoising structures left by the SKR algorithm are sharp details and leads to a lower denoising level. This general behavior will be discussed further in Subsection 4.4.8: an image process that creates phase-coherent artifacts may increase the S index.

As the sharpness metrics SI and S, the Q metric is sensitive to blur and noise. However, it is not sensitive to ringing, so that the parametric deconvolution process described in Subsection 4.4.6 cannot be achieved with the Q index, as shown in Fig. 4.11. This is a crucial difference between these two indices.

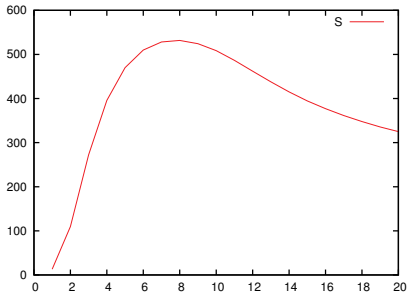
4.4.8 Perceptual sharpness and Visual Summation

Even if GPC, SI and S are sensitive to noise, blur and ringing, we should not forget that they were initially designed to measure phase coherence, and that it only appears that they can be *interpreted* as image quality indices. Thus, contrary to image quality metrics designed on purpose, there is no reason *a priori* that these indices reflect accurately our visual perception of sharpness. An interesting illustration of this is brought by image compression. For example, JPEG compression is known to produce artificial edges (in particular along the boundaries of the 8×8 blocks used for DCT), and as these edges require global phase coherence, one can logically



Evolution of Q

Q optimized, 14 iterations



Evolution of S

S optimized, 8 iterations

Figure 4.10: **Parameter selection in SKR denoising: Q versus S.** The plots on the left report the evolution of Q and S as functions of the number of iterations in the SKR denoising. The input is the image *Cheetah* corrupted by a white Gaussian noise with standard deviation 18. Both indices are able to select an optimal number of iterations, and the resulting images are shown in the middle column (with some close-up views on the right). Note that the residual phase-coherent artifacts left by the SKR algorithm are considered as sharp by the S index, which thus selects a number of iterations that is significantly smaller. In that particular application, the Q metric is best suited to denoise uniform zones, while the S index leads to better texture preservation.

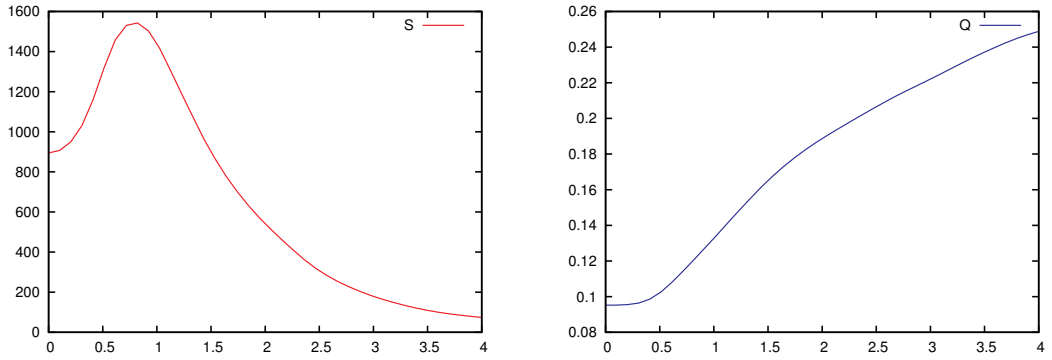


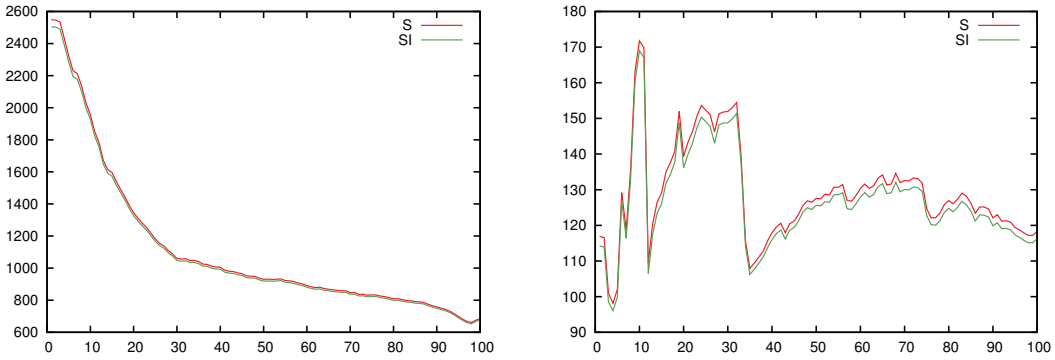
Figure 4.11: **S versus Q**. These diagrams represent the values of S (left) and the values of the metric Q of Zhu and Milanfar (right) computed on the H^1 regularization $u_{\lambda,\rho}$ defined by (4.27) with fixed $\lambda = 0.01$ and varying ρ (horizontal axis), for *Lena* image. One can see that S admits an optimal value whereas Q does not. Therefore, contrary to S, the metric Q cannot be used for parametric blind deblurring, as it does not consider that ringing artifacts decrease image quality. This limitation of Q is studied more deeply in [Liu *et al.* 2013].

expect them to produce high values of GPC, SI and S. Fig. 4.12 confirms this analysis. Note, however, that one could probably adapt the sharpness indices we defined to reflect more accurately the quality of compressed images. One possible solution would be to define the sharpness $S_C(a)$ of a compressed image $a = C(u)$ (here C denotes the compression operator) by the minimum sharpness found among all possible uncompressed versions of a, that is

$$S_C(a) = \min_{v, C(v)=a} S(v).$$

Such a definition could reflect more accurately our perception of image quality, and would in particular satisfy the desirable property $S_C(C(u)) \leqslant S(u)$ (that is, compression cannot increase image quality).

If we follow the idea of relating the sharpness indices GPC, SI and S to perceptual sharpness, the issue of normalization with respect to image size must be addressed. As we saw in Subsection 4.4.3, these indices tend to grow rapidly with the size of an image, which does not really correspond to our visual perception. One possibility to deal with this problem could be to use a "visual summation" principle [Vu & Chandler 2009], and define the overall sharpness of an image as the maximal sharpness of all its fixed-size (say, 32×32) sub-parts. A less extreme variant could be to weight the sharpness of each sub-part by some sort of saliency measure. These solutions would solve the size-dependence issue, and thus probably increase the similarity between the proposed indices and our visual perception of sharpness. However, the obtained indices would be analytically more complicated and probably less stable when addressing restoration problems like the blind deblurring application we consider in the next section.



Barbara *House*

Figure 4.12: **Sharpness indices and JPEG compression.** These diagrams show the evolution of SI (in green) and S (in red) when an image (respectively, *Barbara* on the left, and *House* on the right) is compressed using the JPEG standard. The horizontal scale refers to the JPEG quality parameter. One can see that S and SI do not reflect our perception of image quality in this case: they increase as the image compression rate increases. This phenomenon, due to the artificial phase coherence brought by the image uncompression scheme, could be avoided by considering instead, for a given compressed image, the minimum sharpness of all possible original images.

4.5 An Application to Blind Deblurring

In Subsection 4.4.6, we saw that the S index could be used to select a parameter in a deconvolution process. In this section, we will show that it can drive much more general blind deblurring algorithms. Blind deblurring consists in sharpening an image without knowing precisely the blurring process involved in the image acquisition. We here focus on linear and spatially-invariant blur, which can be modeled by a convolution operator. There is an abundant literature on that subject, and regular advances. We will compare the results we obtain with the method recently proposed by Levin et al. [Levin *et al.* 2011], which can produce impressive results.

To design blind deblurring algorithms based on the S index, we will follow the general scheme proposed in [Blanchet & Moisan 2012]. Let us denote by u_0 the image to recover, by φ an unknown convolution kernel and by n an additive noise. Instead of trying to recover the kernel φ and then invert the image formation process $u = \varphi * u_0 + n$, we will select a restoration kernel k that maximizes $S(k * u)$, the sharpness of the restored image $k * u$. In this framework, k can be interpreted as a regularized inverse of φ that is supposed to mitigate the effects of the noise. Of course, the linearity of the deblurring process is a limitation of this approach, but as we shall see, a well-chosen linear filter may perform surprisingly well compared to more sophisticated non-linear image transforms. Moreover, linearity has several advantages like stability, computational efficiency, and the fact that deconvolution artifacts (and in particular the effect on noise) are much better understood in the

linear case.

4.5.1 Remarks on $k \mapsto S(k * u)$

As mentioned above, the idea underlying the algorithms that will follow is the maximization of the function

$$F_u : k \mapsto S(k * u) \tag{4.28}$$

on a given set $\mathcal{K}$ of deconvolution kernels. Since the function S is quite singular, it is worth discussing the existence of maxima. First, Proposition 4.3.3 ensures that, as soon as the set $\{k * u \ , \ k \in \mathcal{K}\}$ does not contain any image which is constant in the x or y direction, F_u is continuous on $\mathcal{K}$. Moreover, since $S(\lambda k * u) = S(k * u)$ for any $\lambda \neq 0$, the maximization of F_u can be equivalently realized on the bounded set

$$\mathcal{K}' = \{k/\|k\|_2, k \in \mathcal{K}\}.$$

Thus, if $\mathcal{K}'$ is closed (which is an easily achievable condition), F_u has to be maximized on a compact set and we can thus guarantee the existence of a solution. It seems difficult to obtain any guarantee of uniqueness in general (recall that the function S is not concave), but we can at least hope to design algorithms that converge to an interesting local maximum of F_u. Among them, Algorithm 2 below (a direct adaptation of the algorithm proposed in [Blanchet & Moisan 2012]) is very flexible since it can handle various types of kernels, as we will see in the next subsections.

Algorithm 2

- Begin with $k = \delta_0$
- Repeat n times
  - ▷ Define k' from a random perturbation of k
  - ▷ If $S(k' * u) > S(k * u)$ then $k \leftarrow k'$
- Return k and $k * u$

4.5.2 Kernels with compact support

A first interesting case is the set of symmetric kernels with a fixed support, e.g. a 11×11 square. One possible perturbation strategy at each iteration consists in adding a random number uniformly distributed in $[-\alpha, \alpha]$ (say, $\alpha = 0.05$) to a randomly chosen coefficient of the kernel (see [Blanchet & Moisan 2012]). As shown in Fig. 4.13, this simple stochastic algorithm already gives interesting sharpening results. However, it may also lead to failure cases, in particular when the image contains some high-frequency structured textures [Leclaire & Moisan 2013a]. We believe that these failure cases are mostly due to the fact that this set of kernels contains candidates which are not plausible as deconvolution kernels.

Figure 4.13: **Blind deblurring results** obtained by running Algorithm 2 on the set of 11×11 kernels. The original (unprocessed) images are shown on the left column (from top to bottom: *Yale*, *Caps* (cropped), *Room*), and the sharpened images are displayed on the right column. In the first two cases, the output image is sharper than the original one and presents a limited quantity of ringing artifacts. However, the result is not satisfactory for the *Room* image.

4.5.3 Kernel with a radial-unimodal Fourier transform

To cope with the failure cases of fixed support kernels, we suggested in [Leclaire & Moisan 2013a] to consider another class of kernels, whose shape is built in Fourier domain by rotating a radial profile defined by d values

$$r(0) = 1, r(1), r(2), \ldots, r(d-2), r(d-1) = 0.$$

More precisely, we consider the deconvolution kernel k_r defined in Fourier domain by

$$\widehat{k_r}(\xi_1, \xi_2) = L_r \left((d-1) \sqrt{2 \left(\left(\frac{\xi_1}{M} \right)^2 + \left(\frac{\xi_2}{N} \right)^2 \right)} \right) ,$$

where $L_r : [0, d-1] \longrightarrow \mathbb{R}$ denotes the piecewise affine interpolation of r. We also suggested to constrain the discrete profile r to be unimodal, which means that there exists a value m such that

$$\forall i < m,\ r(i+1) \geqslant r(i) \ , \text{ and } \ \forall i \geqslant m,\ r(i+1) \leqslant r(i) \ .$$

The set U of unimodal profiles is rich enough to provide interesting deblurring kernels, and constrained enough to limit distortions in Fourier domain (as large differences in the amplification factor applied to neighboring frequencies tend to produce ringing artifacts). In practice, enforcing the unimodality constraint (by performing a projection on U for example) appeared to be rather inefficient in terms of convergence, and we chose to relax the constraint by incorporating the Euclidean distance[3] $d(r, U)$ between r and the set U in the objective function. We also decided to constrain the profile r to be smooth with the additional term

$$\|r\|_{H^1}^2 = \sum_{i=0}^{d-2} \big(r(i+1) - r(i) \big)^2 \ .$$

Finally, the function to optimize is

$$\mathcal{F}_u(r) = S(k_r * u) - \lambda_{um}\, d(r, U) - \lambda_{reg}\, \|r\|_{H^1}^2 \ , \tag{4.29}$$

where λ_{um} and λ_{reg} are two weighting parameters. The maximization of $\mathcal{F}_u$ is realized with Algorithm 3.

[3] See Appendix 4.D for the numerical computation of $d(r, U)$.

Algorithm 3

- Initialize r with the piecewise affine profile defined by

$$r(0) = 1 \ , \ r(m_{init}) = 2 \ , \text{ and } r(d-1) = 0 \ .$$

- Repeat n times
  - ▷ Pick a random index $i \in \{1, 2, \dots d-2\}$
  - ▷ Draw a uniform random value $\varepsilon \in [-a/2, a/2]$
  - ▷ Set $r' \leftarrow r$, and then $r'(i) \leftarrow r(i) + \varepsilon$
  - ▷ If $\mathcal{F}_u(r') > \mathcal{F}_u(r)$ then $r \leftarrow r'$
- Return r, k_r and $k_r * u$

We observed in practice that Algorithm 3 reached a stable state in less than 10000 iterations (which, on a 512×512 image takes about 4 minutes with a parallel C implementation using a dual-core processor). Although $\mathcal{F}_u$ may have several local maxima, several realizations of the algorithm would always return approximately the same profile r, which demonstrates its stability.

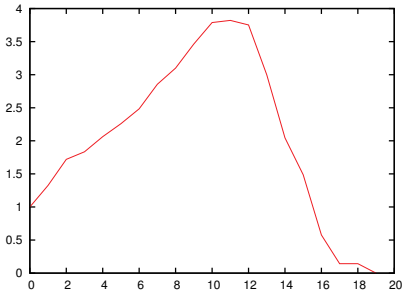
Algorithm 3 involves several constants (λ_{um}, λ_{reg}, d, m_{init}, n, a), but in practice only λ_{reg} is a real parameter. The value d can be set to 20, which achieves a goof trade-off between the dimension of the parameter space and the accuracy of the radial profile. The setup $a = 0.1$ led to an efficient proposition strategy in all cases. As mentioned before, the value $n = 10000$ seems to be sufficient for convergence, in the sense that the average rate of convergence $\left\| \frac{r_{new} - r_{old}}{r_{old}} \right\|_\infty$ was in general less than 10^{-3} after 10000 iterations. To force r to be as close to unimodal as possible, we affected to λ_{um} a high value (10000 in our experiments); we could have made it grow to $+\infty$ in the last iterations. As concerns m_{init} (the initial mode index), we observed that the different possibilities of initialization (any integer between 1 and $d-2$) could lead to two (or three in a few cases) different radial profiles. A systematic strategy would be to try all these indices and select the one leading to the maximum value of $\mathcal{F}_u$. In practice, we observed that this maximum value was obtained for an index $m_{init} \in [d/4, 3d/4]$. Besides, in the case where 2 or 3 different radial profiles were obtained (depending on the initialization), we observed that they lead to similar deblurring results. For the sake of simplicity, all the experiments shown in this paper were run with $m_{init} = d/4$ (that is, 5).

In Fig. 4.14, we show some results obtained with Algorithm 3 (for $\lambda_{reg} = 0$) on the original images *Yale* and *Barbara* (no blur or noise added). In both cases, the resulting image is clearly sharper than the original one and the edges are nicely enhanced, even on the image *Barbara* which is a difficult case for it contains high-frequency textures.

To assess more precisely the performances of Algorithm 3, we also ran it on artificially degraded images. We transformed each original image u_0 into a blurry


Original

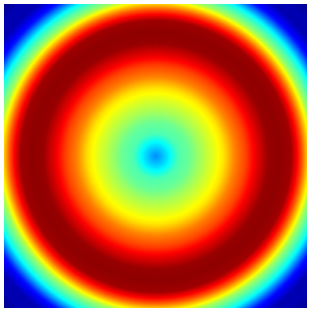

Original, close-up



Radial profile (r)

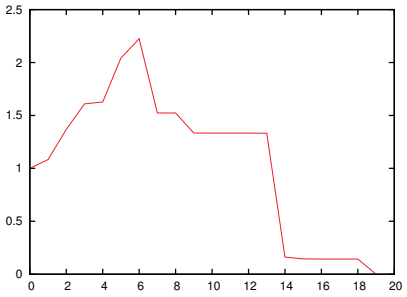

Deblurred


Deblurred, close-up


Spectrum of deconvolution kernel $(\widehat{k_r})$


Original


Original, close-up



Radial profile (r)


Deblurred


Deblurred, close-up

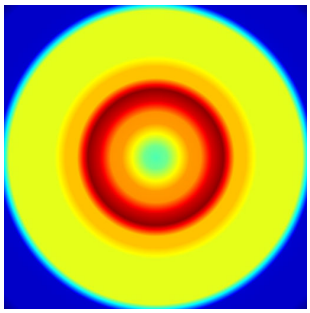

Spectrum of deconvolution kernel $(\widehat{k_r})$

Figure 4.14: **Blind deblurring of unprocessed images.** Algorithm 3 is applied (with $\lambda_{reg} = 0$, and $n = 10000$ iterations) to the images *Yale* (top 2 rows) and *Barbara* (bottom rows). In each case, the obtained radial profile r is displayed, as well as the Fourier transform of the corresponding deconvolution kernel k_r. It is interesting to observe the stability of the proposed algorithm: the deblurred images are much sharper than the original ones, but do not present ringing artifacts or excessive noise amplification. Notice also how the deconvolution kernel adapts itself to each image, leading, in the case of *Barbara*, to a quite irregular profile.

and noisy image

$$u = \kappa_1 * u_0 + n \,, \tag{4.30}$$

where κ_1 is the Gaussian kernel (4.25) obtained for $\rho = 1$ and n is a realization of a Gaussian white noise with standard deviation $\sigma = 1$. This setup allowed us to build two oracle deblurring filters: the Wiener filter (4.27) associated to the (supposedly unknown) kernel κ_1, and the oracle radial filter minimizing the expected l^2 risk, defined by

$$k_0 = \arg\min_{k_r} \mathbb{E}\left(\|u_0 - k_r * (\kappa_1 * u_0 + W)\|^2\right) \,, \tag{4.31}$$

where W is a white Gaussian noise with variance $\sigma^2 = 1$ and the arg min is taken over all kernels k_r obtained from an arbitrary radial profile r with d points[4].

A comparison of the effect of these filters (including Algorithm 3 with several values of the λ_{reg} parameter) is shown on *Parrots* image in Fig. 4.15. We can see that Algorithm 3 manages to find a kernel that is close to the Wiener filter associated to the true level of blur ($\rho = 1$). The oracle output reveals slightly more details, but also leaves on the image some undesirable structured noise (which is not costly for the l^2 risk function that it optimizes). The comparison with [Levin *et al.* 2011] is also interesting: compared to Algorithm 3, it manages to clean uniform zones better, but tends to reveal less details in more complex areas (geometric structures or textures). In terms of PSNR (which use is questionable since the original image itself could be noisy and blurry), Algorithm 3 performs better (for $\lambda_{reg} = 10$) that [Levin *et al.* 2011] and the Wiener oracle, but does not attain the ultimate performance given by the oracle radial filter.

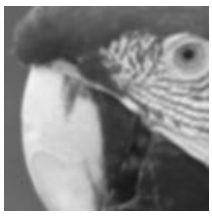
To end this section, we now discuss the influence of the regularity parameter λ_{reg}. As expected, increasing λ_{reg} tends to smooth the radial profile r (see Fig. 4.15 and 4.16). One can also see that this regularity prior constrains the overall energy of the kernel, so that when λ_{reg} increases, the kernel values tend to decrease. The *Room* image (see Fig. 4.16) is difficult to process because it contains different high-frequency textures that are likely to produce ringing artifacts. In this particular case, the regularity constraint is mandatory: the disappointing result obtained for $\lambda_{reg} = 0$ is greatly improved for $\lambda_{reg} = 100$. For the other images we considered (and that are not displayed here), we noticed that the choice $\lambda_{reg} = 100$ always led to visually satisfying results, and $\lambda_{reg} \in [0, 25]$ gave even better results with images that were not too prone to ringing artifacts.

4.6 Perspectives

In this chapter, we discussed and compared the phase coherence indices GPC, SI and S, and provided some mathematical results as well as several experiments demonstrating their usefulness for no-reference image quality assessment and blind deblurring. The more explicit and simple variants SI and S are clearly an improvement over the original GPC, but many questions remain. The decrease of these

[4] The computation of this oracle kernel is detailed in Appendix 4.E.


Blurred and noisy input
PSNR = 30.5, S = 140

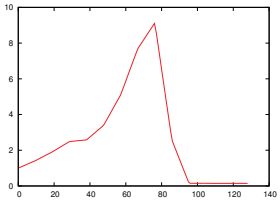

Close-up


Levin et al. (close-up)

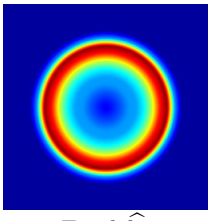

Levin et al.
PSNR = 27.9, S = 605


Deblurred ($\lambda_{reg} = 0$)
PSNR = 24.5, S = 440


Close-up

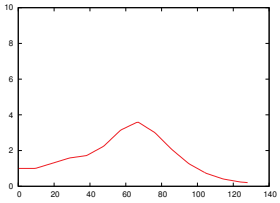


Radial profile

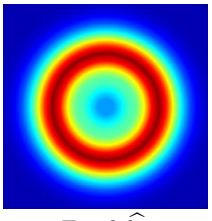

Final $\widehat{k_r}$


Deblurred ($\lambda_{reg} = 10$)
PSNR = 34.2, S = 394


Close-up

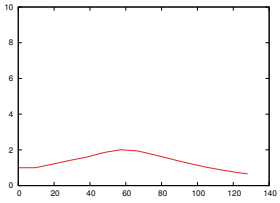


Radial profile

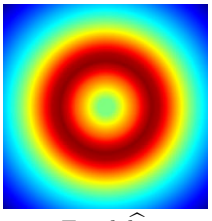

Final $\widehat{k_r}$


Deblurred ($\lambda_{reg} = 100$)
PSNR = 33.8, S = 300


Close-up

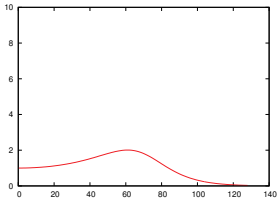


Radial profile

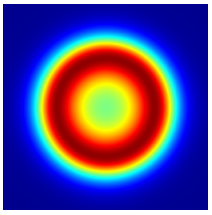

Final $\widehat{k_r}$


Wiener oracle
PSNR = 33.7, S = 316


Close-up

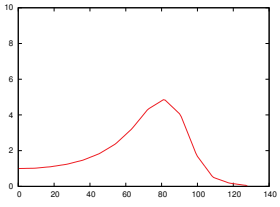


Wiener radial profile


Wiener oracle filter


Radial oracle
PSNR = 35.6, S = 370


Close-up



Oracle radial profile

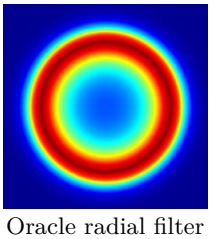

Oracle radial filter

Figure 4.15: **Blind deblurring of a blurry and noisy version of *Parrots*.** The first row displays the degraded image (used as input), and the deblurred image obtained with Levin et al. algorithm [Levin *et al.* 2011]. Each other row is devoted to a different linear algorithm based on a radial kernel (in each case, the radial profile and the Fourier transform of the kernel are displayed). The PSNR values are computed with respect to the original *Parrots* image. The result obtained with Levin et al. algorithm is cleaner in uniform regions, but slightly less detailed than the one obtained with Algorithm 3 when $\lambda_{reg} = 10$. Notice also the similarity between the filter obtained with $\lambda_{reg} = 10$ and the Wiener oracle filter. Algorithm 3 was used with 10000 iterations.

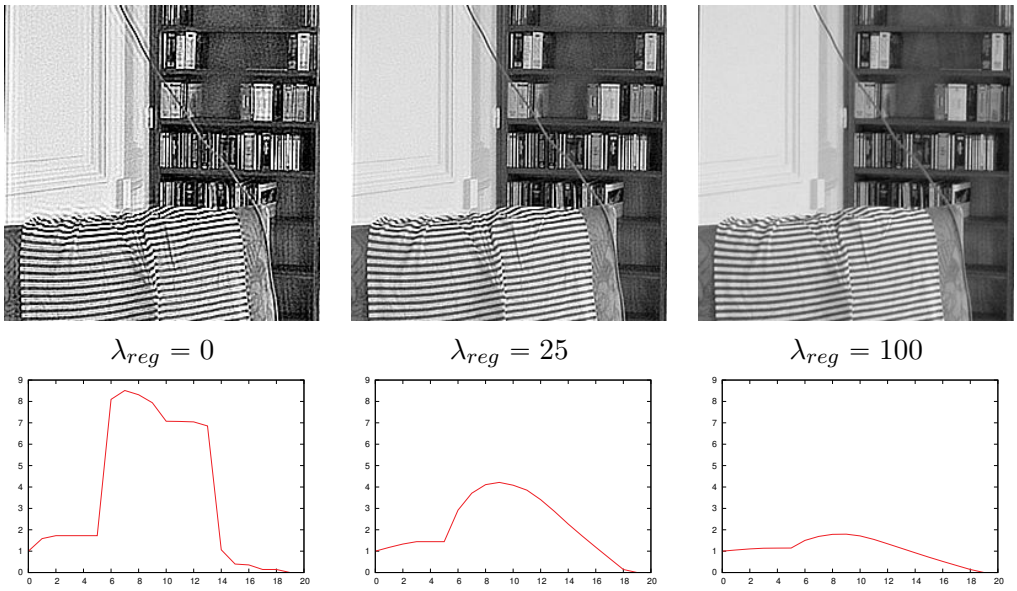


Figure 4.16: **Blind deblurring of the original *Room* image** for three different levels of regularization of the Fourier profile. On the top row, we display a close-up of the result of the blind deblurring Algorithm 3, which selects (and applies) an optimal radial convolution filter (the corresponding radial profile is shown on the bottom row in each case). The strong ringing artifacts that appear for $\lambda_{reg} = 0$ (left column) are greatly attenuated for $\lambda_{reg} = 25$ (middle) and disappear almost completely for $\lambda_{reg} = 100$. On this kind of images presenting a strong high-frequency content (here, the stripes of the piece of clothing in particular), the parameter λ_{reg} plays a crucial role. Algorithm 3 was used with 10000 iterations.

indices with respect to noise and blur is easy to check numerically, but a mathematical proof is still to be established. Also, it would be interesting to understand, from an analytical (non-probabilistic) point of view, why the formulae obtained for SI and S are efficient for image quality assessment and blind deblurring. This could be a way to design non-probabilistic variants, very different from classical analytical regularizers like TV or more generally sparsity-promoting priors. The optimization of S also brings interesting issues, and it seems very likely that the simple iterative stochastic optimization we proposed could be greatly improved, which should increase even further the attractiveness of these indices.

Software resources

Source codes to compute the GPC, SI and S metrics and images files used in the experiments are freely available on the web page
http://www.mi.parisdescartes.fr/~moisan/sharpness/

Appendices

4.A Estimation of the mean TV of a RPN

We saw in Theorem 4.2.1 (Equation (4.10)) that

$$\mathbb{E}(\mathrm{TV}(u * W)) = (\alpha_x + \alpha_y)\sqrt{\frac{2}{\pi}}\sqrt{MN}\,. \tag{4.32}$$

The right-hand term of (4.32) appears to be a good approximation of $\mathbb{E}(\mathrm{TV}(u_\psi))$, that is, the mean TV in the RPN model. As noticed in [Blanchet & Moisan 2012], for most images the relative error is around 1% or below. In this Appendix, we will exhibit an upper bound of the absolute difference.

With the definition of TV, one can write

$$\mathbb{E}(\mathrm{TV}(u_\psi)) = \sum_{\mathbf{x}\in\Omega} \mathbb{E}|\partial_x \dot{u}_\psi(\mathbf{x})| + \mathbb{E}|\partial_y \dot{u}_\psi(\mathbf{x})|$$

so that it is sufficient to show that $\mathbb{E}|\partial_x \dot{u}_\psi(\mathbf{x})| \approx \alpha_x\sqrt{\frac{2}{\pi MN}}$ for each $\mathbf{x} \in \Omega$. This will follow from a Gaussian approximation of $\partial_x \dot{u}_\psi(\mathbf{x})$ which implies

$$\mathbb{E}(|\partial_x \dot{u}_\psi(\mathbf{x})|) \approx \sqrt{\frac{2}{\pi}}\sqrt{\mathbb{E}((\partial_x \dot{u}_\psi(\mathbf{x}))^2)} \tag{4.33}$$

(notice that the equality holds for a zero-mean Gaussian r.v., as shown by Lemma 4.C.1 of Appendix 4.C).

With the Fourier reconstruction formula, one can write that for all $\mathbf{x} \in \Omega$,

$$\partial_x \dot{u}_\psi(\mathbf{x}) = \frac{1}{MN}\sum_{\boldsymbol{\xi}\in\Omega} |\hat{u}(\boldsymbol{\xi})| e^{i\psi(\boldsymbol{\xi})} e^{i\langle \mathbf{x},\boldsymbol{\xi}\rangle}(e^{\frac{2i\pi x_1}{M}} - 1)\,. \tag{4.34}$$

For any $\mathbf{x} \in \Omega$, the set $(e^{i\psi(\boldsymbol{\xi})}e^{i\langle \mathbf{x},\boldsymbol{\xi}\rangle})_{\boldsymbol{\xi}\in\Omega}$ is a random phase field. It follows that the r.v. $|\partial_x \dot{u}_\psi(\mathbf{x})|$ are identically distributed, but they are not independent *a priori*. This is why we cannot use the central limit theorem directly on the sum $\sum_{\mathbf{x}\in\Omega} |\partial_x \dot{u}_\psi(\mathbf{x})|$. Instead we will use a Gaussian approximation of each $\partial_x \dot{u}_\psi(\mathbf{x})$ in order to derive a bound for the Gaussian approximation of $\sum_{\mathbf{x}\in\Omega} |\partial_x \dot{u}_\psi(\mathbf{x})|$.

The Gaussian approximation of $\partial_x \dot{u}_\psi(\mathbf{x})$ will be precised with a Berry-Esseen theorem. First, to cope with the Hermitian dependence, we have to introduce a subset Ω_+ of Ω that contains exactly one point in each pair of symmetrical points, that is, such that

$$\Omega \backslash \{0, \boldsymbol{\eta}_x, \boldsymbol{\eta}_y, \boldsymbol{\eta}_{xy}\} = \Omega_+ \cup (-\Omega_+)$$

and the union is disjoint. To make the following proof lighter, we will assume that if they exist, the Nyquist coefficients $\hat{u}(\boldsymbol{\eta}_x)$, $\hat{u}(\boldsymbol{\eta}_{xy})$, and $\hat{u}(\boldsymbol{\eta}_y)$ are equal to zero (in general, in natural images these coefficients are very small). Then we can write

$$u_\psi(\mathbf{x}) = |\hat{u}(\mathbf{0})|(-1)^{\varepsilon_0} + \frac{1}{MN} \sum_{\boldsymbol{\xi}\in\Omega_+} 2|\hat{u}(\boldsymbol{\xi})| \cos(\psi(\boldsymbol{\xi}) + \langle \mathbf{x}, \boldsymbol{\xi}\rangle) \ ,$$

and therefore

$$u_\psi(x_1 + 1, x_2) - u_\psi(x_1, x_2) = \frac{1}{MN} \sum_{\boldsymbol{\xi}\in\Omega_+} X_{\boldsymbol{\xi}} \ ,$$

where we set for all $\boldsymbol{\xi} \in \Omega_+$,

$$\begin{aligned} X_{\boldsymbol{\xi}} &= 2|\hat{u}(\boldsymbol{\xi})| \left(\cos\left(\psi(\boldsymbol{\xi}) + \langle \mathbf{x}, \boldsymbol{\xi}\rangle + \frac{2\pi\xi_1}{M}\right) - \cos\left(\psi(\boldsymbol{\xi}) + \langle \mathbf{x}, \boldsymbol{\xi}\rangle\right)\right) \\ &= -4|\hat{u}(\boldsymbol{\xi})| \sin\left(\psi(\boldsymbol{\xi}) + \langle \mathbf{x}, \boldsymbol{\xi}\rangle + \frac{\pi\xi_1}{M}\right) \sin\left(\frac{\pi\xi_1}{M}\right) . \end{aligned}$$

Since the $X_{\boldsymbol{\xi}}$ are independent and centered r.v., we can apply the following generalization of Berry-Esseen Theorem (for non identically distributed r.v.):

Theorem 4.A.1 (Berry-Esseen, 1942)**.** *Let $X_1, \ldots, X_n$ be independent and centered r.v. in L^3. Let us denote $\sigma_i^2 = \mathbb{E}(X_i^2)$ and $\rho_i = \mathbb{E}(|X_i|^3)$. Let F_n be the cumulative distribution function of*

$$\frac{X_1 + \ldots + X_n}{(\sigma_1^2 + \ldots + \sigma_n^2)^{1/2}} \ .$$

Then there exists a positive universal constant C_0 such that

$$\forall t \in \mathbb{R}, \quad |F_n(t) - \mathbb{P}(Y \leqslant t)| \leqslant C_0 \psi_0$$

where $Y \sim \mathcal{N}(0,1)$ and $\psi_0 = \left(\sum_{i=1}^n \sigma_i^2\right)^{-3/2} \left(\sum_{i=1}^n \rho_i\right)$.

Concerning the value of C_0, some recent papers (e.g. [Shevtsova 2010]) have shown that the best constant C_0 is below 0.56.

Let us apply this theorem to the r.v. $X_{\boldsymbol{\xi}}$, $\boldsymbol{\xi} \in \Omega_+$. Remark that if the r.v. U is uniformly distributed on $[0, 2\pi]$, then $\mathbb{E}(\sin^2(U)) = \frac{1}{2}$ and $\mathbb{E}(|\sin(U)|^3) = \frac{4}{3\pi}$. Thus, we have for all $\boldsymbol{\xi} \in \Omega_+$,

$$\sigma_{\boldsymbol{\xi}}^2 := \mathbb{E}(X_{\boldsymbol{\xi}}^2) = 8|\hat{u}(\boldsymbol{\xi})|^2 \sin^2\left(\frac{\pi\xi_1}{M}\right) \ ,$$

$$\rho_{\boldsymbol{\xi}} := \mathbb{E}(|X_{\boldsymbol{\xi}}|^3) = \frac{4^4}{3\pi}|\hat{u}(\boldsymbol{\xi})|^3 \left|\sin\left(\frac{\pi\xi_1}{M}\right)\right|^3 .$$

Consequently,

$$\begin{aligned}\sum_{\boldsymbol{\xi}\in\Omega_+} \sigma_{\boldsymbol{\xi}}^2 &= \sum_{\boldsymbol{\xi}\in\Omega_+} 8|\hat{u}(\boldsymbol{\xi})|^2 \sin^2\left(\frac{\pi\xi_1}{M}\right) \\ &= \sum_{\boldsymbol{\xi}\in\Omega} 4|\hat{u}(\boldsymbol{\xi})|^2 \sin^2\left(\frac{\pi\xi_1}{M}\right) \\ &= \sum_{\boldsymbol{\xi}\in\Omega} |\hat{u}(\boldsymbol{\xi})|^2 \left|e^{\frac{2i\pi\xi_1}{M}} - 1\right|^2 = \|\widehat{\partial_x \dot{u}}\|_2^2 = MN\|\partial_x \dot{u}\|_2^2 ,\end{aligned}$$

and

$$\sum_{\boldsymbol{\xi}\in\Omega_+} \rho_{\boldsymbol{\xi}} = \frac{4^4}{3\pi} \sum_{\boldsymbol{\xi}\in\Omega_+} |\hat{u}(\boldsymbol{\xi})|^3 \left|\sin\left(\frac{\pi\xi_1}{M}\right)\right|^3 = \frac{128}{3\pi}\|\widehat{\partial_x \dot{u}}\|_3^3 .$$

Hence, noticing that

$$\frac{1}{\sqrt{MN}\|\partial_x \dot{u}\|_2} \sum_{\boldsymbol{\xi}\in\Omega_+} X_{\boldsymbol{\xi}} = \frac{\sqrt{MN}}{\|\partial_x \dot{u}\|_2}\, \partial_x \dot{u}_\psi(\mathbf{x}) ,$$

and setting

$$\psi_0 = \frac{K(u)}{(MN)^{3/2}} \quad \text{with} \quad K(u) = \frac{\frac{128}{3\pi}\|\widehat{\partial_x \dot{u}}\|_3^3}{\|\partial_x \dot{u}\|_2^3} ,$$

Theorem 4.A.1 ensures that for all $t \in \mathbb{R}$,

$$\left|\mathbb{P}\left(\frac{\sqrt{MN}}{\|\partial_x \dot{u}\|_2}\, \partial_x \dot{u}_\psi(\mathbf{x}) \geqslant t\right) - \mathbb{P}(Y \geqslant t)\right| \leqslant \frac{C_0 K(u)}{(MN)^{3/2}} . \tag{4.35}$$

Now, we write

$$\mathbb{E}\left(\frac{\sqrt{MN}}{\|\partial_x \dot{u}\|_2}\, |\partial_x \dot{u}_\psi(\mathbf{x})|\right) = \int_0^{+\infty} \mathbb{P}\left(\frac{\sqrt{MN}}{\|\partial_x \dot{u}\|_2}\, |\partial_x \dot{u}_\psi(\mathbf{x})| \geqslant t\right) dt ,$$

$$\text{and } \mathbb{E}(Y) = \int_0^{+\infty} \mathbb{P}(Y \geqslant t)\, dt ,$$

and we split the integral into two parts: $\int_0^{+\infty} = \int_0^A + \int_A^{+\infty}$. Inequality (4.35) can be integrated between 0 and A to give an upper bound of $\int_0^A$, whereas the tail $\int_A^{+\infty}$ can be treated using Bienaymé-Tchebitchev inequality:

$$\begin{aligned}\mathbb{P}\left(\frac{\sqrt{MN}}{\|\partial_x \dot{u}\|_2}\, |\partial_x \dot{u}_\psi(\mathbf{x})| \geqslant t\right) &\leqslant \frac{1}{t^2 MN\|\partial_x \dot{u}\|_2^2}\mathbb{E}\left(\sum_{\boldsymbol{\xi}\in\Omega_+} X_{\boldsymbol{\xi}}\right)^2 \\ &= \frac{1}{t^2 MN\|\partial_x \dot{u}\|_2^2} \sum_{\boldsymbol{\xi}\in\Omega_+} \sigma_{\boldsymbol{\xi}}^2 = \frac{1}{t^2} .\end{aligned}$$

Putting the two terms together, we have for all $A > 0$,

$$\left| \mathbb{E}\left(\frac{\sqrt{MN}}{\|\partial_x \dot{u}\|_2} \, |\partial_x \dot{u}_\psi(\mathbf{x})| \right) - \mathbb{E}(|Y|) \right| \leqslant \frac{2C_0 K(u)}{(MN)^{3/2}} A + \frac{2}{A} \, ,$$

and then, choosing the best A,

$$\left| \mathbb{E}\left(\frac{\sqrt{MN}}{\|\partial_x \dot{u}\|_2} \, |\partial_x \dot{u}_\psi(\mathbf{x})| \right) - \sqrt{\frac{2}{\pi}} \right| \leqslant 4 \frac{\sqrt{C_0 K(u)}}{(MN)^{3/4}} \, .$$

Therefore, for all $\mathbf{x}$,

$$\left| \mathbb{E}\left(\sqrt{MN} \, |\partial_x \dot{u}_\psi(\mathbf{x})| \right) - \alpha_x \sqrt{\frac{2}{\pi}} \right| \leqslant \frac{C_x(u)}{(MN)^{3/4}} \, ,$$

$$\text{where } \; C_x(u) = 4\sqrt{C_0} \sqrt{\frac{\frac{128}{3\pi} \|\widehat{\partial_x \dot{u}}\|_3^3}{\|\partial_x \dot{u}\|_2}} \, .$$

Recalling that $\alpha_x = \|\partial_x \dot{u}\|_2$, one has

$$\begin{aligned} &\left| \mathbb{E}(\|\partial_x \dot{u}_\psi\|_1) - \alpha_x \sqrt{MN} \sqrt{\frac{2}{\pi}} \right| \\ \leqslant &\frac{1}{\sqrt{MN}} \sum_{\mathbf{x} \in \Omega} \left| \mathbb{E}\left(\sqrt{MN} \, |\partial_x \dot{u}_\psi(\mathbf{x})| \right) - \alpha_x \sqrt{\frac{2}{\pi}} \right| \\ \leqslant &\frac{1}{\sqrt{MN}} \sum_{\mathbf{x} \in \Omega} \frac{C_x(u)}{(MN)^{3/4}} \, , \end{aligned}$$

and thus,

$$\left| \mathbb{E}(\|\partial_x \dot{u}_\psi\|_1) - \alpha_x \sqrt{MN} \sqrt{\frac{2}{\pi}} \right| \leqslant \frac{C_x(u)}{(MN)^{1/4}} \, . \tag{4.36}$$

Finally, we obtain the following

Theorem 4.A.2. *If ψ is a discrete random phase field, then*

$$\left| \mathbb{E}(\mathrm{TV}(u_\psi)) - (\alpha_x + \alpha_y) \sqrt{MN} \sqrt{\frac{2}{\pi}} \right| \leqslant \frac{C_x(u) + C_y(u)}{(MN)^{1/4}} \, ,$$

$$\textit{where } \; \forall a \in \{x, y\}, \quad C_a(u) = 32 \sqrt{\frac{2C_0}{3\pi}} \sqrt{\frac{\|\widehat{\partial_a \dot{u}}\|_3^3}{\|\partial_a \dot{u}\|_2}} \, .$$

Theorem 4.A.2 provides an explicit bound on the absolute error between the mean TV of a RPN and the exact formula (4.32) obtained for the associated Gaussian field, but this error bound depends on the considered image and all terms tend to increase with the image size. We can write a normalized inequality by dividing (4.36) by $\alpha_x \sqrt{2MN/\pi}$, so that

$$\left| \frac{\mathbb{E}(\|\partial_x \dot{u}_\psi\|_1)}{\alpha_x \sqrt{MN}} \sqrt{\frac{\pi}{2}} - 1 \right| \leqslant c_x(u) \, , \tag{4.37}$$

where the relative error bound is now

$$c_x(u) := \frac{32}{(MN)^{3/4}}\sqrt{\frac{C_0}{3}}\sqrt{\frac{\|\widehat{\partial_x \dot{u}}\|_3^3}{\|\partial_x \dot{u}\|_2^3}} = 32\sqrt{\frac{C_0}{3}}\sqrt{\frac{\|\widehat{\partial_x \dot{u}}\|_3^3}{\|\widehat{\partial_x \dot{u}}\|_2^3}}$$

(of course, one would obtain a similar inequality for the y component).

Taking $C_0 = 0.56$, one can compute values of c_x for different natural images. For example, $c_x(u) \approx 1.025$ for the 512×512 *Lena* image, while $c_x(u) \approx 0.337$ for the 13 Mpixels *Lotriver* image[5]. The bound is quite useless for *Lena*, and still far from sharp for *Lotriver* (numerical computations seem to indicate that the true values of the left-hand term of (4.37) are below 10^{-4} for these two images).

Even if it does not provide an accurate error bound, Theorem 4.A.2 remains interesting because it indicates that (4.32) provides the correct asymptotical estimate of the mean TV of a RPN when the image size tends to infinity. Indeed, it has been known for a long time that natural images statistically exhibit a power-law Fourier spectrum (see [Deriugin 1956] and other references in [Ruderman 1994]), that is,

$$|\hat{u}(\boldsymbol{\xi})| \propto |\boldsymbol{\xi}|^{-\alpha} \tag{4.38}$$

in average, where α is a bit larger than 1 in general. Using (4.38) in the expression of c_x above, one easily obtains that for a $R \times R$ image, $c_x \propto R^{-1/2}$ as $R \longrightarrow \infty$, provided that $\alpha < 5/3$. This suggests that the bound c_x tends to decrease to 0 when the size of the considered image increases.

4.B Gaussian approximation of TV(W)

We would like to prove that TV(u_ψ) and TV($u * W$) approximately (or asymptotically) follow Gaussian distributions. Unfortunately, as we already said in the previous Appendix, we cannot apply a classical central limit theorem because the r.v. appearing in the TV formula are not independent. These dependencies introduce a lot of difficulties and this is why we shall here focus on a much simpler problem, that is, the asymptotical distribution of TV(W) (which is the TV of the Gaussian model in the particular case $u = \delta_0$).

Proposition 4.B.1. *Let $(\Omega_n)_{n\geqslant 0}$ be a sequence of rectangular domains of $\mathbb{Z}^2$ such that $|\Omega_n| \to \infty$ when n tends to ∞, and let $(W_n(\mathbf{x}))_{\mathbf{x}\in\Omega_n}$ be a set of i.i.d. r.v. with distribution $\mathcal{N}\left(0, |\Omega_n|^{-1/2}\right)$. Then one has*

$$\mathrm{TV}(W_n) - \mathbb{E}(\mathrm{TV}(W_n)) \xrightarrow{d} \mathcal{N}(0, \sigma^2)\ , \quad \textit{where}$$

$$\mathbb{E}(\mathrm{TV}(W_n)) = \frac{4|\Omega_n|^{1/2}}{\sqrt{\pi}} \quad \textit{and} \quad \sigma^2 = \frac{8}{\pi}\left(\omega(1) + 6\cdot\omega\Big(\frac{1}{2}\Big)\right)\ .$$

To prove this result, we will use the central limit theorem given in [Janson 1988], which applies to a set of r.v. whose dependencies are controlled through their dependency graph.

[5] This image is available on the web site http://www.mi.parisdescartes.fr/~moisan/sharpness/

Definition 4.B.1 ([Janson 1988])**.** A graph Γ is a dependency graph for a set of r.v. if the following two conditions are satisfied:

1. There exists a one-to-one correspondence between the r.v. and the vertices of the graph.
2. If V_1 and V_2 are two disjoint sets of vertices of Γ such that no edge of Γ has one endpoint in V_1 and the other in V_2, then the corresponding sets of r.v. are independent.

Now we can recall the

Theorem 4.B.1 (Janson [Janson 1988])**.** *Suppose, for each integer n, that $(X_{n,i})_{i=1,\dots,N_n}$ is a set of r.v. satisfying $|X_{n,i}| \leqslant A_n$ a.s. for all i. Suppose further that Γ_n is a dependency graph for this set and let M_n be the maximal degree[6] of Γ_n (unless Γ_n has no edges at all, in which case we set $M_n = 1$). Let $S_n = \sum_{i=1}^{N_n} X_{n,i}$ and $\sigma_n^2 = \mathrm{Var}(S_n)$. If there exists an integer m such that*

$$\left(\frac{N_n}{M_n}\right)^{1/m} \frac{M_n A_n}{\sigma_n} \to 0 \quad \text{as} \quad n \to \infty\,, \tag{4.39}$$

then $\dfrac{S_n - \mathbb{E}(S_n)}{\sigma_n} \to \mathcal{N}(0,1)$ in distribution as $n \to \infty$.

First, we will clarify the remark following this theorem in [Janson 1988]. It states that we can replace the boundedness hypothesis

$$\forall n, \quad \forall i, \quad |X_{n,i}| \leqslant A_n \quad \text{a.s.}$$

$$\text{by} \quad \frac{M_n}{\sigma_n^2} \sum_{i=1}^{N_n} \mathbb{E}(X_{n,i}^2 \mathbf{1}_{|X_{n,i}|>A_n}) \to 0 \quad \text{as } n \to \infty\,. \tag{4.40}$$

Indeed, assume that (4.40) is true. We use the truncation argument suggested in [Janson 1988] and set

$$X_{n,i}^T = X_{n,i}\, \mathbf{1}_{|X_{n,i}|\leqslant A_n}\,,$$

$$S_n^T = \sum_{i=1}^{N_n} X_{n,i}^T\,, \quad \text{and} \quad (\sigma_n^T)^2 = \mathrm{Var}(S_n^T)\,.$$

It is clear that the variables $X_{n,i}^T$ have the same dependency degree than the $X_{n,i}$. We will see that (4.39) is still true for σ_n^T so that Janson's Theorem will give

$$\frac{S_n^T - \mathbb{E}(S_n^T)}{\sigma_n^T} \xrightarrow{d} \mathcal{N}(0,1)\,.$$

But first let us explain how we control the residual sum. One can write

$$\frac{S_n - \mathbb{E}(S_n)}{\sigma_n} - \frac{S_n^T - \mathbb{E}(S_n^T)}{\sigma_n} = \frac{1}{\sigma_n} \sum_{i=1}^{N_n} \Big(X_{n,i}\, \mathbf{1}_{|X_{n,i}|>A_n} - \mathbb{E}(X_{n,i}\, \mathbf{1}_{|X_{n,i}|>A_n}) \Big)\,.$$

[6] We recall that the maximal degree of a graph is the maximal number of edges incident to a single vertex.

For a fixed n, setting

$$T_i = X_{n,i}\,\mathbf{1}_{|X_{n,i}|>A_n} - \mathbb{E}(X_{n,i}\,\mathbf{1}_{|X_{n,i}|>A_n})$$

(which again have a dependency degree smaller than M_n) and writing $i \sim j$ if T_i and T_j are not independent, one can write

$$\begin{aligned}
\mathbb{E}\left(\left(\sum_i T_i\right)^2\right) &= \sum_{i,j}\mathbb{E}(T_iT_j)\\
&= \sum_i\sum_{j\sim i}\mathbb{E}(T_iT_j)\\
&\leqslant \frac{1}{2}\sum_i\sum_{j\sim i}\mathbb{E}(T_i^2)+\mathbb{E}(T_j^2)\\
&= \frac{1}{2}\sum_i\sum_{j\sim i}\mathbb{E}(T_i^2)+\frac{1}{2}\sum_j\sum_{i\sim j}\mathbb{E}(T_j^2)\\
&\leqslant (M_n+1)\sum_i\mathbb{E}(T_i^2)\\
&\leqslant 2M_n\sum_i\mathbb{E}(T_i^2)\,,
\end{aligned}$$

which gives

$$\begin{aligned}
&\mathbb{E}\left(\frac{1}{\sigma_n^2}\left(\sum_{i=1}^{N_n}X_{n,i}\,\mathbf{1}_{|X_{n,i}|>A_n}-\mathbb{E}(X_{n,i}\,\mathbf{1}_{|X_{n,i}|>A_n})\right)^2\right)\\
\leqslant& 2\,\frac{M_n}{\sigma_n^2}\sum_{i=1}^{N_n}\mathrm{Var}(X_{n,i}\,\mathbf{1}_{|X_{n,i}|>A_n})\\
\leqslant& 2\,\frac{M_n}{\sigma_n^2}\sum_{i=1}^{N_n}\mathbb{E}(X_{n,i}^2\,\mathbf{1}_{|X_{n,i}|>A_n})\,.
\end{aligned}$$

Therefore, (4.40) gives that

$$\frac{S_n-\mathbb{E}(S_n)}{\sigma_n}-\frac{S_n^T-\mathbb{E}(S_n^T)}{\sigma_n}\xrightarrow{L^2}0\,. \tag{4.41}$$

To conclude, it remains to show that $\frac{\sigma_n^T}{\sigma_n}\to 1$ as n tends to ∞. Indeed, it is thus equivalent to check condition (4.39) for σ_n or σ_n^T so that we are able to apply Janson's theorem to obtain

$$\frac{S_n^T-\mathbb{E}(S_n^T)}{\sigma_n^T}\xrightarrow{d}\mathcal{N}(0,1)\,. \tag{4.42}$$

Moreover it implies that the distributional convergence of $\dfrac{S_n^T-\mathbb{E}(S_n^T)}{\sigma_n^T}$ is equivalent to the one of $\dfrac{S_n^T-\mathbb{E}(S_n^T)}{\sigma_n}$. To show that σ_n and σ_n^T are equivalent, notice that (4.41) and the reverse Minkowski inequality (in L^2) give

$$\left\|\frac{S_n-\mathbb{E}(S_n)}{\sigma_n}\right\|_{L^2}-\left\|\frac{S_n^T-\mathbb{E}(S_n^T)}{\sigma_n}\right\|_{L^2}\to 0\,,$$

which is exactly

$$1 - \frac{\sigma_n^T}{\sigma_n} \to 0 \,. \tag{4.43}$$

Finally, putting together (4.41), (4.42), (4.43), we obtain that

$$\frac{S_n - \mathbb{E}(S_n)}{\sigma_n} \xrightarrow{d} 0 \,. \qquad \square$$

Let us now get into the details of the application to the TV of a white Gaussian noise. For $\mathbf{x} \in \Omega_n$, we will set

$$Z_{n,\mathbf{x}} = |\dot{W}_n(x+1,y) - \dot{W}_n(x,y)| + |\dot{W}_n(x,y+1) - \dot{W}_n(x,y)| \,,$$

so that $\mathrm{TV}(W_n) = \sum_{\mathbf{x}\in\Omega_n} Z_{n,\mathbf{x}}$. With these notations, we will be able to apply Janson's theorem on this sum with $M_n = 6$. Indeed, for a fixed $\mathbf{x} = (x,y) \in \Omega_n$, the variables $\dot{W}_n(x+1,y)$, $\dot{W}_n(x,y+1)$ and $\dot{W}_n(x,y)$ appear in $Z_{n,\mathbf{x}}$. These two variables also appear in $Z_{n,(x-1,y)}$, $Z_{n,(x-1,y+1)}$, $Z_{n,(x,y-1)}$, $Z_{n,(x+1,y-1)}$, $Z_{n,(x+1,y)}$, $Z_{n,(x,y+1)}$, and do not appear in any other $Z_{n,\mathbf{x}}$, $\mathbf{x} \in \Omega_n$. That is why we can set $M_n = 6$.

Next, to apply the theorem, we also need to know the variance of the sum. It is actually independent of n and given by Theorem 4.2.1:

$$\sigma^2 = \sigma_n^2 = \mathrm{Var}(\mathrm{TV}(W_n)) = \frac{8}{\pi}(\omega(1) + 6 \cdot \omega(1/2)) \,.$$

Notice that the theorem also gives

$$\mathbb{E}(\mathrm{TV}(W_n)) = \frac{4}{\sqrt{\pi}} |\Omega_n|^{1/2} \,.$$

Now, it remains to find a sequence A_n which satisfies both (4.39) and (4.40). Since in our case M_n and σ_n are constant, we must find A_n and an m such that

$$|\Omega_n|^{1/m} A_n \to 0 \quad \text{and} \quad \sum_{\mathbf{x}\in\Omega_n} \mathbb{E}(Z_{n,\mathbf{x}}^2 \mathbf{1}_{|Z_{n,\mathbf{x}}|>A_n}) \to 0$$

as $n \to \infty$. Since all the $Z_{n,\mathbf{x}}$ follow the Gaussian distribution with standard deviation $2|\Omega_n|^{-1/2}$, the second condition is equivalent to

$$\mathbb{E}\left(Z^2 \mathbf{1}_{|Z|>A_n|\Omega_n|}\right) \to 0 \,.$$

Hence, it suffices to find A_n and an m such that

$$|\Omega_n|^{1/m} A_n \to 0 \quad \text{and} \quad A_n|\Omega_n| \to \infty \,.$$

We can take $m = 3$ and $A_n = |\Omega_n|^{-1/2}$. The two conditions are satisfied, and with Janson's theorem we obtain the result of Proposition 4.B.1.

Remark: One can point out that we applied a powerful central limit theorem in order to prove a very specific case. In fact, one can adapt the preceding proof to show that, as soon as u has compact support in Ω_n with $|\Omega_n| \to \infty$, we have normal convergence of $\mathbb{E}(\mathrm{TV}(u * W))$ after centralization and normalization.

4.C Proof of Theorem 4.2.1

Before proving Theorem 4.2.1, let us give two lemmas about Gaussian random vectors.

Lemme 4.C.1. *Let X be a Gaussian r.v. with zero mean and variance σ^2. Then*

$$\mathbb{E}(|X|) = \sigma\sqrt{\frac{2}{\pi}}\,.$$

Proof. Since $X \sim \mathcal{N}(0,\sigma^2)$, one can write

$$\begin{aligned} E(|X|) &= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}|x|e^{-\frac{x^2}{2\sigma^2}}dx = \frac{2}{\sigma\sqrt{2\pi}}\int_0^{+\infty}xe^{-\frac{x^2}{2\sigma^2}}dx \\ &= \frac{2}{\sigma\sqrt{2\pi}}\left[-\sigma^2e^{-\frac{x^2}{2\sigma^2}}\right]_0^{+\infty} = \sigma\sqrt{\frac{2}{\pi}}\,. \end{aligned}$$

□

Lemme 4.C.2. *Let $Z = (X,Y)^T$ be a Gaussian random vector with zero mean and covariance matrix*

$$\mathbb{E}(ZZ^T) = \begin{pmatrix} a^2 & ab\sin\theta \\ ab\sin\theta & b^2 \end{pmatrix}\,,$$

with $\theta \in [-\frac{\pi}{2},\frac{\pi}{2}]$. Then, one has

$$\mathbb{E}(|XY|) = \frac{2|ab|}{\pi}(\cos\theta + \theta\sin\theta)\,.$$

Proof. If $a = 0$ or $b = 0$, then $\mathbb{E}(XY) = 0$ so there is nothing more to prove. Hence we can assume that $ab \neq 0$ and set $X' = X/a$, $Y' = Y/b$, so that

$$\mathbb{E}|XY| = |ab|\cdot\mathbb{E}|X'Y'|\,, \tag{4.44}$$

where the covariance of $Z' = (X',Y')^T$ is

$$C = \mathbb{E}(Z'Z'^T) = \begin{pmatrix} 1 & \sin\theta \\ \sin\theta & 1 \end{pmatrix}\,.$$

If $|\sin\theta| = 1$, then $Y' = X'\sin\theta$ almost surely, so that $\mathbb{E}|X'Y'| = \mathbb{E}X'^2 = 1$ and $\mathbb{E}|XY| = |ab|$ by (4.44). Hence, we assume in the following that $|\theta| < \frac{\pi}{2}$. Now we have

$$C^{-1} = \frac{1}{\cos^2\theta}\begin{pmatrix} 1 & -\sin\theta \\ -\sin\theta & 1 \end{pmatrix}\,,$$

so that $\mathbb{E}|X'Y'|$ equals

$$\frac{1}{2\pi\cos\theta}\int_{\mathbb{R}^2}|xy|\exp\left(-\frac{x^2+y^2-2xy\sin\theta}{2\cos^2\theta}\right)dxdy\,.$$

Using symmetry considerations, this formula can be rewritten under the form

$$\mathbb{E}|X'Y'| = \frac{I(\theta)+I(-\theta)}{\pi\cos\theta} \tag{4.45}$$

$$\text{with } I(\theta) = \int_0^{+\infty}\int_0^{+\infty} xy \exp\left(-\frac{x^2+y^2-2xy\sin\theta}{2\cos^2\theta}\right) dxdy.$$

Using polar coordinates, we then get

$$\begin{aligned} I(\theta) &= \int_0^{+\infty}\int_0^{\frac{\pi}{2}} r^2\cos\varphi\sin\varphi \\ &\qquad \exp\left(-\frac{r^2}{2\cos^2\theta}(1-2\cos\varphi\sin\varphi\sin\theta)\right) r\, drd\varphi \\ &= \int_0^{\frac{\pi}{2}} \left(\cos\varphi\sin\varphi\int_0^{+\infty} r^3 e^{-\alpha(\varphi)r^2} dr\right) d\varphi\ , \end{aligned}$$

$$\text{with } \alpha(\varphi) = \frac{1-2\cos\varphi\sin\varphi\sin\theta}{2\cos^2\theta} \geqslant 0\ .$$

Integrating by part the inside integral yields

$$\begin{aligned} &\int_0^{+\infty} r^3 e^{-\alpha(\varphi)r^2} dr \\ &= \left[r^2\cdot\frac{1}{-2\alpha(\varphi)} e^{-\alpha(\varphi)r^2}\right]_0^{+\infty} - \frac{1}{-2\alpha(\varphi)}\int_0^{+\infty} 2re^{-\alpha(\varphi)r^2} dr \\ &= \frac{1}{2\alpha(\varphi)^2}\ . \end{aligned}$$

Thus we have

$$\begin{aligned} I(\theta) &= \int_0^{\frac{\pi}{2}} \cos\varphi\sin\varphi\cdot\frac{(2\cos^2\theta)^2}{2(1-2\cos\varphi\sin\varphi\sin\theta)^2} d\varphi \\ &= 2\cos^4\theta\cdot\int_0^{\frac{\pi}{2}} \frac{\tan\varphi}{(\cos^{-2}\varphi - 2\tan\varphi\sin\theta)^2}\,\frac{d\varphi}{\cos^2\varphi} \\ &= 2\cos^4\theta\cdot\int_0^{+\infty} \frac{t}{(1+t^2-2t\sin\theta)^2} dt \quad (t=\tan\varphi) \\ &= 2\cos^4\theta\cdot\int_0^{+\infty} \frac{t}{((t-\sin\theta)^2+\cos^2\theta)^2} dt \\ &= 2\cos^4\theta\cdot\int_{-\sin\theta}^{+\infty} \frac{u+\sin\theta}{(u^2+\cos^2\theta)^2} du \quad (u=t-\sin\theta)\ . \end{aligned}$$

Now usual integration formulae give (for $a>0$),

$$\int \frac{u}{(u^2+a^2)^2} du = \frac{-1}{2(u^2+a^2)}$$

$$\text{and} \quad \int \frac{1}{(u^2+a^2)^2} du = \frac{1}{2a^3}\arctan\frac{u}{a} + \frac{u}{2a^2(u^2+a^2)}\ ,$$

so that $I(\theta)$ equals

$$\begin{aligned} I(\theta) &= 2\cos^4\theta \Bigg(\left[\frac{-1}{2(u^2+\cos^2\theta)^2}\right]_{-\sin\theta}^{+\infty} \\ &+ \sin\theta \left[\frac{1}{2\cos^3\theta}\arctan\frac{u}{\cos\theta} + \frac{u}{2\cos^2\theta(u^2+\cos^2\theta)}\right]_{-\sin\theta}^{+\infty} \Bigg) \\ &= 2\cos^4\theta \left(\frac{1}{2} + \sin\theta\left(\frac{\pi}{2\cos^3\theta} + \frac{\theta}{2\cos^3\theta} + \frac{\sin\theta}{2\cos^2\theta}\right)\right) \\ &= \cos^4\theta + \pi\sin\theta\cos\theta + \theta\sin\theta\cos\theta + \sin^2\theta\cos^2\theta \\ &= \cos^2\theta + \pi\sin\theta\cos\theta + \theta\sin\theta\cos\theta \,. \end{aligned}$$

Then, $I(\theta) + I(-\theta) = 2\cos\theta(\cos\theta + \theta\sin\theta)$ and we conclude by (4.44) and (4.45) that

$$\mathbb{E}|XY| = \frac{2|ab|}{\pi}(\cos\theta + \theta\sin\theta) \,.$$

□

Proof of Theorem 4.2.1

Writing $U = u * W$, we have by linearity

$$\partial_x \dot{U} = (\partial_x \dot{u}) * W \,,$$

so that the discrete random field $\partial_x \dot{U}$ is a stationary Gaussian field whose marginal distributions have zero mean and variance

$$\mathbb{E}((\partial_x \dot{U}(\mathbf{x}))^2) = \frac{1}{MN}\sum_{\mathbf{y}\in\Omega}(\partial_x \dot{u}(\mathbf{x}-\mathbf{y}))^2 = \frac{\alpha_x^2}{MN} \,.$$

From Lemma 4.C.1, we hence get that for any $\mathbf{x} \in \Omega$,

$$\mathbb{E}(|\partial_x \dot{U}(\mathbf{x})|) = \frac{\alpha_x}{\sqrt{MN}}\sqrt{\frac{2}{\pi}} \,,$$

and by using a similar reasoning on $\partial_y \dot{U}$, we obtain (4.10).

We now consider the variance of $\mathrm{TV}(U)$. We have

$$\begin{aligned} \mathbb{E}(\mathrm{TV}(U)^2) = \sum_{\mathbf{x},\mathbf{y}\in\Omega} &\mathbb{E}|\partial_x \dot{U}(\mathbf{x})\partial_x \dot{U}(\mathbf{y})| + \mathbb{E}|\partial_x \dot{U}(\mathbf{x})\partial_y \dot{U}(\mathbf{y})| \\ &+ \mathbb{E}|\partial_y \dot{U}(\mathbf{x})\partial_x \dot{U}(\mathbf{y})| + \mathbb{E}|\partial_y \dot{U}(\mathbf{x})\partial_y \dot{U}(\mathbf{y})| \,. \end{aligned}$$

Writing $\mathbf{z} = \mathbf{y} - \mathbf{x}$ and using the stationarity of $\nabla \dot{U}$, the quantity $\mathbb{E}(\mathrm{TV}(U)^2)$ can be rewritten

$$\begin{aligned} MN \sum_{\mathbf{x}\in\Omega,\mathbf{y}\in\Omega} &\mathbb{E}|\partial_x \dot{U}(0)\partial_x \dot{U}(\mathbf{z})| + \mathbb{E}|\partial_x \dot{U}(0)\partial_y \dot{U}(\mathbf{z})| \\ &+ \mathbb{E}|\partial_y \dot{U}(0)\partial_x \dot{U}(\mathbf{z})| + \mathbb{E}|\partial_y \dot{U}(0)\partial_y \dot{U}(\mathbf{z})| \,. \end{aligned} \tag{4.46}$$

Each term of this sum can be written under the form $\mathbb{E}|XY|$ where (X, Y) is a zero-mean 2-dimensional Gaussian vector with covariance matrix

$$\begin{pmatrix} \mathbb{E}(X^2) & \mathbb{E}(XY) \\ \mathbb{E}(XY) & \mathbb{E}(Y^2) \end{pmatrix} .$$

For the second term of (4.46) for example, we have $X = \partial_x \dot{U}(0)$ and $Y = \partial_y \dot{U}(\mathbf{z})$, thus

$$\begin{aligned} \mathbb{E}(XY) &= \mathbb{E}\left(\sum_{\mathbf{x}\in\Omega, \mathbf{y}\in\Omega} \partial_x \dot{u}(-\mathbf{x}) \partial_y \dot{u}(\mathbf{z}-\mathbf{y}) W(\mathbf{x}) W(\mathbf{y}) \right) \\ &= \frac{1}{MN} \sum_{\mathbf{x}\in\Omega} \partial_x \dot{u}(\mathbf{x}) \partial_y \dot{u}(\mathbf{z}+\mathbf{x}) = \frac{1}{MN} \Gamma_{xy}(\mathbf{z}) \end{aligned}$$

and the covariance matrix of (X, Y) is

$$\frac{1}{MN} \begin{pmatrix} \alpha_x^2 & \Gamma_{xy}(\mathbf{z}) \\ \Gamma_{xy}(\mathbf{z}) & \alpha_y^2 \end{pmatrix} ,$$

so that thanks to Lemma 4.C.2 we obtain

$$\mathbb{E}|XY| = \frac{2\alpha_x \alpha_y}{\pi MN} \cdot \widetilde{\omega}\left(\frac{\Gamma_{xy}(\mathbf{z})}{\alpha_x \alpha_y} \right) ,$$

with $\widetilde{\omega}(t) = t \arcsin t + \sqrt{1-t^2} = \omega(t) + 1$. Combining all terms arising from (4.46), we finally obtain that

$$\begin{aligned} \mathbb{E}(\mathrm{TV}(U)^2) = \frac{2}{\pi} \sum_{\mathbf{z}\in\Omega} & \alpha_x^2 \widetilde{\omega}\left(\frac{\Gamma_{xx}(\mathbf{z})}{\alpha_x^2} \right) \\ & + 2\alpha_x \alpha_y \widetilde{\omega}\left(\frac{\Gamma_{xy}(\mathbf{z})}{\alpha_x \alpha_y} \right) + \alpha_y^2 \widetilde{\omega}\left(\frac{\Gamma_{yy}(\mathbf{z})}{\alpha_y^2} \right) \end{aligned} \tag{4.47}$$

and the announced result follows from

$$\mathrm{Var}(\mathrm{TV}(U)) = \mathbb{E}(\mathrm{TV}(U)^2) - (\mathbb{E}(\mathrm{TV}(U)))^2 ,$$

which simply amounts to change $\widetilde{\omega}$ into ω in (4.47).

4.D Unimodal regression

In this appendix, we detail an algorithm to compute the distance from a signal $s = (s(1), s(2), \ldots, s(n)) \in \mathbb{R}^n$ to the set U of unimodal signals of size n, defined by

$$U = \bigcup_{1\leqslant i\leqslant n} C_i \cap D_i ,$$

$$\text{where } C_i = \{p \in \mathbb{R}^n,\ p(1) \leqslant p(2) \leqslant \ldots \leqslant p(i)\}$$

$$\text{and } D_i = \{p \in \mathbb{R}^n,\ p(i) \geqslant p(i+1) \geqslant \ldots \geqslant p(n)\}$$

(with the natural convention $C_1 = D_n = \mathbb{R}^n$). The algorithm we use is due to Frisen [Frisen 1986]. It is based on the fact that U can also be written

$$U = \bigcup_{1 \leqslant i \leqslant n-1} C_i \cap D_{i+1} \,,$$

which entails $d(s,U) = \min_{1 \leqslant i \leqslant n-1} d_i$ with

$$\begin{aligned} d_i^2 &= \min_{p \in C_i \cap D_{i+1}} \|p - s\|_2^2 \\ &= \min_{p \in C_i} \sum_{k=1}^{i} (p(k) - s(k))^2 + \min_{q \in D_{i+1}} \sum_{k=i+1}^{n} (q(k) - s(k))^2 \,. \end{aligned}$$

These two monotone regression problems are independent, and can be solved in time $\mathcal{O}(n)$ using the simple *Pool Adjacent Violators* algorithm described in [Ayer *et al.* 1955] (see Algorithm 4). Thus, the computation of $d(s,U)$ can be realized in time $\mathcal{O}(n^2)$ (Algorithm 5). Note that in fact the unimodal regression problem can be solved in time $\mathcal{O}(n)$ with a more sophisticated algorithm (see [Stout 2008]), but considering the small value of n we use in Subsection 4.5.3 ($n = 20$), the gain obtained with this algorithm would be negligible compared to other steps (e.g., Fourier transforms) of the deblurring process.

4.E Oracle deconvolution filter

Consider a blurry and noisy image $v = \kappa * u_0 + n$, obtained from an image u_0 after a convolution by a kernel κ and the addition of a Gaussian white noise n with standard deviation σ^2. In this appendix, we show how to compute the oracle kernel k_0 which provides, in average with respect to n, the best linear estimate of u_0 that can be computed from v. This oracle kernel is defined by

$$k_0 = \arg\min_k \; \mathbb{E}\left(\|u_0 - k * (\kappa * u_0 + W)\|_2^2 \right) \,, \tag{4.48}$$

where W is a Gaussian white noise with variance σ^2. The arg min can be taken over various kernel spaces, here we consider the set of kernels obtained by rotating a radial linearly interpolated profile, that is

$$\forall \boldsymbol{\xi} \in \Omega, \quad \hat{k}(\boldsymbol{\xi}) = r(\lfloor |\boldsymbol{\xi}| \rfloor)(\lceil |\boldsymbol{\xi}| \rceil - |\boldsymbol{\xi}|) + r(\lceil |\boldsymbol{\xi}| \rceil)(|\boldsymbol{\xi}| - \lfloor |\boldsymbol{\xi}| \rfloor) \,,$$

where $(r(0), \ldots, r(d-1)) \in \mathbb{R}^d$,

$$|\boldsymbol{\xi}|^2 = 2(d-1)^2 \left(\left(\frac{\xi_1}{M} \right)^2 + \left(\frac{\xi_2}{N} \right)^2 \right) \,,$$

and $\lfloor t \rfloor$ and $\lceil t \rceil$ denote respectively the lower and upper integer part of $t \in \mathbb{R}$ (we also set $\hat{k}(\boldsymbol{\xi}) = 0$ when $|\boldsymbol{\xi}| > d-1$). This interpolation formula naturally involves the disjoint subsets

$$\widehat{\Omega}_l = \{ \boldsymbol{\xi} \in \Omega, l \leqslant |\boldsymbol{\xi}| < l+1 \} \,. \tag{4.49}$$

Algorithm 4: Monotone regression [Ayer *et al.* 1955]

- Inputs: $s \in \mathbb{R}^n$, $\varepsilon \in \{-1, 1\}$
- Output : non-decreasing (case $\varepsilon = 1$) or non-increasing (case $\varepsilon = -1$) regression p of s.
- $k \leftarrow 1$
- For each $i = 1, \ldots, n$
  - ▷ $\sigma_k \leftarrow s(i)$
  - ▷ $n_k \leftarrow 1$
  - ▷ While $k > 1$ and $\left(\frac{\sigma_{k-1}}{n_{k-1}} - \frac{\sigma_k}{n_k}\right)\varepsilon > 0$
    - · $\sigma_{k-1} \leftarrow \sigma_{k-1} + \sigma_k$
    - · $n_{k-1} \leftarrow n_{k-1} + n_k$
    - · $k \leftarrow k - 1$
  - ▷ $k \leftarrow k + 1$
- $i \leftarrow 1$
- For $l = 1, \ldots, k$, repeat n_l times the steps
  - ▷ $p(i) \leftarrow \sigma_l$
  - ▷ $i \leftarrow i + 1$

Algorithm 5: Unimodal regression distance [Frisen 1986]

- Input: $s \in \mathbb{R}^n$
- Output: $d(s, U)$
- For each $i = 1, \ldots, n$
  - ▷ $p \leftarrow$ non-decreasing regression of $(s(k))_{1 \leqslant k \leqslant i}$
  - ▷ $q \leftarrow$ non-increasing regression of $(s(k))_{i+1 \leqslant k \leqslant n}$
  - ▷ $d_i^2 \leftarrow \sum_{k=1}^{i} (s(k) - p(k))^2 + \sum_{k=1}^{n-k-1} (s(i+1+k) - q(k))^2$
- return $\min_i d_i$.

Since W is a white Gaussian noise, the cost function of (4.48) can be written

$$\begin{aligned}&\|u_0 - k * \kappa * u_0\|_2^2 + \sigma^2 MN \|k\|_2^2 \\ =&\frac{1}{MN}\sum_{\boldsymbol{\xi}\in\Omega} |\widehat{u_0}(\boldsymbol{\xi})|^2 |1 - \hat{k}(\boldsymbol{\xi})\hat{\kappa}(\boldsymbol{\xi})|^2 + \sigma^2 MN |\hat{k}(\boldsymbol{\xi})|^2 \,, \end{aligned} \tag{4.50}$$

which, when $\hat{k}$ is radial and when κ is supposed to be symmetrical, transforms into

$$\begin{aligned}\frac{1}{MN}\sum_{l=0}^{d-1}\sum_{\boldsymbol{\xi}\in\widehat{\Omega}_l} |\widehat{u_0}(\boldsymbol{\xi})|^2 \bigg(1 - \kappa(\boldsymbol{\xi}) r(l)(l+1-|\boldsymbol{\xi}|) - \kappa(\boldsymbol{\xi}) r(l+1)(|\boldsymbol{\xi}| - l)\bigg)^2 \\ + \sigma^2 MN \bigg(r(l)(l+1-|\boldsymbol{\xi}|) + r(l+1)(|\boldsymbol{\xi}| - l)\bigg)^2 \,.\end{aligned}$$

This is a quadratic function in r, and its unique minimum is characterized by the vanishing-gradient condition, which can be written $Ar = b$, where $A = ((a_{k,l}))_{0\leqslant k,l\leqslant d-1}$ and $b = (b_l)_{0\leqslant l\leqslant d-1}$ are defined by

$$\begin{aligned}a_{l,l} &= \sum_{\boldsymbol{\xi}\in\widehat{\Omega}_l} (l+1-|\boldsymbol{\xi}|)^2 (|\kappa(\boldsymbol{\xi})|^2 |\widehat{u_0}(\boldsymbol{\xi})|^2 + \sigma^2 MN) + \sum_{\boldsymbol{\xi}\in\widehat{\Omega}_{l-1}} (|\boldsymbol{\xi}| - l + 1)^2 (|\kappa(\boldsymbol{\xi})|^2 |\widehat{u_0}(\boldsymbol{\xi})|^2 + \sigma^2 MN) \\ a_{l,l+1} &= \sum_{\boldsymbol{\xi}\in\widehat{\Omega}_l} (l+1-|\boldsymbol{\xi}|)(|\boldsymbol{\xi}| - l)(|\kappa(\boldsymbol{\xi})|^2 |\widehat{u_0}(\boldsymbol{\xi})|^2 + \sigma^2 MN) \\ a_{l,l-1} &= \sum_{\boldsymbol{\xi}\in\widehat{\Omega}_{l-1}} (|\boldsymbol{\xi}| - l + 1)(l - |\boldsymbol{\xi}|)(|\kappa(\boldsymbol{\xi})|^2 |\widehat{u_0}(\boldsymbol{\xi})|^2 + \sigma^2 MN) \\ a_{l,m} &= 0 \quad \text{for} \quad |l-m| > 1 \\ b_l &= \sum_{\boldsymbol{\xi}\in\widehat{\Omega}_l} (t+1-|\boldsymbol{\xi}|)^2 (|\kappa(\boldsymbol{\xi})||\widehat{u_0}(\boldsymbol{\xi})|^2) + \sum_{\boldsymbol{\xi}\in\widehat{\Omega}_{l-1}} (|\boldsymbol{\xi}| - l + 1)^2 (|\kappa(\boldsymbol{\xi})||\widehat{u_0}(\boldsymbol{\xi})|^2) \,.\end{aligned}$$

This linear system associated to the tridiagonal matrix A can be solved with standard numerical techniques. The solution is the oracle radial profile r_0, from which the DFT of the oracle kernel k_0 can be defined by

$$\forall l, \forall \boldsymbol{\xi} \in \widehat{\Omega}_l, \quad \widehat{k_0}(\boldsymbol{\xi}) = r_0(l)(l+1-|\boldsymbol{\xi}|) + r_0(l+1)(|\boldsymbol{\xi}| - l) \,.$$

Remark: One can also consider the minimization problem (4.48) on the set of all kernels k. It is easy to deduce from (4.50) that the corresponding oracle kernel is given in Fourier domain by

$$\forall \boldsymbol{\xi} \in \Omega, \quad \hat{k}(\boldsymbol{\xi}) = \frac{\hat{\kappa}(\boldsymbol{\xi})^* \, |\widehat{u_0}(\boldsymbol{\xi})|^2}{|\kappa(\boldsymbol{\xi})|^2 |\widehat{u_0}(\boldsymbol{\xi})|^2 + \sigma^2 MN} \,.$$

One can notice that, making the assumption $|\hat{u}(\boldsymbol{\xi})|^2 = c\left(4\pi^2\left(\frac{\xi_1^2}{M^2} + \frac{\xi_2^2}{N^2}\right)\right)^{-1}$ (see the discussion at the end of Appendix A), and setting $\lambda = \sigma^2 MN/c$, the corresponding filter is exactly the one that optimizes the criterion (4.26).

Chapter 5

Random Fields with Structured Phase

Contents

The objective of this chapter is to discuss the possibility of direct phase analysis or synthesis. After illustrating that direct phase synthesis is not an appropriate way to tackle macrotexture synthesis, we propose two extensions of the random phase texture models of Chapter 2 which can deal with more structured textures.

One important goal of this thesis was to design new models of random fields which can be easily simulated and which are richer than the uniform random phase models RPN and ADSN (which are unable to reproduce sharp edges). Of course, several texture synthesis algorithms are able to deal with structured textures, for example [Efros & Leung 1999], [Efros & Freeman 2001], [Kwatra *et al.* 2005], or [Lefebvre & Hoppe 2005]. Even if these algorithms are very efficient and allow to reproduce a large variety of textures, one common drawback is that the output distribution of these algorithms is difficult to describe. More precisely, these algorithms are motivated by a Markov model, but the properties of the output random field are hardly examined: for example, the output of [Efros & Leung 1999] is clearly not stationary, and even if it is sampled by a scheme that respects the Markov assumption, the local specification cannot be made explicit (due to the approximation of the Gibbs sampling). Notice however that the authors of the later

article [Levina & Bickel 2006] proved the consistency of this algorithm in a resampling framework (based on some assumptions on the input texture and provided that the size of the observation grows to infinity).

It is actually a difficult (and still open) problem to design a texture synthesis algorithm that truly respects a random model, and still allows for macrotexture synthesis. In this chapter, we will tackle this problem by pursuing the works of [Van Wijk 1991], [Kwatra *et al.* 2005] and [Galerne *et al.* 2011b]. Let us recall that the RPN model is based on a very transparent analysis/synthesis pipeline: the Fourier modulus is extracted, and the phase component is replaced by a uniform random phase. In other words, we sample according to the maximum entropy distribution for a fixed Fourier modulus. Therefore, one can hope to improve the RPN model by preserving more than the Fourier modulus. But the questions are then: which additional relevant features can be extracted? and is it possible to sample a texture which preserves these new features?

A naive approach is to search for these features directly in the phase information. But, because of its link to spatial translations, the direct analysis of the phase information is difficult. In particular, in the circular framework, we will show in Subsection 5.1.1 that no relevant phase constraint can be drawn only from the stationarity assumption. A way to represent the phase up to the spatial translations is to consider the phase of the bispectrum, as mentioned in Subsection 5.1.2, but we will see that the bispectrum data are too large (which makes it very difficult to handle in practice), and that the preservation of the whole bispectrum does not leave any room for innovation in the textural content. Finally, we also show in Subsection 5.1.3 that the phase coherence indices introduced in Chapter 4 do not suffice to measure precisely the plausibility of a texture image in the uniform random phase model.

Therefore, in the lack of a simple richer extension of the RPN algorithm, we explored variants of the ADSN model. The advantage of this approach is that it can rely on several results given for Poisson spot noises in [Galerne 2010], and that it allows for easy experiments. In Section 5.2 we derive a local spot noise model which allows to resynthesize a given exemplar texture by preserving its local aspect. This model is not stationary, but is able to reproduce non random phase features, thus demonstrating the benefit of relaxing the stationarity constraint in the Gaussian model.

In Section 5.3, we also propose a general methodology for macrotexture synthesis, by defining bi-level models which consist in a stationary low-resolution component on which textural details are added. According to the observation that any texture, seen from sufficiently far away, can be considered as a microtexture, the low-resolution component can be synthesized with an ADSN field. Then, based on this coarse synthesis, fine-scale details may be added using a kind of texture refinement [Lefebvre & Hoppe 2005], [Chainais *et al.* 2011]. Here, this refinement step is emulated with a local patch-based operator inspired by [Kwatra *et al.* 2005]. The resulting algorithm (which corresponds only to a single instance of the bi-level model) is able to synthesize macrotextures and still has clear mathematical guarantees.

5.1 Phase Sensitive Texture Analysis

The main objective of this section is to illustrate the difficulty of incorporating the phase information in by-example texture synthesis. Our original idea was to seek simple extensions of the RPN model that can preserve something more than the modulus information. A first step, that is exposed in Subsection 5.1.1 was to understand more precisely the phase constraints that are imposed by the stationarity assumption. We will show in particular that these phase constraints are actually very weak and cannot inspire us to address direct synthesis of the phase information. Seeking other phase analysis tools, we turned to the bispectrum phase, which is exposed in Subsection 5.1.2. We will see that it is indeed more relevant in the phase analysis of texture images (because it is invariant to spatial translations), but also that it is much more redundant than the DFT phase so that its practical use is made even more intricate. In Subsection 5.1.3, we illustrate that the phase coherence indices of Chapter 4 do not suffice to measure the plausibility of a texture image in the RPN model.

In this section,

$$\Theta = \mathbb{Z}/M\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$$

will denote a periodic rectangular discrete domain of size $M \times N$.

5.1.1 Phase Information and Stationarity

In this subsection, we justify the difficulty of direct phase analysis, and clarify the impact of stationarity in the phase constraints of circular stationary random fields.

We recall again that the phase of an image $u : \Theta \longrightarrow \mathbb{R}$ is the argument φ of its discrete Fourier transform $\hat{u} = |\hat{u}|e^{i\varphi}$. It is thus an angular function $\varphi : \Theta \longrightarrow \mathbb{T}$ (where $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$) which satisfies $\varphi(-\boldsymbol{\xi}) = -\varphi(\boldsymbol{\xi})$. If $\hat{u}(\boldsymbol{\xi}) \neq 0$, the corresponding phase coefficient $\varphi(\boldsymbol{\xi})$ is uniquely defined modulo 2π whereas any value can be chosen if $\hat{u}(\boldsymbol{\xi}) = 0$.

A first naive difficulty is that the visualization of the phase information suffers from the fact that the angular values are wrapped onto $\mathbb{R}/2\pi\mathbb{Z}$. Indeed, in practice, when visualizing the phase information of a natural image, one usually computes the phase values $\varphi(\boldsymbol{\xi}) \in [-\pi, \pi[$. Actually, algorithms exist for unwrapping the phase information [Ghiglia & Pritt 1998] but they become inefficient when dealing with too irregular phase functions.

But the main difficulty of the phase information is that it is indeed very irregular because of its link with spatial translations: a translation of the image in spatial domain corresponds in Fourier domain to the addition of a linear function (sometimes called *ramp function*) $\boldsymbol{\xi} \mapsto \langle \mathbf{v}, \boldsymbol{\xi} \rangle$ to the phase function (where $\mathbf{v}$ is the translation vector). If the vector $\mathbf{v}$ is large, the addition of the ramp function completely changes the behavior of the phase function, even in the case of a very simple image (for example an elementary shape like a disc or an elongated blob). If one considers images obtained as addition of several elementary shapes, the situation is even worse because the computation of the phase information does not respect the addition.

This link between phase and translations has some consequences on the structure of the phase component of circular stationary random fields. For example, the following result was proved in [Matsubara 2007]: if the random field $F : \Theta \longrightarrow \mathbb{R}$ with phase ψ has a distribution that is invariant to the subpixel translations of Θ (defined by (4.22)), and if $\boldsymbol{\xi}$, $\boldsymbol{\zeta}$ are two linearly independent frequencies such that $\hat{F}(\boldsymbol{\xi})$ and $\hat{F}(\boldsymbol{\zeta})$ are almost surely non zero, then the corresponding phase coefficients $\psi(\boldsymbol{\xi})$, $\psi(\boldsymbol{\zeta})$ are necessarily independent and uniform on $\mathbb{T}$.

With the underlying desire to synthesize plausible texture phase functions, it may seem interesting to precise this result by looking for necessary and sufficient phase constraints that ensure the stationarity of the random field. Actually, the following theorem brings a disappointing answer to that question.

Theorem 5.1.1. *Let $F : \Theta \longrightarrow \mathbb{R}^d$ be a random field. Then F is circular stationary if and only if F has the same distribution as $F(\mathbf{V} + \cdot)$ where $\mathbf{V}$ follows the uniform distribution on Θ and is independent of F.*

Proof. Let us assume that F and $F(\mathbf{V} + \cdot)$ have the same distribution, where $\mathbf{V}$ is independent of F and follows the uniform distribution on Θ. The random field $F(\mathbf{V} + \cdot)$ is circular stationary because for each $\mathbf{v} \in \Theta$, $(\mathbf{V}, F)$ has the same distribution as $(\mathbf{V} + \mathbf{v}, F)$ and thus $F(\mathbf{V} + \cdot)$ has the same distribution as $F(\mathbf{V} + \mathbf{v} + \cdot)$. Therefore, F is also circular stationary.

Conversely, assume that F is circular stationary and that $\mathbf{V}$ is a random vector which is independent of F and has uniform distribution on Θ. Since F and $\mathbf{V}$ are independent, the distribution of $(F, \mathbf{V})$ is the tensor product between the distribution of F and the distribution $\frac{1}{|\Theta|}\sum_{\mathbf{v}\in\Theta} \delta_{\mathbf{v}}$ of $\mathbf{V}$. Thus Fubini's theorem gives that for each measurable function $h : \mathbb{R}^\Theta \longrightarrow \mathbb{R}_+$,

$$\mathbb{E}\left[h(F(\mathbf{V} + \cdot))\right] = \frac{1}{|\Theta|}\sum_{\mathbf{v}\in\Theta} \mathbb{E}\left[h(F(\mathbf{v} + \cdot))\right] = \frac{1}{|\Theta|}\sum_{\mathbf{v}\in\Theta} \mathbb{E}\left[h(F)\right] = \mathbb{E}[h(F)] \ ,$$

so that $F(\mathbf{V} + \cdot)$ has the same distribution as F. □

The last theorem states that any random field can be made circular stationary by applying a random translation of this domain, and also that, in distribution, any circular stationary random field can be obtained in this way. In other words, the phase of a circular stationary random field F can always be understood as the addition of itself and an independent ramp function whose gradient is chosen uniformly in Θ. This shows that the phase constraints induced by the stationarity assumption are somehow independent with the phase constraints that are needed to produce salient features in the spatial domain.

5.1.2 Bispectrum

One possibility to extend texture analysis to higher-order statistics is to consider higher-order spectra, and in particular the bispectrum. We refer to [Picinbono 1998], [Nikias & Mendel 1993], [Hall & Giannakis 1995] or [Collis *et al.* 1998] for a detailed presentation of the higher-order spectra. Here,

we only present basic properties of the bispectrum, and explain why it has not proven to be useful in designing non random phase texture models.

Definition 5.1.1. Given a gray-level image $u : \Theta \longrightarrow \mathbb{R}$, one can define the bicorrelation $C^2 : \Theta \times \Theta \longrightarrow \mathbb{R}$ of u by

$$C^2(\mathbf{v}, \mathbf{w}) = \sum_{\mathbf{x} \in \Theta} u(\mathbf{x}) u(\mathbf{x} + \mathbf{v}) u(\mathbf{x} + \mathbf{w}) \ . \tag{5.1}$$

Then, the bispectrum $B : \Theta \times \Theta \longrightarrow \mathbb{C}$ of u is defined as the discrete Fourier transform of C^2 with respect to both variables $\mathbf{v}, \mathbf{w}$, that is

$$B(\boldsymbol{\xi}, \boldsymbol{\zeta}) = \sum_{\mathbf{v}, \mathbf{w} \in \Theta} C^2(\mathbf{v}, \mathbf{w}) e^{-i\langle \boldsymbol{\xi}, \mathbf{v} \rangle - i \langle \boldsymbol{\zeta}, \mathbf{w} \rangle} \ . \tag{5.2}$$

A simple calculation gives the useful formula

$$B(\boldsymbol{\xi}, \boldsymbol{\zeta}) = \hat{u}(\boldsymbol{\xi}) \hat{u}(\boldsymbol{\zeta}) \overline{\hat{u}(\boldsymbol{\xi} + \boldsymbol{\zeta})} \ , \tag{5.3}$$

which establishes a link between the bispectrum and the usual DFT of u.

One can first draw several remarks on this definition. First, we see that the bispectrum indeed encodes a third-order information and in particular, in contrast with the Fourier modulus (or autocorrelation), it is sensitive to the symmetry $u \mapsto -u$. Furthermore, the bispectrum of a real image satisfies several constraints, for example

$$\begin{aligned} B(\boldsymbol{\xi}, \boldsymbol{\zeta}) &= B(\boldsymbol{\zeta}, \boldsymbol{\xi}) \ , \\ B(-\boldsymbol{\xi} - \boldsymbol{\zeta}, \boldsymbol{\zeta}) &= B(\boldsymbol{\xi}, \boldsymbol{\zeta}) \ , \\ B(\boldsymbol{\xi}, -\boldsymbol{\xi} - \boldsymbol{\zeta}) &= B(\boldsymbol{\xi}, \boldsymbol{\zeta}) \ , \\ B(-\boldsymbol{\xi}, -\boldsymbol{\zeta}) &= \overline{B(\boldsymbol{\xi}, \boldsymbol{\zeta})} \ . \end{aligned}$$

Subsequently, the bispectrum is a very redundant information (much more than the DFT which is only constrained by $\hat{u}(-\boldsymbol{\xi}) = \overline{\hat{u}(\boldsymbol{\xi})}$ for real images u). Notice that it is possible to derive subdomains of $\Theta \times \Theta$ on which the bispectrum is not redundant [Chandran & Elgar 1994], but these so-called principal domains still occupies a considerable volume of $\Theta \times \Theta$, making the bispectrum not easy to store or interpret.

Notice also that the modulus of the bispectrum $|\hat{B}|$ depends only on $|\hat{u}|$ so that the real benefit of the bispectrum lies in its phase. More precisely, if φ is a phase function for u, a phase function for the bispectrum is obtained as

$$\arg(B(\boldsymbol{\xi}, \boldsymbol{\zeta})) = \varphi(\boldsymbol{\xi}) + \varphi(\boldsymbol{\zeta}) - \varphi(\boldsymbol{\xi} + \boldsymbol{\zeta}) \ . \tag{5.4}$$

One can observe that if one adds a linear function to φ, the bispectrum phase is unchanged, which is equivalent to assert that the bispectrum is not affected by a translation of the spatial domain. Therefore, the bispectrum phase (5.4) allows to represent the phase in a way that is insensitive to the addition of the ramp function, as suggested in Subsection 5.1.1, at the cost of one additional frequency variable. The following theorem (which is a discrete version of [Jaming & Kolountzakis 2003, Lemma 2.3]) actually shows that, up to the translation invariance, there is no loss of information between the phase and the phase of the bispectrum.

Theorem 5.1.2. *Let $u_1, u_2 : \Theta \longrightarrow \mathbb{R}$ be two images having the same mean and same Fourier modulus and such that for each frequency $\boldsymbol{\xi}$, $|\widehat{u}_i(\boldsymbol{\xi})| \neq 0$. We assume that u_1 and u_2 have the same bispectrum. Then u_1 and u_2 only differ by a subpixel translation (defined in Fourier domain by (4.22)).*

Proof. Since u_1, u_2 have the same Fourier modulus, by definition (4.22) of the subpixel translations, it is enough to show that the phase functions φ_1, φ_2 of u_1, u_2 differ from a linear function on the domain

$$\Omega = \mathbb{Z}^2 \cap \left(\left[-\frac{M}{2}, \frac{M}{2} \right) \times \left[-\frac{N}{2}, \frac{N}{2} \right) \right) \ .$$

From the equality of the bispectra, setting $\theta = \varphi_1 - \varphi_2$, and from (5.4) we get that

$$\forall \boldsymbol{\xi}, \boldsymbol{\zeta} \in \Theta, \quad \theta(\boldsymbol{\xi} + \boldsymbol{\zeta}) = \theta(\boldsymbol{\xi}) + \theta(\boldsymbol{\zeta}) \qquad \mod 2\pi \ .$$

From this, we get that for each couple of integers $\boldsymbol{\xi} \in \Omega$,

$$\theta(\boldsymbol{\xi}) = \xi_1 \theta(1,0) + \xi_2 \theta(0,1) = 2\pi \left(\frac{v_1 \xi_1}{M} + \frac{v_2 \xi_2}{N} \right) = \langle \mathbf{v}, \boldsymbol{\xi} \rangle \ ,$$

where we chose

$$v_1 = \frac{M}{2\pi} \theta(1,0) \quad , \quad v_2 = \frac{N}{2\pi} \theta(0,1) \ .$$

□

From the result of this theorem, we see that the bispectrum does not seem to be adapted to texture synthesis by example. Indeed, it shows that if a synthesis algorithm preserves all the bispectrum, then it can only apply a subpixel translation to the texture. In other words, preserving the whole bispectrum information does not leave any room for innovation in the textural content. However, one may still want to explore the possibilities of preserving only a part of the bispectrum. But extracting a relevant part of the bispectrum seems at least as complicated as extracting a relevant part of the DFT.

5.1.3 Phase Coherence of Textures

We have seen in Chapter 3 how we could perform texture synthesis using random phase models. In this paragraph, we address the following question: is it possible to measure *a priori* the performance of these synthesis algorithms on a particular exemplar? It amounts to measure the plausibility of the exemplar in a random phase model, or in other words, to give some justification (other than visual) of the statement "the texture I am observing is a random phase texture". About the expression "random phase texture", let us mention that one should make a clear distinction between random phase models (which are random images whose Fourier transform has uniform random phase) and random phase textures (which are natural texture images that appear plausible in a uniform random phase model).

One naive approach (which will be quickly proved wrong) is to measure the plausibility of a particular element in a fixed random model by computing the corresponding likelihood value. One may wonder if the plausibility of a texture in

the ADSN model can be measured by the likelihood value obtained after the maximization of the likelihood function with respect to the model parameters (explained in Subsection 3.1.1). Actually, one can see from the calculation given in Subsection 3.1.1 that for an observation u whose DFT does not vanish at any frequency, the maximum value of the likelihood function in the ADSN model is equal to

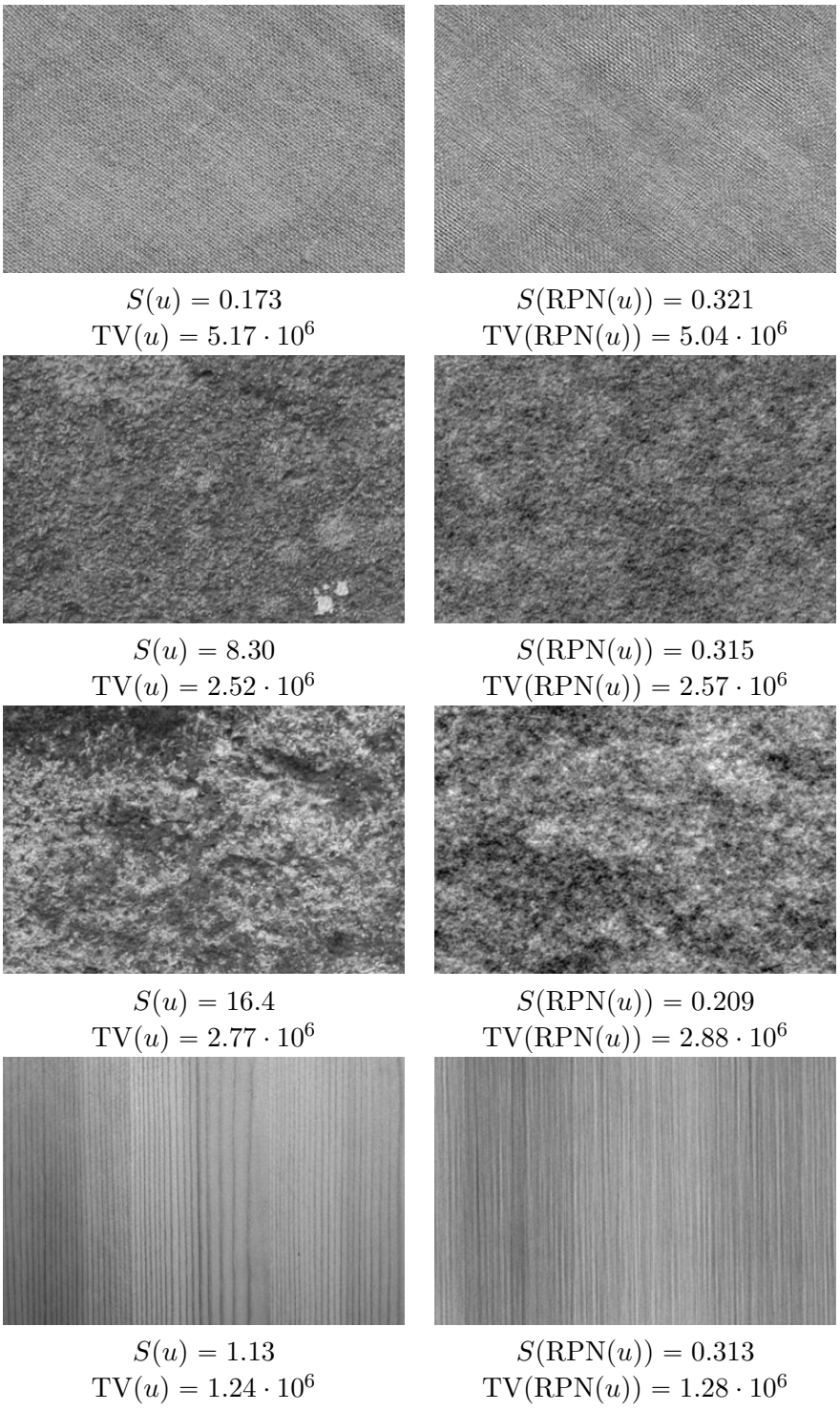
$$\prod_{\boldsymbol{\xi}\in\Theta\setminus\{0\}} \frac{1}{|\hat{u}(\boldsymbol{\xi})|}$$

up to some multiplicative constant that does not depend on u. It is clear that this value does not reflect the plausibility in the random phase model, should it be only for the reason of homogeneity (if u is multiplied by λ, this value is multiplied by $\lambda^{1-|\Theta|}$), and especially because it does not depend on the phase of u. In fact, the maximum likelihood procedure, which was designed for estimation purpose, does not allow to measure the adequation of the fitted model to the observation. This is confirmed by another simple remark: one would not say that 0 is a plausible sample of $\mathcal{N}(0,1)$ even if it realizes the maximum likelihood of the corresponding model.

Actually, the intuitive notion of plausibility cannot be assessed only through the random model but must rely on geometric measures that are somehow linked to our texture perception. If one relies on the total variation to measure the texture regularity, then one can question the plausibility of a texture in the random phase model by using the phase coherence indices introduced in Chapter 4. Indeed, we analyzed in Subsection 4.3.3 the values obtained with the phase coherence indices GPC, SI and S computed over random phase fields. Precisely, we observe numerically that the distribution of the S value of a random phase field was concentrated around the value 0.3. Therefore, an image that is plausible in the random phase model is expected to have a low value of S. This is confirmed by the examples of Fig 5.1 in which we gave some examples of RPN results, each time with the corresponding value for S and TV. Notice that in contrast, the TV value does not indicate anything about the plausibility of the texture in the random phase model.

But is the converse true?: if an image has a small value of S, can it be considered to be a random phase texture? Unfortunately, this is not so simple, as illustrated in Fig. 5.2. In this figure, we propose the same experiment as in Fig. 5.1, but with non random phase textures. Of course, textures with large scale geometry and sharp edges (like in the first and fourth row of Fig. 5.2) will have a large value of S. But not all the non random phase textures have such cartoon elements.

For example, the second row of Fig 5.2 exhibits a fabric texture on which the random phase synthesis fails (because of the complicated mixtures of directions and because this texture has a quasi-periodic salient pattern), and yet, this image has a small value of S. This small value of S is explained by the severe oscillations of this texture which are reflected by the high value of TV. So there exist oscillating non random phase textures having a small value of S. One could wonder why we cannot apply this argument on the vegetation texture of the third row of Fig 5.2. In this case, one can observe that the phase randomization makes the TV increase more than in the second row, so that the TV of the vegetation image is still low amongst the TV of its phase randomizations. We can give the following explanation: in the



$S(u) = 0.173$
$\mathrm{TV}(u) = 5.17 \cdot 10^6$

$S(\mathrm{RPN}(u)) = 0.321$
$\mathrm{TV}(\mathrm{RPN}(u)) = 5.04 \cdot 10^6$

$S(u) = 8.30$
$\mathrm{TV}(u) = 2.52 \cdot 10^6$

$S(\mathrm{RPN}(u)) = 0.315$
$\mathrm{TV}(\mathrm{RPN}(u)) = 2.57 \cdot 10^6$

$S(u) = 16.4$
$\mathrm{TV}(u) = 2.77 \cdot 10^6$

$S(\mathrm{RPN}(u)) = 0.209$
$\mathrm{TV}(\mathrm{RPN}(u)) = 2.88 \cdot 10^6$

$S(u) = 1.13$
$\mathrm{TV}(u) = 1.24 \cdot 10^6$

$S(\mathrm{RPN}(u)) = 0.313$
$\mathrm{TV}(\mathrm{RPN}(u)) = 1.28 \cdot 10^6$

Figure 5.1: **Some examples of random phase textures.** For each row, the original texture is shown on the left, together with a realization of the associated random phase noise. One can see that the corresponding values of phase coherence are quite low. Let us add that in the second example, the lighter stains on the bottom-right is unlikely to happen in a RPN realization; but except this detail, we can say that it is a random phase texture.

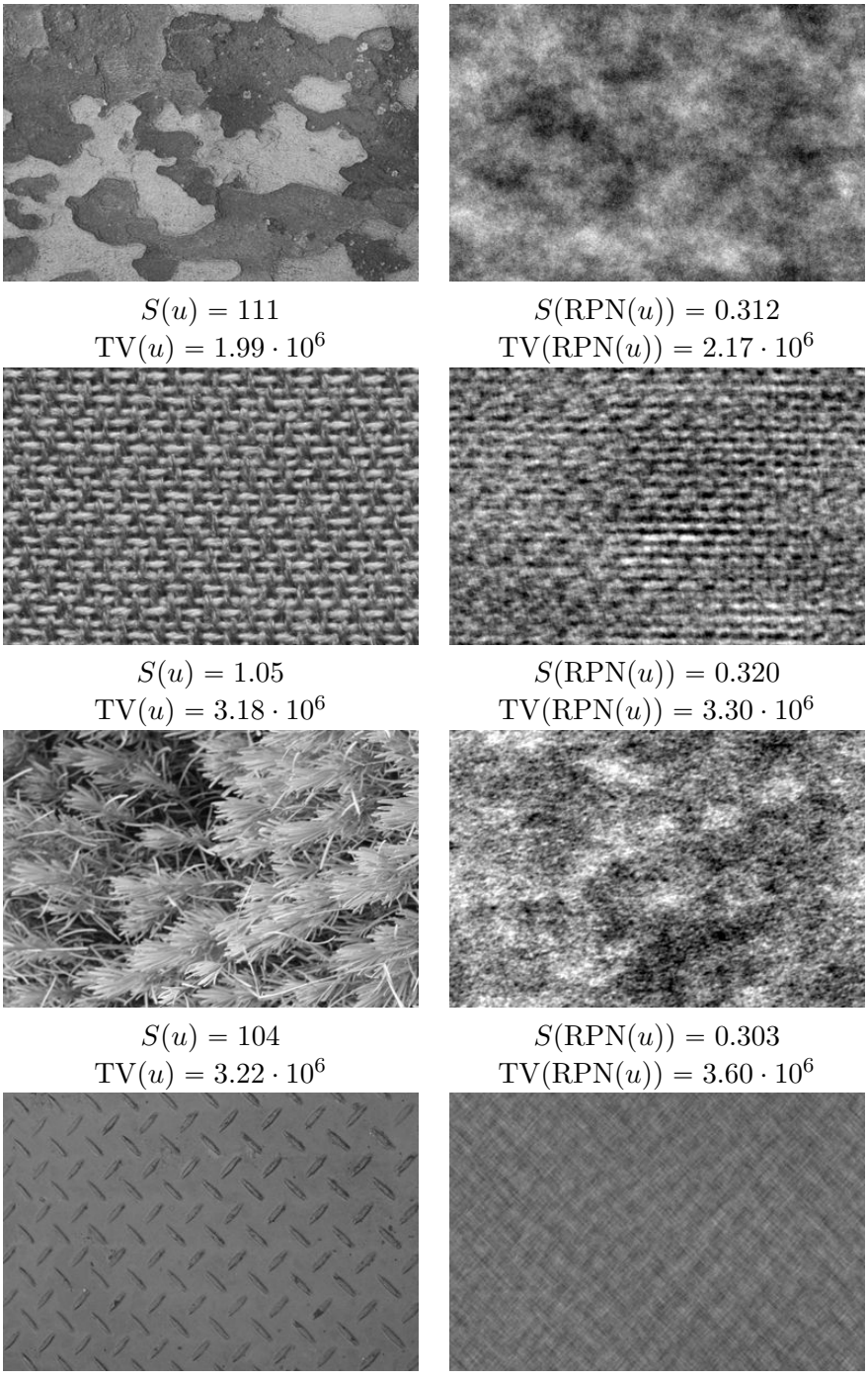


Figure 5.2: **Some examples of non random phase textures.** For each row, the original texture is shown on the left, together with a realization of the associated random phase noise. One can see that the corresponding values of S are high except for the fabric case (second row).

vegetation image, the TV is high because the image exhibits in several regions a lot of small oscillations; but also, the gradient energy (in ℓ^2-norm) is higher than in the fabric image, because large gradients are observed at the boundary of the dark regions at the top-middle and bottom-left of the image. All in all, one must keep in mind that the index S is related to the TV, but also to the gradient energy.

To sum up, on these examples we see that these sharpness indices affect small values to random phase textures. But they are not sufficient to discriminate random phase textures from the others, because a non random phase texture can have a TV that is already large (amongst its phase randomizations) and thus a small S value. Hence, these phase coherence indices can be used as a non random phase test for textures, but not as a precise plausibility measure in the random phase model.

5.2 Local Spot Noise Synthesis

In the last section, we illustrated the difficulty to design simple extensions of the RPN model that deals directly with the phase information. In this section, we will design and study an extension of the ADSN model, benefiting on the fact that ADSN offers larger experimental possibilities (because of the convolutive expression of spot noise models). Since we have seen in Subsection 2.1.4 that no phase coherence can be expected in stationary Gaussian random fields, our extension will consist of a non-stationary Gaussian random field. It will demonstrate that non random phase features can be better reproduced by relaxing the stationarity constraint in the Gaussian model.

More precisely, we propose to define a local spot noise model based on the observation of a texture $u : \Omega \longrightarrow \mathbb{R}^d$. This model allows to resynthesize the texture u on the same domain and respects local second-order properties of the texture. The synthesized image is a non-stationary Gaussian random field $U : \Omega \to \mathbb{R}^d$ built as a limit when $\lambda \to +\infty$ of a localized discrete spot noise $U_\lambda : \Omega \to \mathbb{R}^d$. In this spot noise, the patches thrown at the position $\mathbf{X}$ of the Poisson process are chosen in a spatial neighborhood of $\mathbf{X}$. This local spot noise synthesis can be understood as resynthesis of the texture conditionally to a low-frequency component.

Notice that this local ADSN lies close to the "patch Gaussian model" of [Raad *et al.* 2014] but here, the patches are chosen in a spatial neighborhood of $\mathbf{X}$ and not according to a patch similarity criterion. The local ADSN can also be compared to the non-stationary Gaussian model of [Boussidi *et al.* 2014]; apart from the fact that we work in a discrete framework (for rather practical reasons), the main difference is that the spot noise kernels are directly extracted from the exemplar, and do not belong to a parametric family.

In this section, $\Omega \subset \mathbb{Z}^2$ is a discrete domain of size $M \times N$ and $u : \Omega \longrightarrow \mathbb{R}^d$ is an exemplar texture image. Let us recall that if $A, B \subset \mathbb{Z}^2$, we denote by $|A|$ the cardinal of A, and we set

$$A + B = \{\mathbf{x} + \mathbf{y} \ , \ \mathbf{x} \in A, \mathbf{y} \in B\} \ ,$$

$$A - B = \{\mathbf{x} - \mathbf{y} \ , \ \mathbf{x} \in A, \mathbf{y} \in B\} \ .$$

5.2.1 Local Spot Noise Model

In this subsection, we define and study the local spot noise which is obtained by summing renormalized patches of the exemplar image found around each point of the synthesis.

Let $\omega \subset \mathbb{Z}^2$ be the finite domain on which patches are defined. Let $D \subset \mathbb{Z}^2$ a neighborhood of 0, which represent the research zone of the patches. This domain ω and D can have a general form, but in the following experiments we consider ω to be a disc with center 0 and radius r, and D to be a disc with center 0 and radius ρ.

First, let us consider the image $\bar{u} : \Omega \to \mathbb{R}^d$ that represents the local mean value of u, defined by

$$\bar{u}(\mathbf{y}) = \frac{1}{|(\mathbf{y}+\omega) \cap \Omega|} \sum_{\mathbf{z} \in (\mathbf{y}+\omega)\cap\Omega} u(\mathbf{z}) \; . \tag{5.5}$$

In other words, $\bar{u}(\mathbf{y})$ is the average color value in the vicinity of pixel $\mathbf{y}$. Next, for each $\mathbf{y} \in \mathbb{Z}^2$, let us extract from u the normalized patch $p_u(\cdot, \mathbf{y})$ centered on $\mathbf{y}$ which is defined by

$$p_u(\mathbf{h}, \mathbf{y}) = \begin{cases} \dfrac{1}{|(\mathbf{y}+\omega) \cap \Omega|^{\frac{1}{2}}} \Big(u(\mathbf{y}+\mathbf{h}) - \bar{u}(\mathbf{y}) \Big) & \text{if } \mathbf{h} \in \omega \text{ and } \mathbf{y}+\mathbf{h} \in \Omega \\ 0 \quad \text{otherwise} \end{cases} \; . \tag{5.6}$$

Let us introduce a marked Poisson process $(\mathbf{X}_i, \mathbf{T}_i)$, $(i \geqslant 1)$ with intensity $\lambda|\cdot| \times \mathcal{U}(D)$ where $|\cdot|$ refers to the counting measure on $\mathbb{Z}^2$ and $\mathcal{U}(D)$ to the uniform distribution on D. In other words, $(\mathbf{X}_i)$ is a uniform Poisson process on $\mathbb{Z}^2$ with intensity λ, and with independent marks $\mathbf{T}_i$ which are uniformly distributed on D. We can now introduce the Poisson sum defined for all $\mathbf{x} \in \Omega$ by

$$S_\lambda(\mathbf{x}) = \sum_{i \geqslant 1} p_u(\mathbf{x} - \mathbf{X}_i, \mathbf{X}_i + \mathbf{T}_i) \; . \tag{5.7}$$

It means that around each point $\mathbf{X}_i$ we add a patch $p_u(\,\cdot\,, \mathbf{X}_i + \mathbf{T}_i)$ taken at the position $\mathbf{X}_i + \mathbf{T}_i$. Notice that the Poisson Process $(\mathbf{X}_i)$ can be restricted to

$$\overline{\Omega} = (\Omega - \omega) \cap (\Omega - D) \; .$$

Definition 5.2.1. Given the exemplar image $u : \Omega \longrightarrow \mathbb{R}^d$, the patch domain ω, the research domain D, and the intensity λ, the local discrete spot noise (LDSN) is defined for all $\mathbf{x} \in \Omega$ by

$$U_\lambda(\mathbf{x}) = \bar{u}(\mathbf{x}) + \frac{1}{\sqrt{\lambda}} \Big(S_\lambda(\mathbf{x}) - m(\mathbf{x}) \Big) \; , \tag{5.8}$$

where $m(\mathbf{x}) = \mathbb{E}(S_\lambda(\mathbf{x}))$.

We will show later that when λ tends to $+\infty$, the random field U_λ converges towards a Gaussian random field $U : \Omega \to \mathbb{R}^d$, that is called local asymptotic discrete spot noise.

First-order moments

Grouping the points of the marked Poisson process $(\mathbf{X}_i, \mathbf{T}_i)$ that fall on each point of $\overline{\Omega} \times D$, we can write

$$S_\lambda(\mathbf{x}) = \sum_{(\mathbf{z},\mathbf{t}) \in \overline{\Omega} \times D} N_\lambda(\mathbf{z}, \mathbf{t}) p_u(\mathbf{x} - \mathbf{z}, \mathbf{z} + \mathbf{t}) \tag{5.9}$$

where $(N_\lambda(\mathbf{z}, \mathbf{t}))_{\mathbf{z} \in \overline{\Omega} \times D}$ is a collection of independent random variables with distribution $\mathcal{P}(\frac{\lambda}{|D|})$. We can thus obtain the expectation m of S_λ, as

$$m(\mathbf{x}) = \mathbb{E}(S_\lambda(\mathbf{x})) = \frac{\lambda}{|D|} \sum_{\mathbf{z} \in \overline{\Omega}} \sum_{\mathbf{t} \in D} p_u(\mathbf{x} - \mathbf{z}, \mathbf{z} + \mathbf{t}) \ .$$

Using the notation

$$q(\mathbf{y}) = \begin{cases} \frac{1}{|(y+\omega) \cap \Omega|^{\frac{1}{2}}} & \text{if } \mathbf{y} \in \Omega - \omega \\ 0 \quad \text{otherwise} \end{cases} \ , \tag{5.10}$$

we can give a little more explicit formula:

$$\begin{aligned} m(\mathbf{x}) &= \frac{\lambda}{|D|} \sum_{\mathbf{z} \in \mathbf{x} - \omega} \sum_{\mathbf{t} \in D \cap (\Omega - \mathbf{x})} q(\mathbf{z} + \mathbf{t})(u(\mathbf{x} + \mathbf{t}) - \bar{u}(\mathbf{z} + \mathbf{t})) \\ &= \frac{\lambda}{|D|} \sum_{\mathbf{t} \in D \cap (\Omega - \mathbf{x})} (\mathbf{1}_\omega * q(\mathbf{x} + \mathbf{t}))\, u(\mathbf{x} + \mathbf{t}) - \mathbf{1}_\omega * (q\bar{u})(\mathbf{x} + \mathbf{t}) \ . \end{aligned}$$

In other words,

$$m = \frac{\lambda}{|D|} \mathbf{1}_D * \Big(\mathbf{1}_\Omega \Big((\mathbf{1}_\omega * q) u - \mathbf{1}_\omega * (q\bar{u}) \Big) \Big) \ , \tag{5.11}$$

where the convolutions are computed on Ω with null boundary conditions (which amounts to extend u, $\bar{u}$ and q by zero-padding to $\mathbb{Z}^2$). This shows in particular that the expectation m can be computed in $\mathcal{O}(|\Omega| \log |\Omega|)$ thanks to the DFT.

From Equation (5.9), we also get the spatial covariance of S_λ, which leads to the following result.

Proposition 5.2.1. *The LDSN process defined by (5.8) has expectation $\bar{u}$ and its covariance is given for all $\mathbf{x}, \mathbf{y} \in \Omega$ by*

$$\begin{aligned} \Gamma(\mathbf{x}, \mathbf{y}) = \mathrm{Cov}(U_\lambda(\mathbf{x}), U_\lambda(\mathbf{y})) &= \frac{1}{|D|} \sum_{(\mathbf{z},\mathbf{t}) \in \overline{\Omega} \times D} p_u(\mathbf{x} - \mathbf{z}, \mathbf{z} + \mathbf{t}) p_u(\mathbf{y} - \mathbf{z}, \mathbf{z} + \mathbf{t})^T \\ = \sum_{\mathbf{z} \in (\mathbf{x} - \omega) \cap (\mathbf{y} - \omega)} \frac{1}{|D|} \sum_{\mathbf{t} \in D \cap (\Omega - \mathbf{x}) \cap (\Omega - \mathbf{y})} & \frac{1}{|(\mathbf{z} + \mathbf{t} + \omega) \cap \Omega|} (u(\mathbf{x} + \mathbf{t}) - \bar{u}(\mathbf{z} + \mathbf{t}))(u(\mathbf{y} + \mathbf{t}) - \bar{u}(\mathbf{z} + \mathbf{t}))^T \ . \end{aligned}$$

This result can be rephrased by saying that the covariance is obtained as a local average of patches cross-correlations. It also justifies the adopted patch normalization (which was inspired by Subsection 3.1) at least through the marginal covariance. Indeed, let us assume that $\mathbf{x} = \mathbf{y}$ and that

$$(\mathbf{x} + 2\omega + D) \cup (\mathbf{y} + 2\omega + D) \subset \Omega \ .$$

Then we obtain the marginal covariance of $U_\lambda(\mathbf{x})$ as

$$\frac{1}{|D||\omega|}\sum_{\mathbf{h}\in\omega}\sum_{\mathbf{t}\in D}(u(\mathbf{x}+\mathbf{t})-\bar{u}(\mathbf{x}-\mathbf{h}+\mathbf{t}))(u(\mathbf{x}+\mathbf{t})-\bar{u}(\mathbf{x}-\mathbf{h}+\mathbf{t}))^T\ .$$

It is an unusual estimate of the color covariance of u on $\mathbf{x}+\omega$. But still, it indicates that the normalization that was adopted for U_λ is not absurd in terms of the marginal distributions.

Gaussian limit

Theorem 5.2.1. *When $\lambda\to+\infty$, U_λ converges in distribution to the non-stationary Gaussian random field $U\sim\mathcal{N}(\bar{u},\Gamma)$ whose first-order moments are*

$$\bar{u}(\mathbf{y})=\frac{1}{|(\mathbf{y}+\omega)\cap\Omega|}\sum_{\mathbf{z}\in(\mathbf{y}+\omega)\cap\Omega}u(\mathbf{z})\ ,$$
$$\Gamma(\mathbf{x},\mathbf{y})=\frac{1}{|D|}\sum_{(\mathbf{z},\mathbf{t})\in\overline{\Omega}\times D}p_u(\mathbf{x}-\mathbf{z},\mathbf{z}+\mathbf{t})p_u(\mathbf{y}-\mathbf{z},\mathbf{z}+\mathbf{t})^T\ .$$

This Gaussian random field U is called the ***local asymptotic discrete spot noise (LADSN) associated to u.***

Proof. One can immediately adapt the proof of Theorem 2.1.1 to the non-stationary case (because the Poisson point process can be restricted to a finite set). □

Simulation

Now we will see that the LADSN can be directly sampled (without requiring the heavy simulation of a high-intensity LDSN). For that, let us introduce a family $(W_\mathbf{t})_{\mathbf{t}\in D}$ of independent Gaussian white noise processes defined on $\overline{\Omega}$, with mean 0 and variance 1. For each $\mathbf{t}\in D$, let us consider

$$V_\mathbf{t}(\mathbf{x})=\sum_{\mathbf{z}\in\overline{\Omega}}W_\mathbf{t}(\mathbf{z})p_u(\mathbf{x}-\mathbf{z},\mathbf{z}+\mathbf{t})\ .$$

Notice that $V_\mathbf{t}$ is a Gaussian random field with mean 0 and covariance given by

$$\mathrm{Cov}(V_\mathbf{t}(\mathbf{x}),V_\mathbf{t}(\mathbf{y}))=\sum_{\mathbf{z}\in\overline{\Omega}}p_u(\mathbf{x}-\mathbf{z},\mathbf{z}+\mathbf{t})p_u(\mathbf{y}-\mathbf{z},\mathbf{z}+\mathbf{t})^T\ .$$

Besides, the random fields $\mathbf{x}\mapsto V_\mathbf{t}(\mathbf{x})$ are independent. Therefore,

$$V(\mathbf{x})=\bar{u}(\mathbf{x})+\frac{1}{\sqrt{|D|}}\sum_{\mathbf{t}\in D}V_\mathbf{t}(\mathbf{x}) \tag{5.12}$$

is a Gaussian random field with mean $\bar{u}$ and covariance function

$$\mathrm{Cov}(V_\mathbf{t}(\mathbf{x}),V_\mathbf{t}(\mathbf{y}))=\frac{1}{|D|}\sum_{(\mathbf{z},\mathbf{t})\in\overline{\Omega}\times D}p_u(\mathbf{x}-\mathbf{z},\mathbf{z}+\mathbf{t})p_u(\mathbf{y}-\mathbf{z},\mathbf{z}+\mathbf{t})^T=\Gamma(\mathbf{x},\mathbf{y})\ .$$

Since the distribution of a Gaussian random field is characterized by its second-order moments, it follows that V is the LADSN associated to u.

We will derive from that a simple algorithm for the simulation of the LADSN. Using the definition of the patches p_u, one has

$$\begin{aligned} V_{\mathbf{t}}(\mathbf{x}) &= \sum_{\mathbf{z}\in(\mathbf{x}-\omega)\cap(\Omega-\mathbf{t})} q(\mathbf{z}+\mathbf{t})W_{\mathbf{t}}(\mathbf{z})(u\mathbf{1}_\Omega(\mathbf{x}+\mathbf{t}) - \bar{u}\mathbf{1}_\Omega(\mathbf{z}+\mathbf{t})) \\ &= u\mathbf{1}_\Omega(\mathbf{x}+\mathbf{t})\Big(\mathbf{1}_\omega * (\tau_{-\mathbf{t}}(q\mathbf{1}_\Omega)\cdot W_{\mathbf{t}})(\mathbf{x})\Big) - \mathbf{1}_\omega * (\tau_{-\mathbf{t}}(q\bar{u}\mathbf{1}_\Omega)\cdot W_{\mathbf{t}})(\mathbf{x})\mathbf{1}_\Omega(\mathbf{x}+\mathbf{t}) \,, \end{aligned}$$

where we use the notation $\tau_{\mathbf{h}}v(\mathbf{x}) = v(\mathbf{x}-\mathbf{h})$ (and each time, we also precised $v\mathbf{1}_\Omega$ when we need to consider the zero-padding extension of v outside Ω) . This allows to compute $V_{\mathbf{t}}$ through two convolutions with null-boundary conditions. And after, one can obtain the LADSN using Equation (5.12). Notice that this last operation certainly is the most expansive one: it is not a convolution because the white noise processes W_t are not related.

Circular analog

Notice that one can also define a circular local discrete spot noise model. When referring to the circular model, we will use analog notations except that we add a dot on the letters. In particular, $\dot{u}$ will refer to the (M,N)-periodic extension of u. Besides, the notation $\odot$ will refer to the circular convolution over the domain Ω.

In the circular case, the local mean can be computed by

$$\bar{\dot{u}} = \frac{1}{|\omega|}\mathbf{1}_\omega \odot u \,.$$

Then, one can consider the normalized patches

$$\forall \mathbf{y}\in\Omega, \quad \dot{p}_u(\mathbf{h},\mathbf{y}) = \begin{cases} \dfrac{1}{\sqrt{|\omega|}}\Big(\dot{u}(\mathbf{y}+\mathbf{h}) - \bar{\dot{u}}(\mathbf{y})\Big) & \text{if } \mathbf{h}\in\omega \\ 0 \quad \text{otherwise} \end{cases} \tag{5.13}$$

extracted from the (M,N)-periodic extension $\dot{u}$ of u. And we can define the circular LDSN by

$$\dot{U}_\lambda(\mathbf{x}) = \bar{\dot{u}}(\mathbf{x}) + \frac{1}{\sqrt{\lambda}}\sum_{i\geqslant 1}(\dot{p}_u(\mathbf{x}-\mathbf{X}_i,\mathbf{X}_i+\mathbf{T}_i) - \dot{m}(\mathbf{x})) \,, \tag{5.14}$$

where $(\mathbf{X}_i)$ is a Poisson process on Ω with intensity λ and independent marks $(\mathbf{T}_i)$ which are uniform on D, and where the expectation of the Poisson sum is given by

$$\dot{m} = \frac{\lambda\sqrt{|\omega|}}{|D|}\mathbf{1}_D \odot \left(u - \frac{1}{|\omega|}\mathbf{1}_\omega \odot \bar{\dot{u}}\right) \,.$$

The covariance of $\dot{U}_\lambda$ is given for all $\mathbf{x},\mathbf{y}\in\Omega$ by

$$\begin{aligned} \dot{\Gamma}(\mathbf{x},\mathbf{y}) &= \frac{1}{|D|}\sum_{\mathbf{t}\in D}\sum_{\mathbf{z}\in\Omega}\dot{p}_u(\mathbf{x}-\mathbf{z},\mathbf{z}+\mathbf{t})\dot{p}_u(\mathbf{y}-\mathbf{z},\mathbf{z}+\mathbf{t}) \\ &= \frac{1}{|D|}\sum_{\mathbf{t}\in D}\frac{1}{|\omega|}\sum_{\mathbf{z}\in(\mathbf{x}-\omega)\cap(\mathbf{y}-\omega)}\Big(\dot{u}(\mathbf{x}+\mathbf{t}) - \bar{\dot{u}}(\mathbf{z}+\mathbf{t})\Big)\Big(\dot{u}(\mathbf{y}+\mathbf{t}) - \bar{\dot{u}}(\mathbf{z}+\mathbf{t})\Big) \,. \end{aligned}$$

Again we can show that when $\lambda \to +\infty$, $\dot{U}_\lambda$ converges to the Gaussian random field $\dot{U} \sim \mathcal{N}(\bar{u}, \dot{\Gamma})$ called the circular LADSN associated to u.

With the same computation as above, we show that the circular LADSN can be directly obtained by setting

$$\dot{V}(\mathbf{x}) = \frac{1}{\sqrt{|D|}} \sum_{\mathbf{t} \in D} V_{\mathbf{t}}(\mathbf{x}) \ ,$$

$$\dot{V}_{\mathbf{t}}(\mathbf{x}) = \sum_{\mathbf{z} \in \Omega} \dot{W}_{\mathbf{t}}(\mathbf{z}) \dot{p}_u(\mathbf{x}-\mathbf{z}, \mathbf{z}+\mathbf{t}) = \frac{1}{\sqrt{|\omega|}} \Big(\dot{u}(\mathbf{x}+\mathbf{t})(\mathbf{1}_\omega \odot \dot{W}_{\mathbf{t}}(\mathbf{x})) - \mathbf{1}_\omega \odot (\dot{W}_{\mathbf{t}} \dot{\tau}_{-\mathbf{t}} \bar{\dot{u}})(\mathbf{x}) \Big) \ , \tag{5.15}$$

where $\mathbf{x} \mapsto \dot{W}_{\mathbf{t}}(\mathbf{x})$, $\mathbf{t} \in D$ are independent white Gaussian noise processes on Ω of mean 0 and variance 1, and where $\dot{\tau}_{\mathbf{h}}$ refers to the periodic translation of vector $\mathbf{h}$.

5.2.2 Results and comments

In this subsection, we use the LADSN model to perform resynthesis of given exemplar textures. We also investigate the influence of the synthesis parameters, and discuss the invariance to the addition of a smooth component.

We only consider discrete disc domains

$$\omega = \{\mathbf{x} \in \mathbb{Z}^2, \quad x_1^2 + x_2^2 \leqslant (r + 0.5)^2\}$$

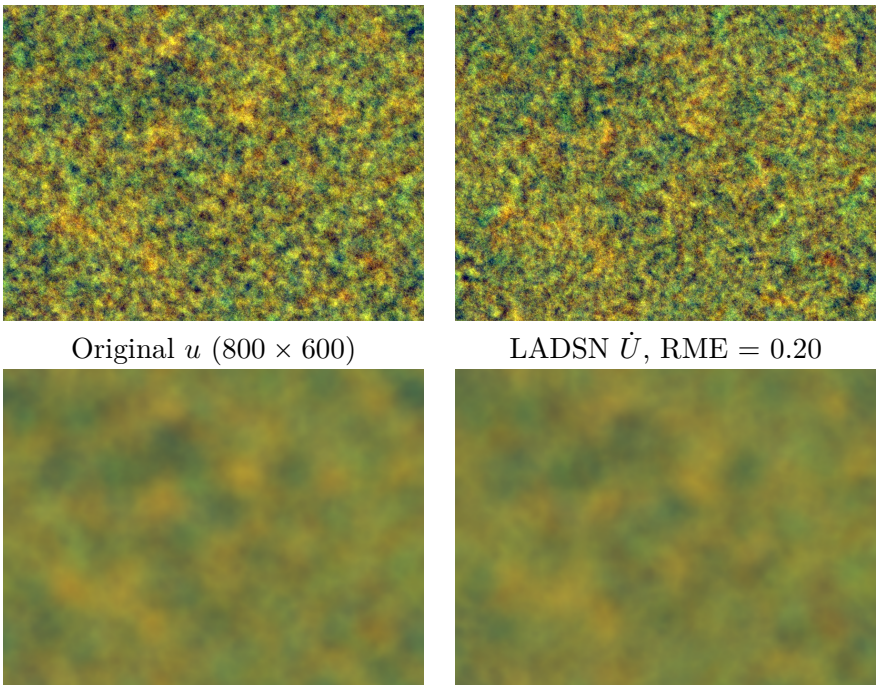
$$D = \{\mathbf{x} \in \mathbb{Z}^2, \quad x_1^2 + x_2^2 \leqslant (\rho + 0.5)^2\}$$

where r and ρ are two parameters (adding 0.5 to r and ρ is a way to obtain discrete discs with a more satisfying circular aspect).

Synthesis examples

Let us first consider in Fig. 5.3 and Fig. 5.4 the result of LADSN synthesis on stationary Gaussian textures. One can notice in Fig. 5.3 that this method allows to resynthesize a Gaussian texture in a faithful manner and also preserves a low-frequency component. This is confirmed by the low values of the relative model error (defined by (3.17)) obtained between the LADSN synthesis and the original Gaussian texture. This shows that, in terms of resynthesis, the local spot noise model is at least as rich as the stationary spot noise model that we studied in Chapter 2. However, as for the spot noise associated to the synthesis-oriented texton (see Section 3.3), the correlations length in the LADSN is constrained by twice the size of the patch domain ω because the LADSN covariance $\Gamma(\mathbf{x}, \mathbf{y})$ vanishes as soon as $\mathbf{x} - \mathbf{y} \notin 2\omega$. Thus, the patch domain must be tuned as the support size of the SOT.

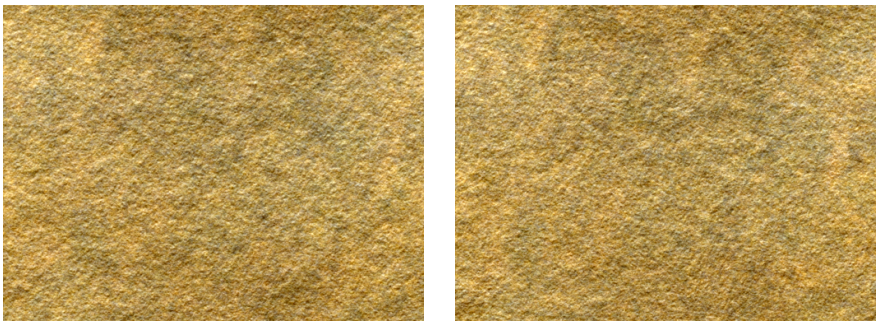
In Fig. 5.5, we show examples of local spot noise associated to four real non-stationary textures. Surprisingly enough, even if the model is only based on second-order statistics, it still allows to reproduce local features of the texture to a certain extent. Indeed, this model is able to preserve local orientations, and also certain fiber structures like the ones of the radiographic image in the third row of Fig. 5.5.



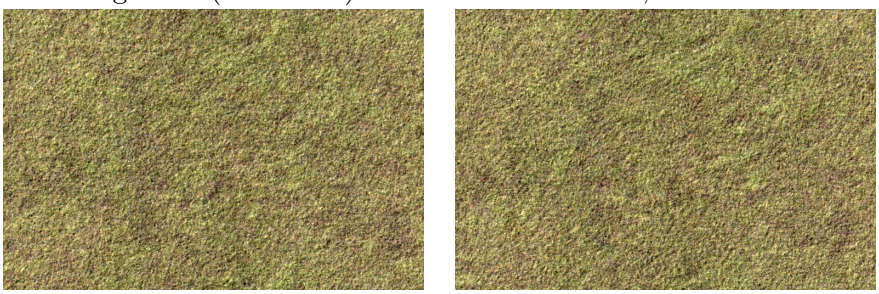
Original u (800×600) LADSN $\dot{U}$, RME = 0.20

Local mean $\bar{\dot{u}}$ Local mean $\bar{\dot{U}}$

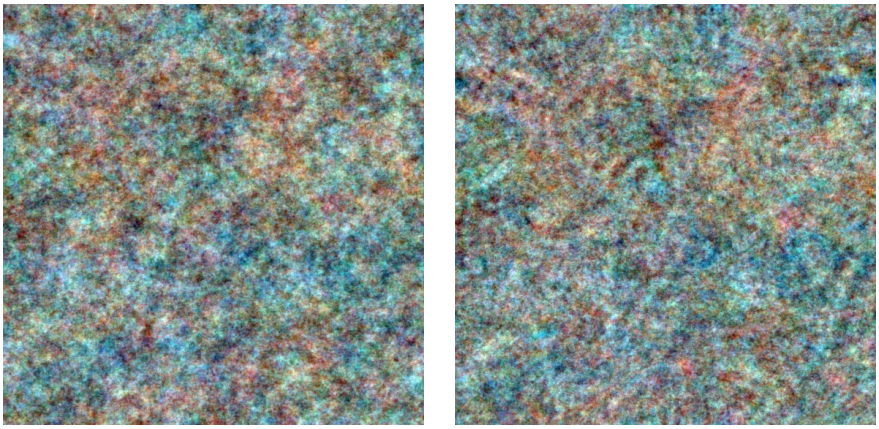
Figure 5.3: **Resynthesis of a homogeneous random phase texture using a local spot noise**. In this figure, we synthesized a circular Gaussian texture using a circular LADSN (with $r = \rho = 30$), and we also computed the local mean images corresponding to u and the realization $\dot{U}$. Notice that $\dot{U}$ and u have similar aspect which is confirmed by a small relative model error (RME) between the Gaussian models associated to u and $\dot{U}$. Besides, as one can see on the local mean images $\bar{\dot{u}}$ and $\bar{\dot{U}}$, the LADSN preserved in a certain sense the low frequency component of the texture.



Original u (512×384) LADSN $\dot{U}$, RME = 0.12



Original u (768×512) LADSN $\dot{U}$, RME = 0.18



Original u (576×576) LADSN $\dot{U}$, RME = 0.18

Figure 5.4: **Resynthesis of random phase textures using a local spot noise**. In this figure, one can observe three other examples of Gaussian texture resynthesis using a circular local spot noise model. This confirms that Gaussian textures can be resynthesized using the LADSN. Notice however that on the first and second line, one can observe slight interference patterns in the synthesis. These patterns disappear with a fine tuning of the parameters r and ρ (see Fig. 5.6 and Fig. 5.7).

However, strongly non-Gaussian features (like very sharp edges) cannot be resynthesized properly (due to the spatial mixing of patches). Besides, one can observe in this figure that the approximation of the LADSN by the LDSN is visually satisfactory when, for a patch domain ω of radius $r = 20$, the mean number of impacts per pixel exceeds 100. Therefore, even if the asymptotic simulation is more difficult in the non-stationary case, the finite-intensity spot noise can still be considered as a very efficient way to perform approximate simulation (as it was the case with the DSN associated to the SOT in Section 3.3).

Influence of the parameters

In Fig. 5.6, one can observe the result of ADSN synthesis for varying size ρ of the research zone D. One can see that when D increases, the local covariances are more and more mixed so that the local orientations are less and less preserved. On the other hand, when D is too small, the synthesis result is troubled with interference patterns. We have not yet come to a satisfying explanation for this artefact. Let us mention that is is also visible in certain Gabor noise textures (see auxiliary material of [Lagae *et al.* 2009]).

In Fig. 5.7, one can also see the influence of the size r of the patch domain ω. Again, one can see on the covariance expression that we have $\Gamma(\mathbf{x}, \mathbf{y}) = 0$ as soon as $\mathbf{x} - \mathbf{y} \notin 2\omega$. Since the LADSN is a Gaussian field, this entails that the samples $U(\mathbf{x})$ and $U(\mathbf{y})$ are independent. In other words, the long-range interactions are only reproduced if they are included in the local mean image $\bar{u}$. This is confirmed by visual inspection of the results shown in Fig. 5.7.

Concerning the parameter ω, we will see that one extreme case is quite easy to interpret. In the circular case, if $\omega = \Omega$, then in Equation (5.15) the local mean image $\bar{u}$ is constant and so are the noise images $\mathbf{1}_\omega \odot \dot{W}_\mathbf{t}$. Thus, we get in this case

$$\dot{V}(\mathbf{x}) = \frac{1}{\sqrt{|\Omega||D|}} \sum_{\mathbf{t} \in D} (\dot{u}(x + t) - \bar{u}) G_\mathbf{t}$$

where $(G_\mathbf{t})_{\mathbf{t} \in D}$ is a Gaussian white noise with mean 0 and variance 1. Therefore, we get that $\dot{V}$ is a circular convolution of the normalized spot $t_u = \frac{1}{\sqrt{|\Omega|}}(u - \bar{u})$ with a spatially localized white noise, which makes this particular case very easy to understand. In particular, if $\omega = D = \Omega$, then the circular LADSN matches the classical circular ADSN.

Invariance to the addition of a smooth component

Since the patches are normalized by the local mean image $\bar{u}$, the local spot noise model is robust to the addition of a very smooth component, which can be observed in Fig. 5.8. This is true provided that the patch domain ω is sufficiently large so that the smooth component will, in a certain sense be absorbed by the local mean image $\bar{u}$. It means that the model is able to capture local variations of the textures, independently of local illumination changes. If the smooth component is not regular enough, then the local autocorrelations are not preserved, as one can see in Fig. 5.9.

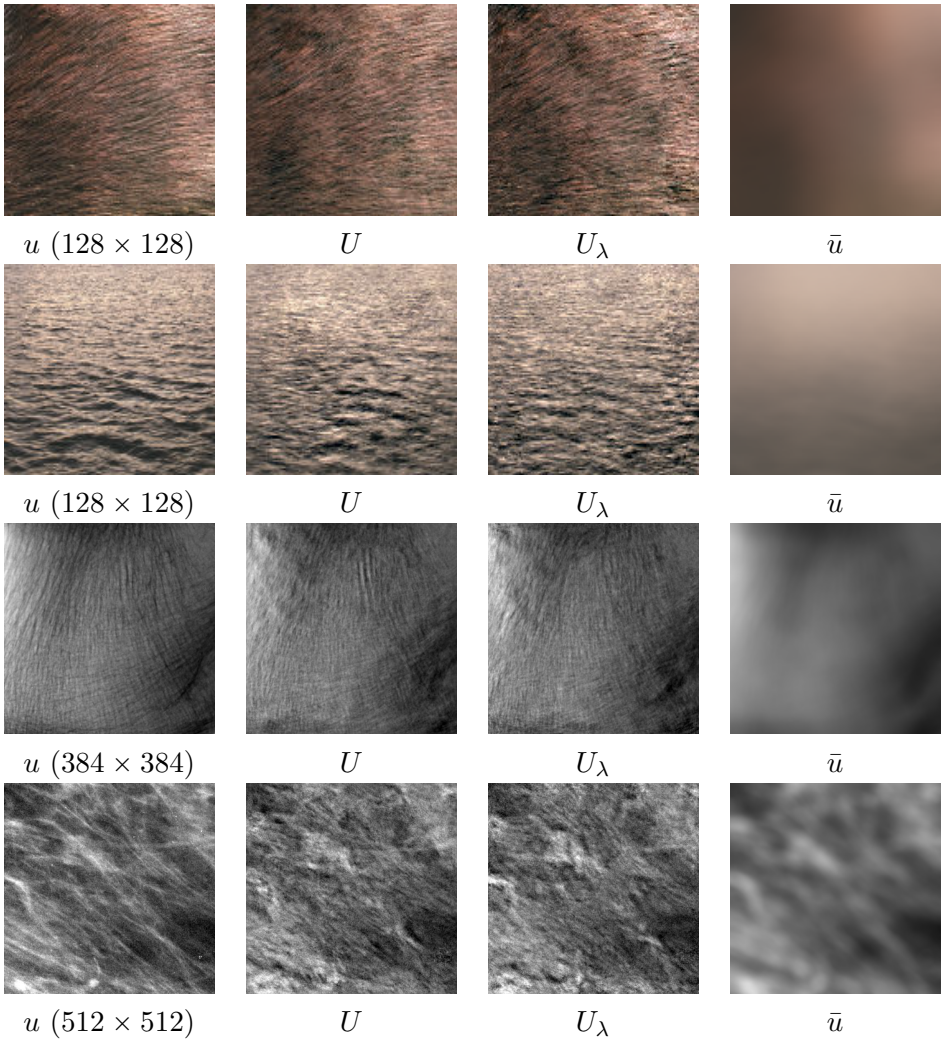


Figure 5.5: **Resynthesis of non-homogeneous textures using a local spot noise**. This figure shows local spot noise synthesis results for natural texture images. From left to right, one can see the original texture u, the corresponding LADSN U, LDSN U_λ, and the local mean image $\bar{u}$. The parameters were set to $r = \rho = 20$, and λ was chosen in order to set the mean number of impacts per pixels to 100. Notice that the synthesis of the three first examples is quite convincing. The model is indeed able to preserve the local orientation of the fur in the first example and of the fiber structures of the third example. However, strongly non-Gaussian features (such as the asymmetry of the distribution in the second case or the fiber structures of the fourth example (mammogram image)) are not preserved.

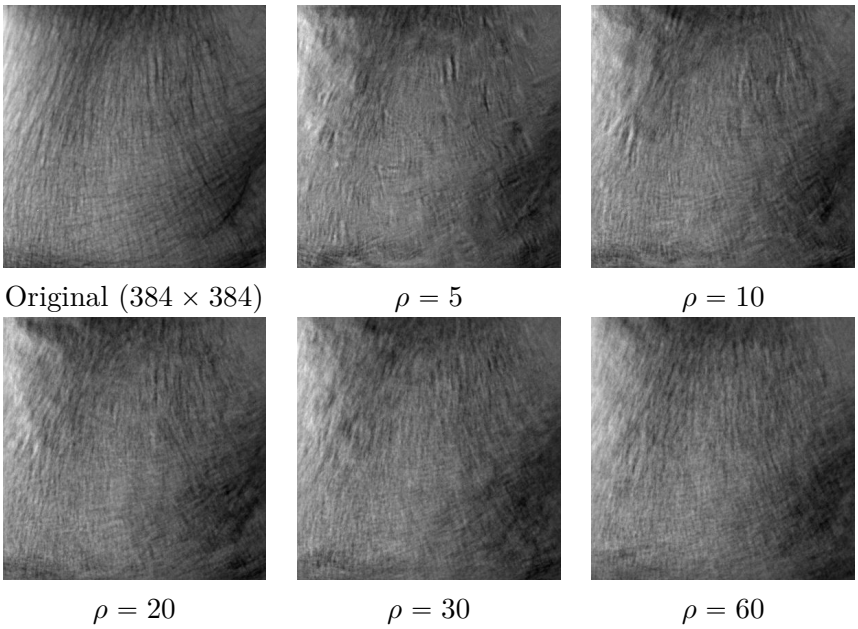


Figure 5.6: **Influence of the size ρ of the research zone.** In this figure, one can see different LADSN synthesis results with a fixed patch radius $r = 20$ and varying the research zone radius ρ. Notice that for low values of ρ, ringing-link patterns appear. When ρ increases, the local characteristics of the texture are less and less preserved.

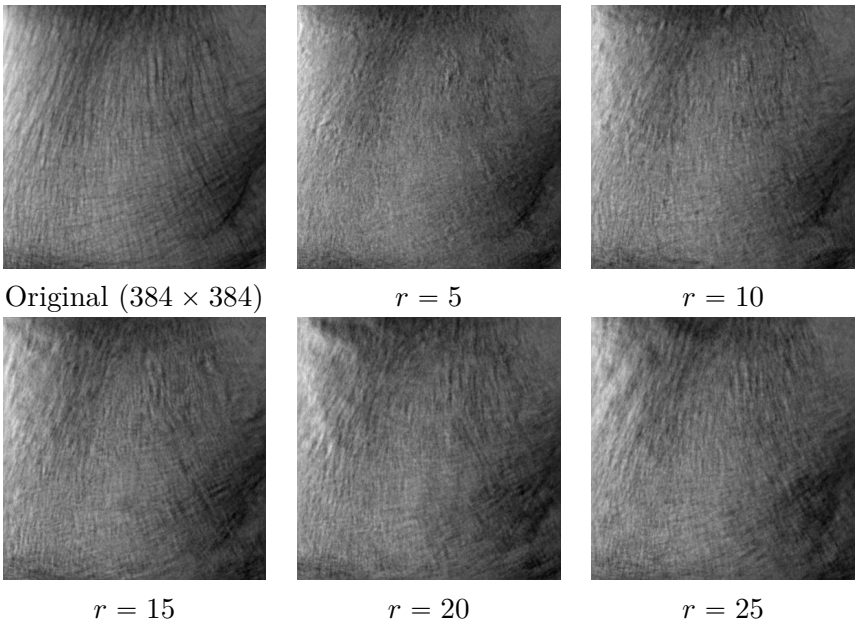


Figure 5.7: **Influence of the size of the patches.** In this figure, one can see different LADSN synthesis results with a fixed research zone radius $\rho = 10$ and varying patch radius r. Notice that r must be large enough in order to preserve the fiber structures and local orientations of the texture.

Since the local mean image is not extracted using an orthogonal decomposition, it may be difficult to go further and write a formal statement about the invariance to a smooth component. However, these experiments raise an interesting question about the invariance of our texture perception up to the addition of a smooth component. More precisely, it would be interesting to investigate the constraints on the smooth component s that make the images u and $u' = u + s$ to be perceived as the same texture (for example, in Fig. 5.9, it does not seem plausible to say that the images u and u' really represent the same texture).

Conclusion

We showed in this section that the local spot noise model could be used for resynthesis of a certain class of non random phase textures. The main interest of that preliminary work is that it allows us to better understand the limitations of non-stationary discrete Gaussian models. The experiments of this section indeed showed that the Gaussian model gets much richer when one relaxes the stationarity constraint, and in particular is able to respect non random phase features like certain fiber structures. The main drawback of the local spot noise model is that it is not defined as a texture model on $\mathbb{Z}^2$. However, in the future, extensions on $\mathbb{Z}^2$ could be designed by adopting the methodology presented in the next section.

5.3 Bi-Level Synthesis

In this section, we propose a methodology for texture synthesis that consists in first synthesizing a low-resolution version of the texture, and next add the details with a local function. We will see that this methodology allows to combine the flexibility of the Gaussian model (for coarse scale synthesis) and the richness of non-parametric sampling (for the refinement step).

5.3.1 Related Works

Before reviewing earlier works about multiscale texture synthesis, let us mention that there is some kind of ambiguity in the words "multilevel", "multiscale" or "multiresolution". They can refer to the progressive synthesis of textures either on finer and finer grids (with different sampling rates) or on finer and finer levels of a wavelet transform (with possibly different sampling rates if the wavelet transform is decimated). In the following, we will try to make a clear distinction between these two possible meanings.

Early attempts of multiscale texture synthesis were inspired by the recursive subdivision algorithm for the simulation of fractional Brownian motions, which amounts to sample the process on finer and finer grids. The one-dimensional method of stochastic interpolation presented in [Mandelbrot & Van Ness 1968] has been used in [Fournier *et al.* 1982] for rendering of stochastic surfaces, and later generalized in [Lewis 1987] to Gaussian processes with more general covariance functions. An undeniable advantage of these stochastic interpolation methods relies in their speed.

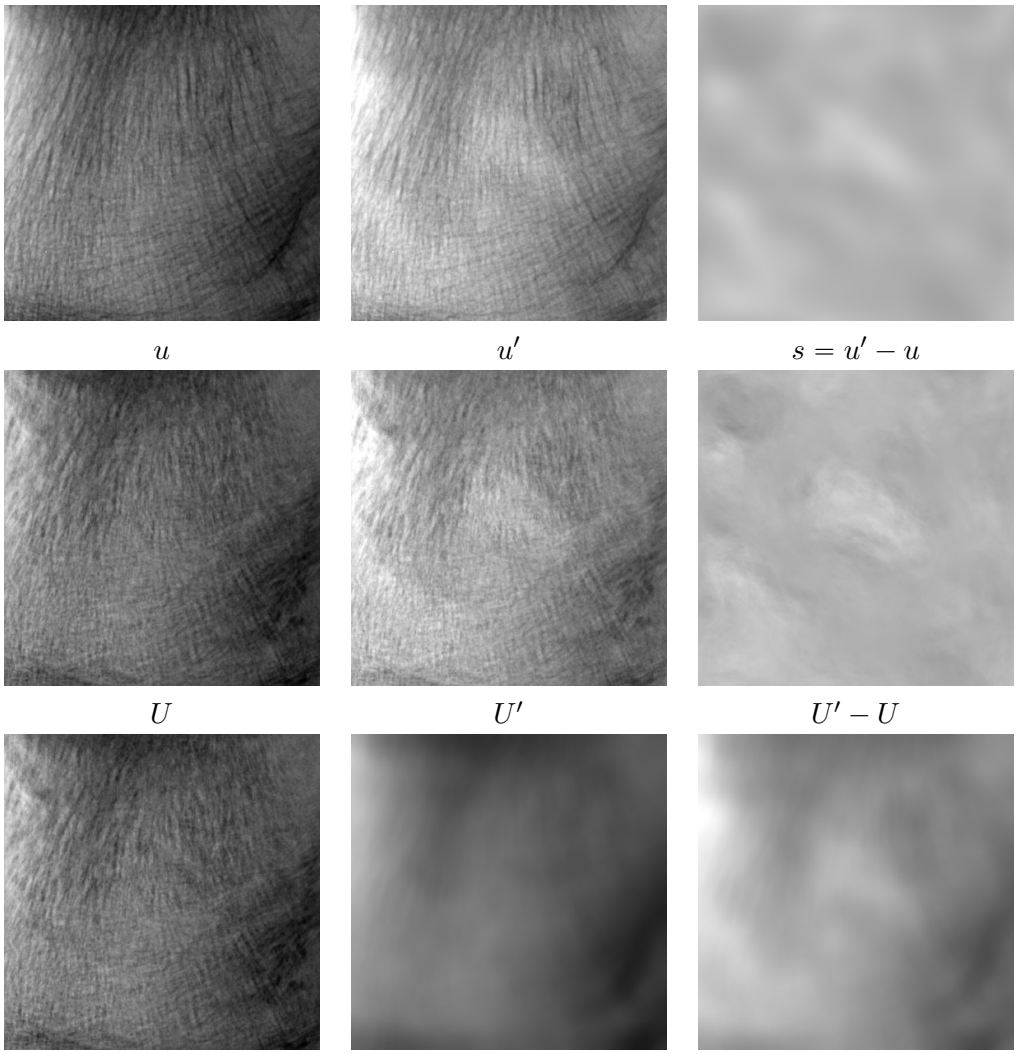


Figure 5.8: **Invariance to the addition of a smooth component.** The image u' was obtained by addition of a smooth component s to the original image u. Then, using the exemplars u and u', we carried LADSN synthesis with the same random seed to obtain the images U and U' (with $r = \rho = 15$). Notice that in this case, the images U and $U' - s$ are very similar. The image s is smooth enough so that the local autocorrelations are not affected by s.

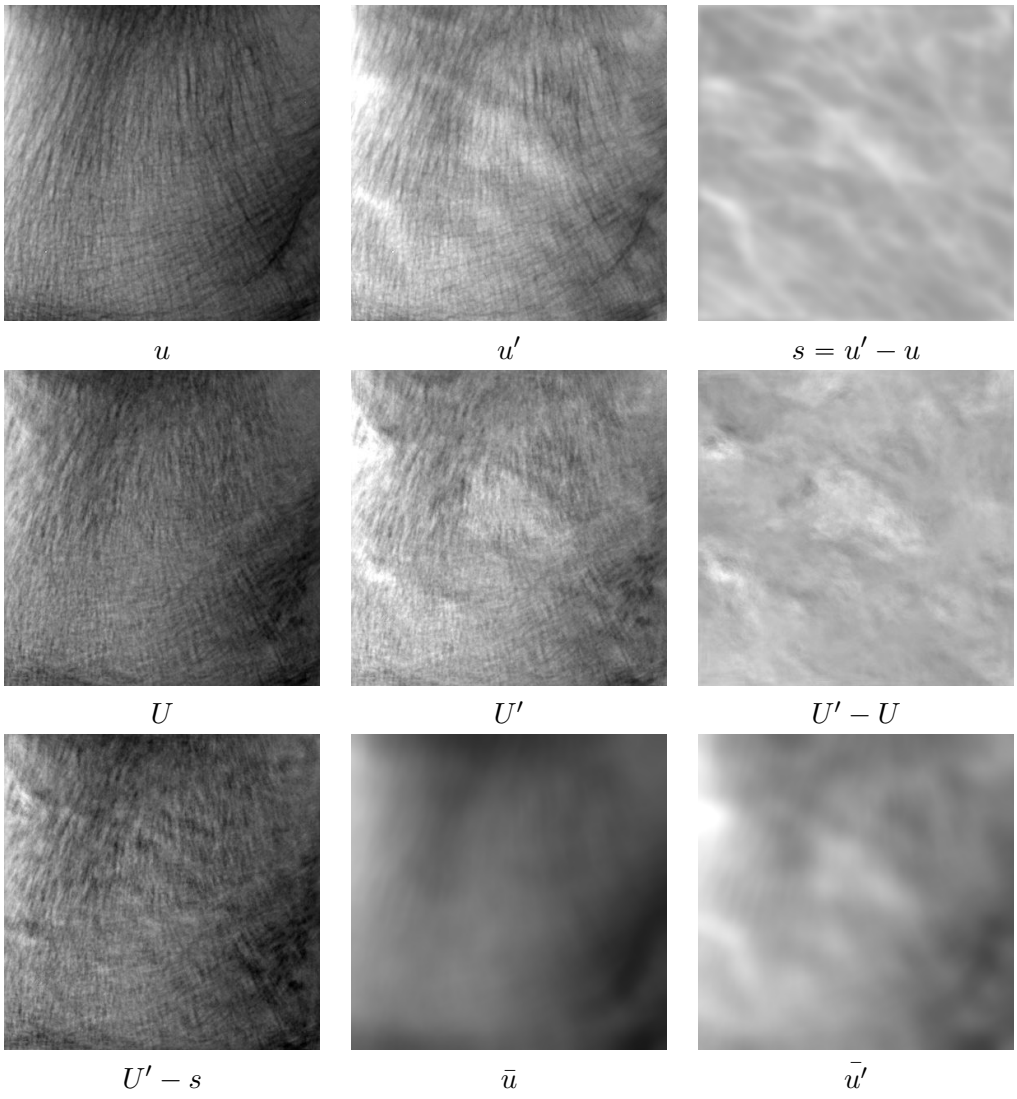


Figure 5.9: **Non-invariance to the addition of a less regular component.** The image u' was obtained by addition of a smooth component s to the original image u. Then, using the exemplars u and u', we carried LADSN synthesis with the same random seed to obtain the images U and U' (with $r = \rho = 15$). The corresponding results U and U' are shown in the second row. Notice that this time, the images U and $U' - s$ do not appear similar. In this case, the smooth component is not regular enough and thus changes the local autocorrelations.

However, these methods are only adapted to parametric covariance functions and thus not adapted *a priori* to by-example synthesis.

The authors of [De Bonet 1997], [Popat & Picard 1993], and [Paget & Longstaff 1998] proposed multiscale algorithms that allow for by-example synthesis. The author of [De Bonet 1997] uses a coarse-to-fine sampling scheme of the Laplacian pyramid associated to the output texture; each new level obtained by sampling the values of the exemplar pyramid that agrees at the parent coarser level. The algorithms [Popat & Picard 1993] and [Paget & Longstaff 1998] are closer to the subdivision methods mentioned above, because they amount to sample the output texture on finer and finer grids, each new pixel being sampled according to an estimated local conditional probability density function; in other words, these algorithms amount to perform multiresolution non-parametric Markov simulation. Unfortunately, these algorithms are much slower than [Fournier *et al.* 1982] and [Lewis 1987] because the estimation of the conditional distribution based on the coarser level is costly.

This conditional simulation step was simplified by the celebrated discovery made by [Efros & Leung 1999] who showed that, instead of estimating those distributions, one could directly perform approximate Markov sampling by drawing values in the exemplar texture at locations that agree with the previously sampled neighboring pixels. The single-resolution algorithm of [Efros & Leung 1999] was followed by multiresolution variants, for example in [Wei & Levoy 2000]. Several later articles presented multiresolution texture synthesis [Wei & Levoy 2002], [Lefebvre & Hoppe 2005], [Kwatra *et al.* 2005], [Han *et al.* 2008] but despite their impressive efficiency for reproducing structured textures, the complexity of these algorithms makes them hardly suited to mathematical analysis. In particular, it may be difficult to see if the output random field truly is a stationary random field. Wavelet representations suffer either from dependence on a grid in the case of decimated transforms (recall our stationarity constraint) or from non-trivial inverse calculation (for undecimated transforms). Let us mention however that the "texture optimization" method proposed in [Kwatra *et al.* 2005] is non-stationary only because of the grid-dependent interpolation procedure chosen by the authors. The following paragraphs build on the method of [Kwatra *et al.* 2005] to produce a truly stationary texture model.

5.3.2 Bi-level Models

Here, we propose to define bi-level models in a general framework, and we will study an instance of such models in the next subsection.

Given a blurring operator B, the associated bi-level model consists of the textures e such that the low-resolution image $e_b = Be$ is plausible in the Gaussian model and such that the details of e can be retrieved from e_b by a well-chosen local correction operation.

It is clear that such bi-level models encompass the Gaussian model for which we can take $B = \mathrm{Id}$. In order to illustrate that these bi-level models are richer than the Gaussian model, recall that a texture, seen from sufficiently far away can be con-

sidered to be a random phase texture [Galerne *et al.* 2011b]. This is corroborated by the fact that the plausibility in the Gaussian model tends to increase after a blurring operation, see Fig. 5.10. In this figure, one can see that several structured textures can be well synthesized at low-resolution by an ADSN model. Considering the work of Chapter 4, this can also be related to the fact that phase coherence indices decrease with blur.

As we have seen in Chapter 3, modelling and synthesizing the low-resolution component e_b can be done in a clear and efficient way using an ADSN random field. The main issue is thus to add the texture details in a way that respects the low-resolution content and that allows for mathematical analysis. Certainly, the better way to do this is to perform conditional sampling based on the low-resolution, as mentioned in [Chainais *et al.* 2011] or [Boussidi *et al.* 2014]. The difficulty of this approach in the example-based context is that it may be difficult to estimate and sample this conditional distribution. Instead of that, inspired by the work of [Efros & Leung 1999] and [Kwatra *et al.* 2005], we propose to emulate texture refinement by applying a local correction operator.

Definition 5.3.1. We say that a random field $V : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$ is a local function of a random field $U : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$ if there exists a finite neighborhood $\omega \subset \mathbb{Z}^2$ of 0 and a (deterministic) function $f : (\mathbb{R}^d)^\omega \longrightarrow \mathbb{R}^d$ such that

$$\forall \mathbf{x} \in \mathbb{Z}^2, \quad V(\mathbf{x}) = f(U_{|\mathbf{x}+\omega}) \ .$$

In the following, we will write $V = f(U)$, with a slight abuse of notation.

In other words, f modifies the value at pixel $\mathbf{x}$ by taking into account the local neighborhood of pixel $\mathbf{x}$. To sum up, a bi-level model is a random field V obtained as a local function f of a low-resolution stationary Gaussian random field U.

The minimal case $\omega = \{0\}$ amounts to apply a univariate function f to U, which can be used to prescribe the marginal distribution of the random field. But more interesting local functions are obtained by taking ω to be a patch domain; in this case f is able to apply geometric local corrections of U. We will see in the next subsection how to derive from an exemplar a patch-based local function.

But before that, keeping with this general setting, let us give simple properties of bi-level models. The most important one is a mathematical guarantee of textural innovation. Let us assume that U is a stationary Gaussian random field on $\mathbb{Z}^2$ whose covariance function C has a finite support S_C (which is always the case in ADSN synthesis as we have seen in 3). Assume also that V is obtained by applying a local function f to U. If $\mathbf{x}, \mathbf{y} \in \mathbb{Z}^2$ are sufficiently far away so that $\mathbf{y} - \mathbf{x}$ does not belong to the Minkowski sum $S_C + (\omega - \omega)$, then the Gaussian vectors $U_{|\mathbf{x}+\omega}$ and $U_{|\mathbf{y}+\omega}$ are independent, and thus, so are the values $V(\mathbf{x})$, $V(\mathbf{y})$. More generally if A, B are two subdomains of $\mathbb{Z}^2$ such that

$$(B - A) \cap (S_C + (\omega - \omega)) = \varnothing \ ,$$

then $V_{|A}$ and $V_{|B}$ are independent. This long-range independence property is a strong guarantee of innovation in the output random field V.

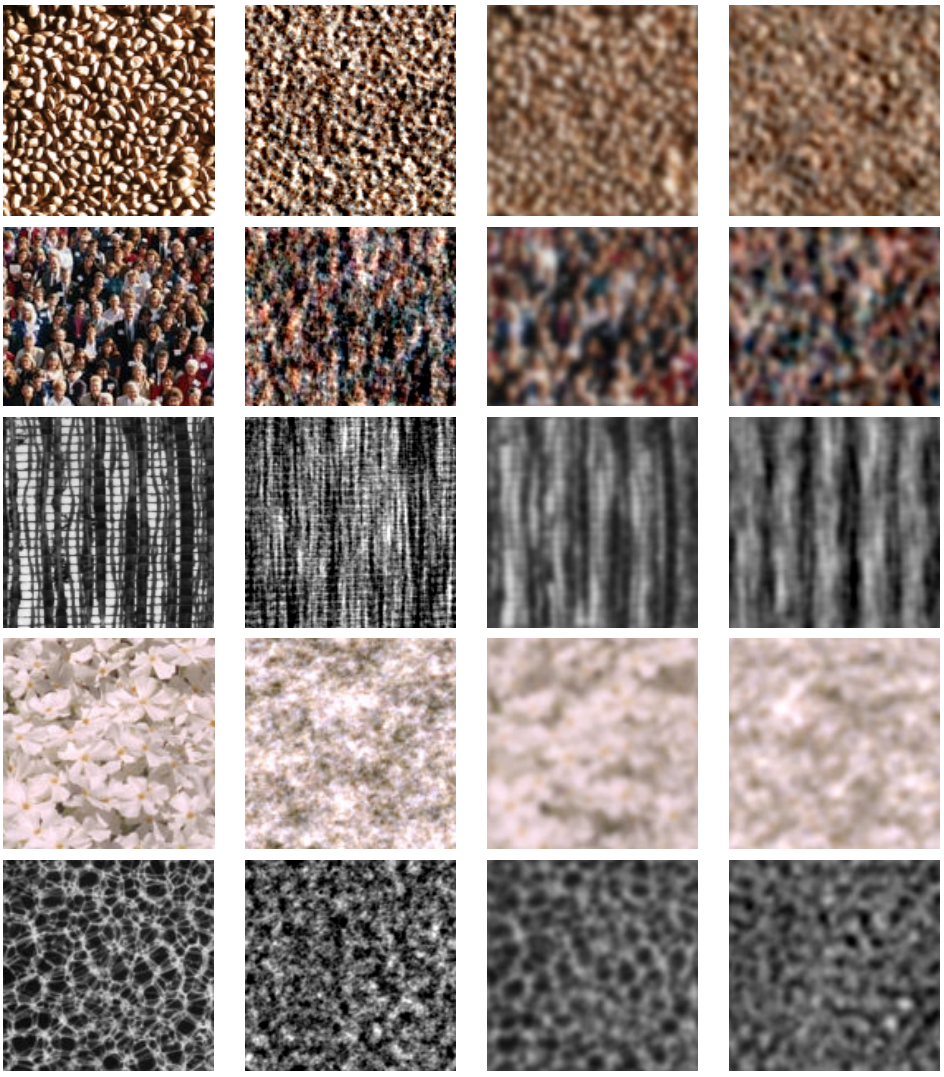


Figure 5.10: **Blur and random phase textures**. For each row and from left to right, we show an exemplar texture, a realization of the corresponding circular ADSN model, a blurred version of the exemplar (using the Gaussian kernel κ_ρ with $\rho = 2$) and a realization of the corresponding ADSN model. One can observe that the plausibility in the ADSN model increases when blur is added on the texture. In the three last cases, strongly non Gaussian features that are still visible in the blurred exemplar could be absorbed by a larger blur kernel.

Notice also that the distribution of $V(\mathbf{x})$ is exactly the push-forward distribution induced by the patch distribution $U_{|\mathbf{x}+\omega}$ and the local function f. Depending on the local function f, this push-forward distribution can be quite complicated, but it still connects the marginal distribution of V with the patch distribution of U. For example, thanks to this property, a univariate function f based on histogram modification can be used to prescribe the marginal distribution of V.

5.3.3 A Bi-level Synthesis Algorithm

In this subsection, we give an instance of bi-level model by precising the local function. In particular, we propose to derive from the exemplar image a patch-based function f that can be used for the refinement step.

Let us first mention that in the following bi-level model, B represent the convolution by a Gaussian kernel κ_σ of width σ pixels. In general, the parameter σ should be adjusted to the exemplar texture, but, for the sake of simplicity, in the following experiments, we always take $\sigma = 2$.

As explained in the last subsection, starting from an exemplar texture $e : \Omega \longrightarrow \mathbb{R}^d$ defined on a finite rectangular domain Ω, we first compute the low-resolution image $e_b = B(e)$. At low-resolution, the image e_b can be synthesized using a stationary Gaussian random field

$$U_b : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$$

as explained in Chapter 3. Next we will apply a local function to U_b in order to retrieve the textural details of the exemplar.

The local function will be obtained by composition of a few elementary patch-based operations inspired by [Kwatra *et al.* 2005]. So let $\omega \subset \mathbb{Z}^2$ be a finite neighborhood of zero (the patch domain), and let

$$\mathring{\Omega} = \{ \mathbf{x} \in \Omega \mid \mathbf{x} + \omega \subset \Omega \}$$

be the set of pixels having a full neighborhood in the exemplar texture. Let us denote by $p_{\mathbf{x}}(u) = u_{|\mathbf{x}+\omega}$ the patch of the image u that is centered on pixel $\mathbf{x}$. In order to compare patches, we will need a reference image e_{ref} (which will be either the exemplar e or the blurred exemplar e_b).

Let us now explain the elementary local functions. If $u : \mathbb{Z}^2 \longrightarrow \mathbb{R}^d$ is the current synthesis, e is the exemplar texture, and e_{ref} is the reference image, then we define the "correspondence map" $m : \mathbb{Z}^2 \longrightarrow \mathring{\Omega}$ by

$$\forall \mathbf{x} \in \mathbb{Z}^2, \quad m(\mathbf{x}) = m(\mathbf{x}|u, e_{ref}, \omega) := \operatorname*{Argmin}_{\mathbf{y} \in \mathring{\Omega}} \|p_{\mathbf{x}}(u) - p_{\mathbf{y}}(e_{ref})\| \ . \tag{5.16}$$

In other words, $m(\mathbf{x})$ is the location in $\mathring{\Omega}$ whose neighborhood in e_{ref} has the best patch similarity with the neighborhood of $\mathbf{x}$ in u. In the case where this minimum is attained for several values $\mathbf{y}$, one can choose one particular value in a deterministic way, for example the first one in raster order. Notice that $m(\mathbf{x})$ only depends on u through the values of $p_{\mathbf{x}}(u)$.

Based on this correspondence map, we can set

$$v(\mathbf{x}) = e\Big(m(\mathbf{x}|u, e_{ref}, \omega)\Big) \ . \tag{5.17}$$

Since $v(\mathbf{x})$ only depends of the values of u in the neighborhood of $\mathbf{x}$, this defines a local function

$$v = f_{e,e_{ref},\omega}(u)$$

(we will write $(v, m) = f_{e,e_{ref},\omega}(u)$ if we want to recall also the underlying map m). Let us emphasize (as in [Kwatra *et al.* 2005]) that this local function $f_{e,e_{ref},\omega}$ processes all the pixels independently which allows for parallel computation.

The overall local function will be obtained as a composition of a few elementary local functions with decreasing patch size. These steps perform geometric corrections of the texture in a more and more local manner, as illustrated in Fig. 5.11. In the following, we will use the square patch domain of size $(2s + 1) \times (2s + 1)$, denoted by

$$\omega_{2s+1} = \{-s, \dots, s\} \times \{-s, \dots, s\} \ .$$

We will use only five local correction steps with patch domains ω_{17}, ω_{17}, ω_9, ω_5, and ω_3. In order to apply the first local function to a low-resolution image, we need to compare patches in the low-resolution exemplar e_b; thus the first local function will be $f_{e,e_b,\omega}$. Let us mention that this choice of successive corrections may seem arbitrary but this choice must be related to the level of blur. Indeed, for the exemplars that we propose in the following, by observing the low-resolution version e_b, it seems reasonable to reinforce the geometry by using patches of dimensions $\leqslant 17$.

To summarize, after synthesizing U_b from e_b with a Gaussian model, we first compute

$$(V_0, M_0) = f_{e,e_b,\omega_{17}}(U_b) \ ,$$

where the patch comparisons are performed at low-resolution. Then, we apply successive local corrections at high-resolution with decreasing patch size.

$$\begin{aligned}
(V_{17}, M_{17}) &= f_{e,e,\omega_{17}}(V_0) \ , \\
(V_9, M_9) &= f_{e,e,\omega_9}(V_{17}) \ , \\
(V_5, M_5) &= f_{e,e,\omega_5}(V_9) \ , \\
(V_3, M_3) &= f_{e,e,\omega_3}(V_5) \ .
\end{aligned}$$

The output random field is then $V = V_3$. Eventually, all these local corrections can be written as a composition

$$f = f_{e,e,\omega_3} \circ f_{e,e,\omega_5} \circ f_{e,e,\omega_9} \circ f_{e,e,\omega_{17}} \circ f_{e,e_b,\omega_{17}} \ .$$

One can see that f is still a local function for the larger patch domain ω_{47}. This bi-level synthesis algorithm is summarized below and illustrated in Fig. 5.11.

Algorithm: Bi-Level Synthesis

- Input: Exemplar texture e.
- Get the low-resolution image: $e_b = \kappa_\sigma * e$.
- Synthesis at low-resolution:

$$u_b = \mathrm{ADSN}(e_b) \ .$$

- Texture refinement: apply the local function

$$v = f(u_b)$$

given by

$$f = f_{e,e,\omega_3} \circ f_{e,e,\omega_5} \circ f_{e,e,\omega_9} \circ f_{e,e,\omega_{17}} \circ f_{e,e_b,\omega_{17}} \ .$$

- Output: Synthesized texture v.

5.3.4 Results

In Fig. 5.11, one can observe a success of bi-level synthesis on a texture that is not a random phase texture (because of the asymmetric shady parts). At low-resolution, the random phase synthesis is satisfactory but not perfect (because of the asymmetry of the marginal distribution). But, as concerns the texture refinement, it is remarkable that a few number of local functions are able to transfer the texture details of the exemplar onto the low-resolution synthesis. Let us remark also that the pixel maps can be more irregular than the corresponding textures, which refutes that it is necessary to copy large portions of the exemplar to produce a convincing texture. In Fig. 5.11, we also illustrate the importance to initialize the refinement process with a low-resolution ADSN. Indeed, if it is initialized with a white noise image (whose marginal distribution is the same than e), then the large-scale structures of the exemplar are lost. In other words, the geometric structures of the synthesis are indeed prescribed by the low-frequency component.

One can observe in Fig. 5.12 and Fig. 5.13 several other results of bi-level synthesis. These examples confirm that the simple patch-based operator f greatly extends the random phase model, and is able to produce perceptually convincing results on a wide class of non random phase textures. Unfortunately, the local function f that we apply is still too complex to give the explicit distribution of the output random field v_1. However, some heuristics can be drawn (which corroborate the last remarks of Subsection 5.3.2): for example, if the patches of e_b are approximately uniformly represented in the synthesis u_b, then the marginal distribution of v_0 will be close to the empirical marginal distribution of e.

Nevertheless, as we have seen in Subsection 5.3.2, the bi-level model has two important assets given by the stationarity of the random field and the long-

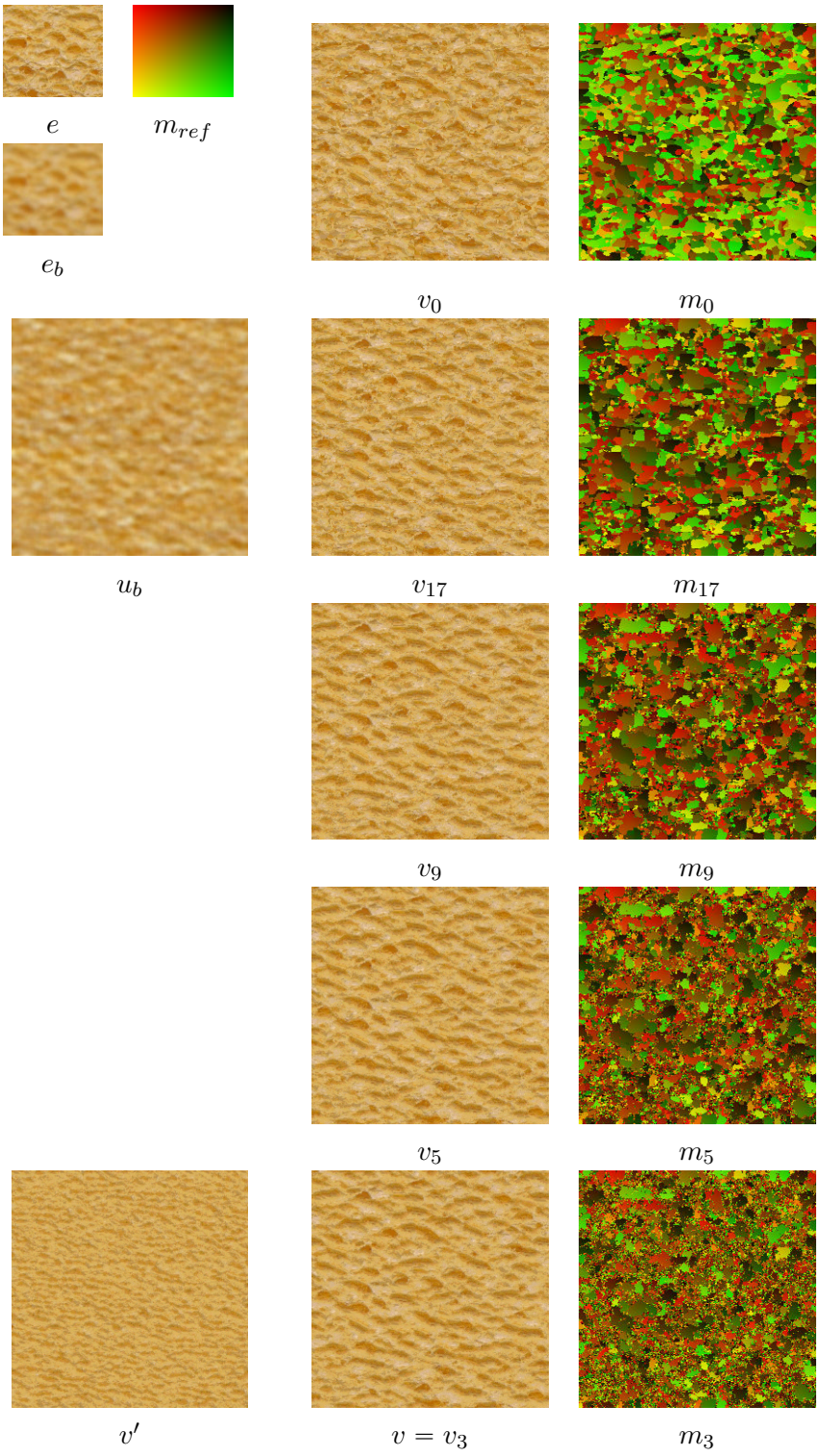


Figure 5.11: **Bi-Level synthesis**. In the upper left, one can see the exemplar texture e (of size 108×99), the reference map m_{ref}, and the low-resolution exemplar e_b (obtained with a Gaussian kernel κ_ρ with $\rho = 2$). Just below, one can see the low resolution synthesis u_b. In the second column, one can see the progressive results of texture refinement. In the bottom left, we also show the result v' of the refinement step applied on a white noise image. The correspondence maps are shown on the third column. The output texture v is a convincing synthesis of the non random phase texture e. Comments are given in the text.

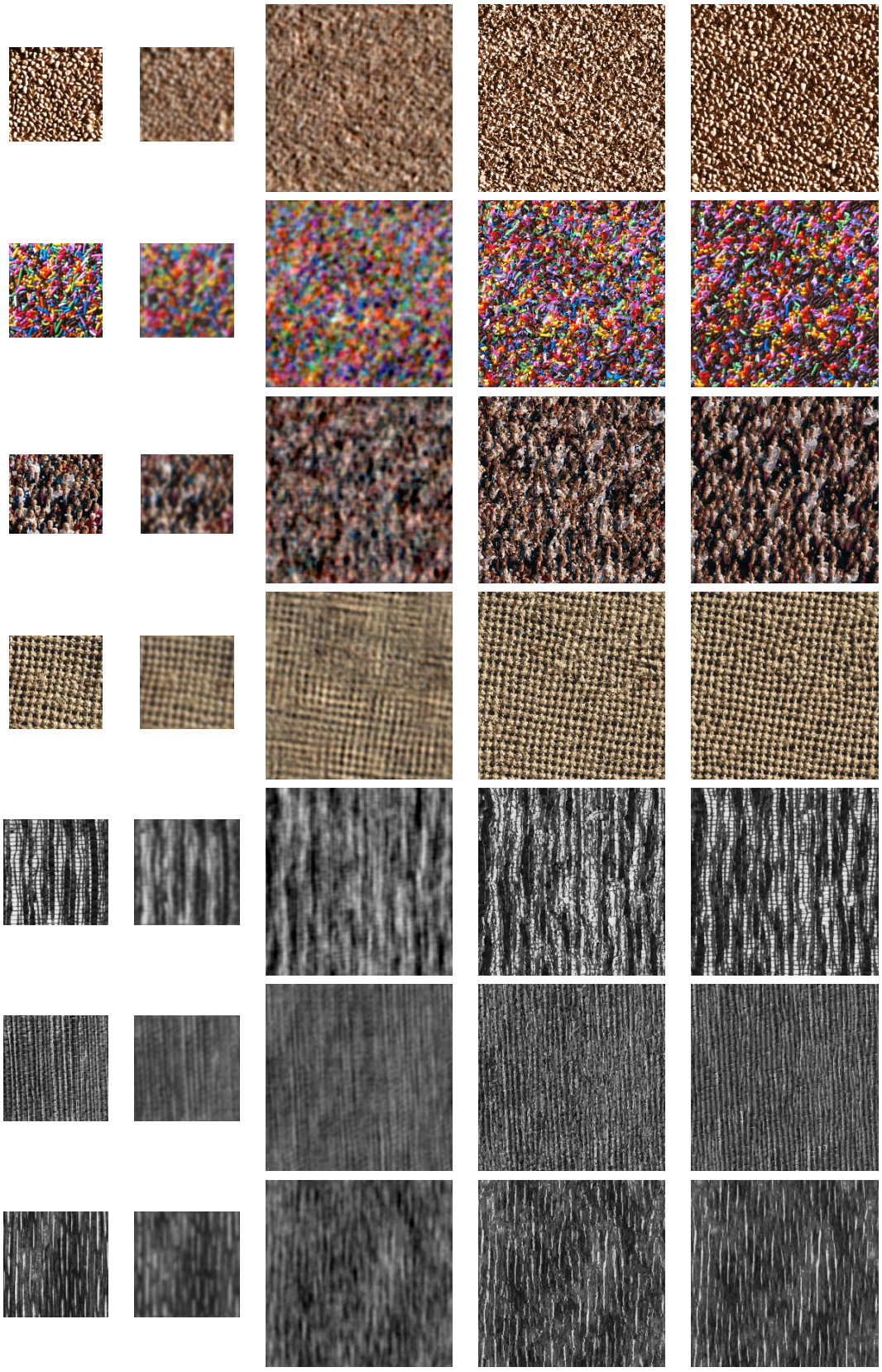


Figure 5.12: **Successful examples of bi-level synthesis**. Each row displays one successful case of bi-level synthesis. For each row, from left to right, one can see the exemplar texture e, its low-resolution version e_b, the low-resolution synthesis u_b, the image $v_0 = f_{e,e_b,\omega_8}(u_b)$, and the image v obtained after the successive local corrections.

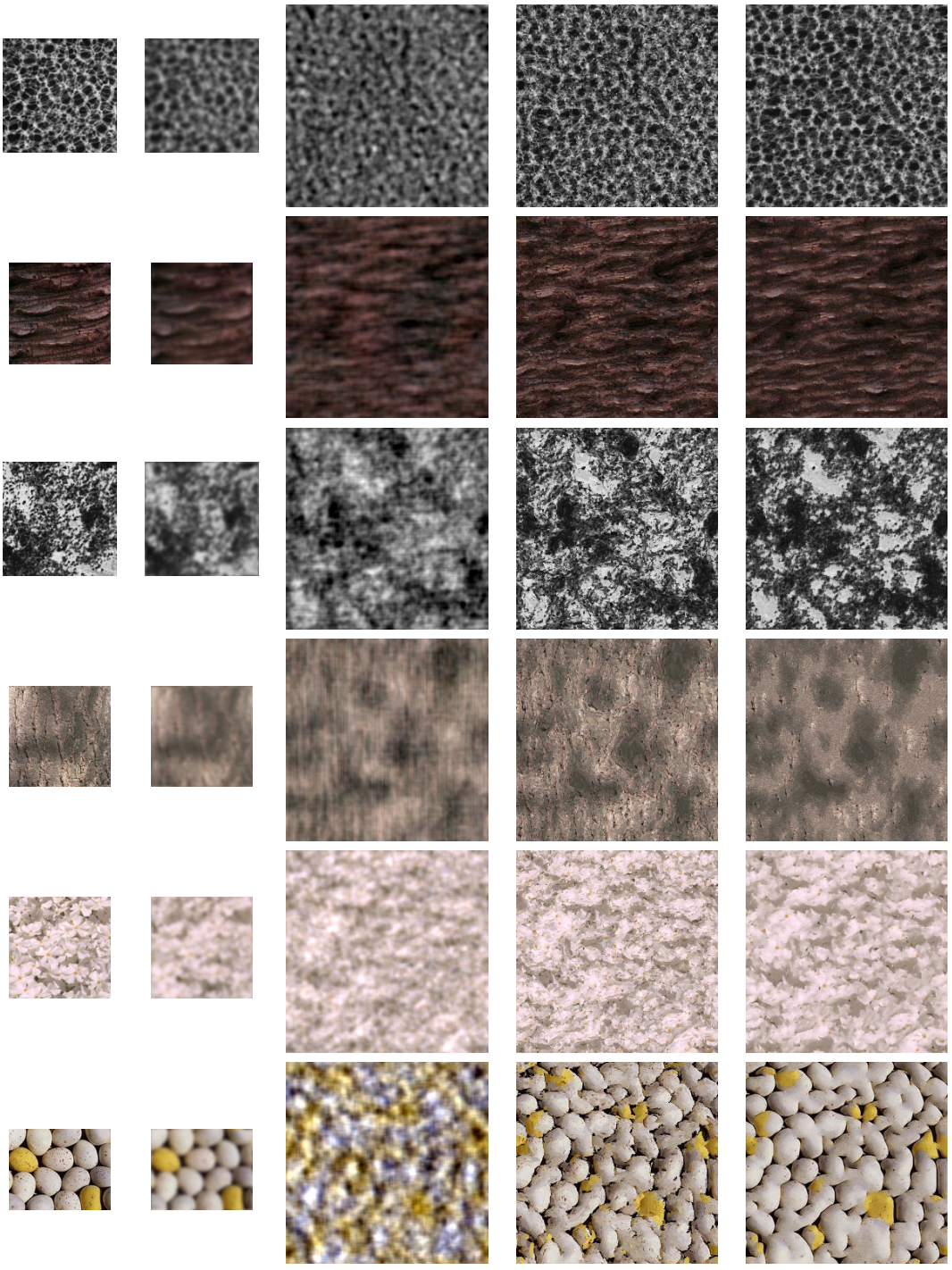
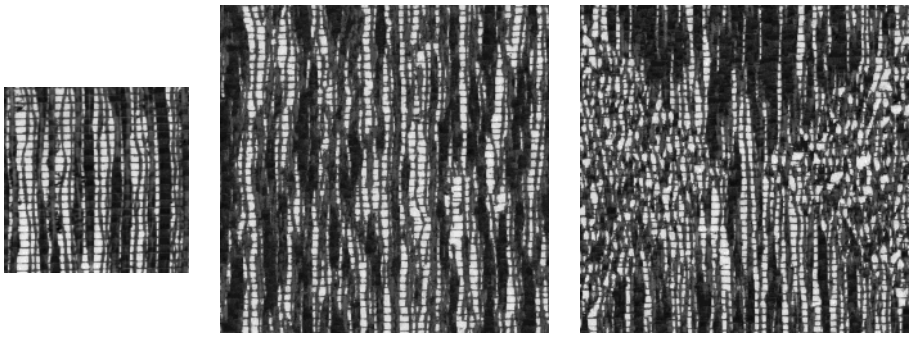


Figure 5.13: **Unsuccessful examples of bi-level synthesis**. Each row displays one failure case of bi-level synthesis. For each row, from left to right, one can see the exemplar texture e, its low-resolution version e_b, the low-resolution synthesis u_b, the image $v_0 = f_{e,e_b,\omega_8}(u_b)$, and the image v obtained after the successive local corrections. In the two first cases, the output texture is more regular than the exemplar. In the third and fourth row, the patch distribution of the exemplar is not respected. And in the two last rows, the exemplar textures have complex geometrical structures that cannot be represented by this model (in the last case, the random phase hypothesis is clearly not satisfied for e_b).



Original Bi-Level Synthesis Efros-Leung

Figure 5.14: **Growing garbage**. From left to right, the original texture, the result of bi-level synthesis, and the result of Efros-Leung algorithm (produced with the online demo of [Aguerrebere *et al.* 2013] with patches of size 7×7 and a tolerance value of 0.2). One can notice that in the result of Efros-Leung algorithm, there is a clear spatial boundary between a subdomain where the synthesis is satisfactory and another subdomain where the synthesis fails. Beyond this boundary, we say that the algorithm is "growing garbage". In contrast, the stationarity ensures that this effect will not appear with the bi-level synthesis algorithm proposed in this section.

range independence, which are good guarantees of stability. Indeed, in contrast to the method of [Efros & Leung 1999] (which can be tested with the online demo [Aguerrebere *et al.* 2013]), bi-level synthesis is unlikely to produce a satisfying texture that degenerates beyond a spatial boundary. In some way, the result with bi-level synthesis is more predictable: either the synthesis completely fails (even on a small domain), either it works on a very large domain. We have shown an example of this "growing garbage effect" in Fig. 5.14. As a result of this stability, the bi-level algorithm can be used to synthesize textures on very large domains, as one can see in Fig. 5.15. With this model, the synthesis on very large domains does not face any theoretical obstacle but is only limited by the available computational time.

Let us now highlight the limitations of this synthesis algorithm. Since the refinement step was not built to solve a well-defined inverse problem, we are not ensured that the patch distribution of the output texture precisely respects the one of the exemplar. In particular, in some cases (for example in the third and fourth row of Fig. 5.13), the algorithm will tend to over-regularize the texture. One reason is that the comparison of noisy patches tends to favor smooth patches: for example, as was discussed in [Newson *et al.* 2014], if P_1 and P_2 are two independent Gaussian white noise processes on the same patch domain, then the expected square distance $\|P_1 - P_2\|_2^2$ between P_1 and P_2 is twice the expected square distance $\|P_1\|_2^2$ between P_1 and 0). Another source of failure concerns textures with a complex local geometry as the fifth example of Fig. 5.13. The flower shapes of the exemplar (each shape occupying approximately 32×32 pixel) are not retrieved by the local function f. For such examples, it is likely that precise (multi-pass) Markov sampling would be required to better reproduce this complex textural unit.

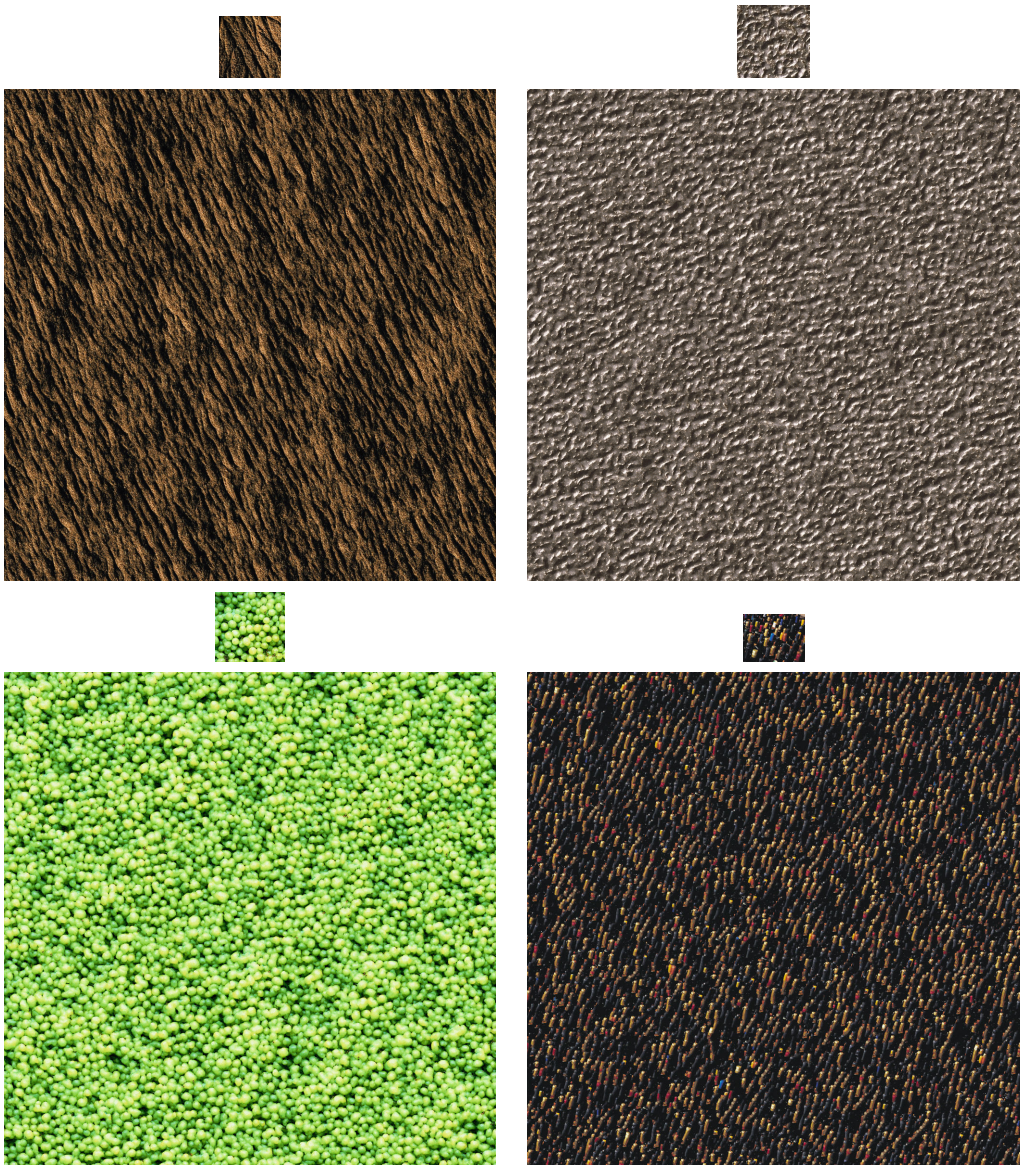


Figure 5.15: **Bi-level synthesis on very large domains**. First and third row: exemplar textures (of size 128×128). Second and fourth row: bi-level synthesis (of size 1024×1024) on the right. These examples illustrate the stability of bi-level synthesis (which comes from the model stationarity and the long-range independence).

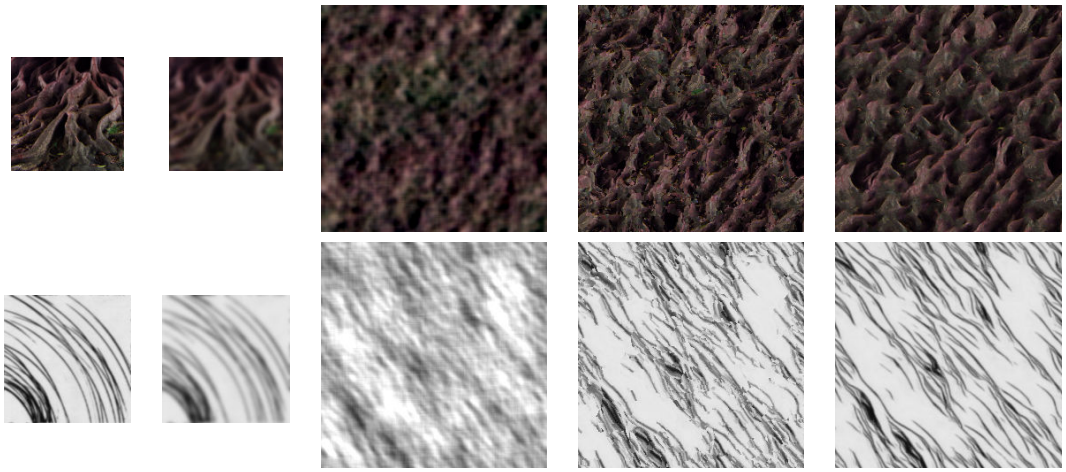


Figure 5.16: **Texturization**. For each row, from left to right, one can see the exemplar texture e, its low-resolution version e_b, the image $v_0 = f_{e,e_b,\omega_{17}}(u_b)$, and the image v obtained after the successive local corrections.

Of course, this algorithm is expected to fail on textures that do not agree with the bi-level model; for example, in the last example of Fig. 5.13, the low-resolution exemplar is clearly not a random phase texture and thus the refinement step creates unrealistic shapes (about that, let us mention that stochastic texture models are certainly not the best tool to deal with textures having strong physical constraints, like images of stacked objects). Similarly, the algorithm will fail on non-homogeneous textures like the examples presented in Fig. 5.16. In particular, one can question the fact that the second exemplar image of Fig. 5.16 can be called texture; apart from the completion of the circular shapes, what we expect of such an image synthesis is not clear at all. However, the first example is more interesting because it modifies the exemplar in such a way to obtain a homogeneous texture; this "texturization" process raises interesting questions about the patch distribution of stationary processes that will be discussed in the end of this thesis.

In conclusion, the purpose of this section was not to propose a universal texture synthesis algorithm, but more to show that by simple extensions of random phase models, one can design richer texture models that allow for a minimal mathematical analysis. It demonstrated once again how texture modelling can benefit from the combination of random phase models and patch-based operations. We used patch-based local functions in the texture refinement. Future research may lead to other ways to solve that step, which could improve the bi-level synthesis algorithm, and also increase the mathematical understanding of such bi-level models.

Chapter 6

Conclusion

The following paragraphs give an overview of the main contributions of this thesis accompanied with perspectives for future research.

6.1 Random Phase Models

In Chapter 2, we presented a unified framework for texture modelling with a particular attention given to spectral representations. In this setting, we were able to compare several well-known stationary random models (discrete spot noise, moving-average, autoregressive, and Markovian models). We emphasized that presenting texture synthesis as sampling of stationary random fields defined on $\mathbb{Z}^2$ imposes an interesting constraint whose practical impact is a guarantee of stability for the synthesis algorithms (thus avoiding the famous growing garbage effect encountered with the non-stationary method of [Efros & Leung 1999]).

In this framework, we recalled the main properties of DSN models, and pursued the works of [Van Wijk 1991], [Galerne *et al.* 2011b], [Desolneux *et al.* 2012], and [Xia *et al.* 2014]. Borrowing the last and informal conclusion of [Van Wijk 1991], we would say that "Spot noise is a hot noise" and we hope that the multiple developments given in the two first chapters of this dissertation suffice to illustrate that it has not cooled down yet.

With the synthesis-oriented texton, we realized one of the perspective suggested in [Van Wijk 1991]. Indeed, the SOT realizes a compact summary of a Gaussian texture, which holds in a prescribed spatial support, and can be used for direct spot noise synthesis. In terms of Gaussian model approximation (measured by the optimal transport distance), the SOT outperforms the canonical (or luminance) texton of [Desolneux *et al.* 2012]. We showed that, using a SOT with small support (31×31), a large class of microtextures can be synthesized with a low complexity (less than 50 operations per pixel) and with the usual benefits of spot noise synthesis (for example the possibility of local parallel evaluations).

The main weakness of the SOT is certainly that we do not properly measure the DSN visual convergence towards its Gaussian limit. Indeed, the efficiency of the SOT in DSN synthesis is mainly due to the random phase initialization of the alternating projections algorithm. If we were able to express the DSN convergence speed as a function $f(h)$ of the kernel h, then we could compute a SOT by optimizing a functional taking into account both $f(h)$ and the distance to the reference Gaussian model. Some work has already be done in this direction by considering the Kolmogorov distance between the marginal distributions. Even if it was more

satisfactory on a mathematical point of view, it did not lead to significant improvement in comparison with the original SOT, because marginal distributions do not suffice to reflect visual convergence. The difficult question of visual convergence is linked to the precise assessment of the plausibility in the RPN or ADSN models, which will be discussed later.

This work on the SOT also triggered many questions about the optimal transport distance between ADSN models [Xia *et al.* 2014], [Desolneux *et al.* 2015]. Our theoretical advances were presented in Chapter 2 where, for example, we extended this optimal transport distance to the case of stationary random fields on $\mathbb{Z}^2$. This extension has not found its applicative context yet. One possible application would be to realize a proper ADSN approximation of general Gaussian random fields defined on $\mathbb{Z}^2$. In the gray-level case, this amounts to approximate a general covariance function by a compactly-supported one, and thus can be seen as a generalization of the SOT with a continuous frequency domain. Also, it could be interesting to explore the dual approach, that is, to perform compact approximation of the convolution inverse of the covariance function; as shown in Subsection 2.2.4, this is equivalent to realize a Markov approximation of the random field. The practical impact of such a Markov approximation is not clear (because sampling a GMRF is not easy); however, the comparison of these dual approaches may help to clarify the concept of texture scale at least in the case of Gaussian textures.

On a more practical point of view, as explained in Subsection 3.3.3, the usual optimal transport distance has no reason to precisely reflect our texture perception. On the one hand, it does not reflect enough the importance of color distribution in texture perception, as illustrated by the color correction needed in the SOT computation and presented in Subsection 3.3.2. And more importantly, it does not take into account that the human visual system is more sensitive to high frequencies, as shown in Subsection 3.3.3. In Subsection 3.3.4, we proposed to integrate frequency weights in the optimal transport distance to get a frequency balance that complies better with our texture perception. This modified distance led to a more robust version of the SOT. It is likely that the SOT methodology could be further improved by minimizing functionals that could deal with the color distribution (using a Wasserstein term as in [Tartavel *et al.* 2014]), and even more complex terms (related to the DSN visual convergence speed, as mentioned above).

Another promising perspective to pursue this work would be to define a continuous SOT in order to address procedural texture synthesis. This would require adapting the SOT methodology to a continuous synthesis domain using interpolation functions. This work would thus be a natural and non-parametric extension of the methods presented in [Lagae *et al.* 2009] and [Galerne *et al.* 2012]. Considering the very low number of impacts per pixel involved in the SOT-based spot noise synthesis, a procedual synthesis algorithm based on a continuous SOT is expected to perform at least 10 times faster than these two former procedural methods. One challenge in this continuous framework is to realize a balance between the irregular aspect of textures and the regularity of interpolation functions.

Finally, we would like to emphasize again that the mathematical benefits of the Gaussian texture model have not been drained out yet. Beyond the availability of

fast exact simulation scheme, the possibility of conditional simulation is certainly under-used as regards image synthesis. In Section 3.4 we proposed a textural inpainting algorithm based on Gaussian conditional simulation. We have seen that this algorithm is able to fill very large holes in microtextures, with a reasonable computation time. Adopting a texture model for the inpainted content is a good way to prevent over-regularization of textures, which is a common drawback of many inpainting algorithms as mentioned in [Newson 2014]. The work reported in Section 3.4 about textural inpainting is still unachieved. Indeed, it remains to adapt the kriging methodology to color textures, and besides, we have to explore the algorithmic improvements that could be drawn from Markov approximations of the model [Hartman & Hössjer 2008]. To end this paragraph, let us mention that, considering the recent work [Raad *et al.* 2015], it seems that conditional Gaussian simulation can also be used to address synthesis of more structured textures, as will be discussed below.

6.2 Phase Coherence Indices

One main goal of this thesis was to design non random phase texture models that would be adapted to macrotextures or at least textures presenting edge-like structures. In the beginning of this PhD, we were forced to observe that the direct analysis or synthesis of the global Fourier phase information is a difficult problem, as illustrated by the experiments reported in Section 5.1.

However, in Chapter 4, we have seen that the *a contrario* methodology allows to tackle this difficult question, by comparing the TV of a given image with the generic TV in an associated random phase model. Following the works of [Blanchet *et al.* 2008] and [Blanchet & Moisan 2012], we provided a thorough study of the phase coherence indices GPC, SI and S, with both theoretical and practical aspects. In particular, we presented several experiments that clarify the link between phase coherence and image quality, and confirmed that these indices can be interpreted as sharpness indices because they are sensitive to blur, noise and ringing artifacts. Also, in Section 4.5, we have shown that these indices could be used to address blind deblurring based on a simple stochastic optimization scheme. The resulting algorithm is able to deal with isotropic blur by choosing an interesting deblurring kernel that enhances the image while keeping a limited amount of noise and ringing.

This work on global phase coherence may be pursued in several ways. On the probabilistic side lies the possibility to study more precisely the TV of random phase fields. We have seen that the construction of the phase coherence indices GPC and SI is based on the random variables $\mathrm{TV}(u_\psi)$ and $\mathrm{TV}(u * W)$ which represent the TV of the RPN and ADSN models, respectively. In the appendices of Chapter 4, we proposed some significant advances on these random variables, showing in particular that the expectation of $\mathrm{TV}(u_\psi)$ can be approximated by the one of $\mathrm{TV}(u * W)$. The bound given in Appendix 4.A is too rough and does not reflect the quality of the approximation that has been observed through Monte-Carlo simulations. Since it

has been obtained by aggregating approximations given by a Berry-Esseen theorem, this bound could certainly be improved using more evolved results from probabilistic ergodic theory.

Not only a further study of $\mathrm{TV}(u_\psi)$ and $\mathrm{TV}(u * W)$ would allow to better understand GPC and SI, but it would especially help to clarify the link between the RPN and ADSN models. In particular, it is a major question to know if there exists an infinite analog of RPN defined on the whole discrete plane $\mathbb{Z}^2$. As explained in Subsection 2.2.3, we conjecture that the answer is no because of the Gaussian asymptotic behavior of RPN fields [Desolneux *et al.* 2015] (in terms of the finite-dimensional marginal distributions). However, since the asymptotic behavior of $\mathrm{TV}(u_\psi)$ cannot be inferred only from finite-dimensional marginal distributions, its study could shed a new light on that question. In particular, as mentioned in Subsection 4.3.3, the second-order moment of $\mathrm{TV}(u_\psi)$ was observed in practice to be much less than the one of $\mathrm{TV}(u * W)$; if we prove that this property persists in an asymptotic framework, then we would claim a crucial asymptotic difference between the RPN and ADSN fields.

Closer to imaging applications, the link between phase coherence and image quality and the success obtained in blind deblurring seem to indicate that these indices can certainly be used to address other image processing tasks. For example, designing phase coherence measures that are localized in the frequency domain would allow to identify the most relevant parts of the image DFT. Such localized measures could be used to define a notion of effective resolution. Considering that the cameras are now able to acquire images with many mega-pixels, it becomes more and more useful to identify and extract the truly relevant content of an image. Beyond data compression, a long-term objective on this topic is to devise innovative adaptive reduction algorithms based on phase coherence measures. The need of a robust and automatic algorithm to reduce an image to its significant content is justified by the consumer need of an efficient storage and faithful printing of pictures.

Another practical avenue of research is to design new variants of SI and S that can be more easily interpreted and optimized. Indeed, SI and S can be computed with closed-form formulae, and experiments show that they are somehow linked to the image quality. But we do not fully understand why these formulae reflect some kind of sharpness. On the contrary, when considering the TV operator, the continuous analysis of the bounded variation functions (and in particular, the coaera formula) helps to understand which images have a small total variation (leading to the widely adopted denomination of "cartoon images"). It would be interesting to derive from the indices SI or S a new simplified index which would still be correlated to the image quality and which would have a more readable link to the geometrical content of the image (as the one we have for TV). In this simplification, one can also hope to gain some more analytical properties, which would allow to integrate this new index as an efficient prior in optimization problems addressing image restoration tasks. Notice by the way that we may not necessarily require both convexity or smoothness since some important progress is currently made in non-convex non-smooth optimization (as in [Bolte *et al.* 2013] for example).

Finally, further research would be needed to extend the *a contrario* methodology adopted for GPC in order to measure more precisely the plausibility of a texture image in the RPN/ADSN model. This would improve the analysis available through the GPC index because, as we have seen in Subsection 4.3.3 and Subsection 5.1.3, a low GPC value only provides a necessary condition to be a realization of a random phase field. To deal with this problem, we would have to identify more precisely which geometrical details of a texture makes it non-plausible in the Gaussian model. This question is thus related to texture perception, and stochastic geometry. Inspired by [Julesz 1981] and [Malik *et al.* 1999], it seems relevant to adopt an *a contrario* framework based on several simultaneous filter responses, which could profit from recent advances in extended *a contrario* approaches [Myaskouvskey *et al.* 2013]. Answering this question would define a kind of distance to the ADSN model and thus allow for *a priori* evaluation of the performance of ADSN synthesis. Besides, if the corresponding methodology could be adapted to other texture models (possibly patch-based models), the corresponding distances could be used to choose a synthesis algorithm for a particular texture sample, and could also be aggregated to compute a measure reflecting the complexity of the texture sample.

6.3 Non Random Phase Texture Models

The phase coherence indices presented in Chapter 4 represent a significant step in the analysis of the phase information. As concerns phase-sensitive texture synthesis, the explorative work presented in Chapter 5 must be pursued. In other words, it is still a widely open question to design clear texture models that allow for efficient macrotexture synthesis while keeping strong mathematical guarantees. In Section 5.1, we discussed the phase constraints due to stationarity, and concluded that direct phase synthesis is not an appropriate way to tackle the problem.

Considering their success in texture synthesis, we thus tried to design texture models that combine the benefits of spot noise models and patch-based models. In Section 5.2, we proposed a local spot noise model that can be used for resynthesis of non-stationary microtextures. We illustrated that this model is able to reproduce certain non random phase features, thus showing that Gaussian models with relaxed stationarity constraints can outperform the ADSN model. This corroborates the findings of [Raad *et al.* 2014]. This work opens three perspectives.

The first perspective is to use this model in order to tackle texture analysis. Indeed, the synthesis results show that this model is sensitive to certain local characteristics of the texture that are accessible through the patches autocorrelation. Finding a way to extract and represent these local characteristics would define local textons that could be used for example in segmentation [Malik *et al.* 1999].

The second perspective is to derive from the local spot noise an extended stationary model on $\mathbb{Z}^2$ that enters the framework of Chapter 2. One possibility would be to perform synthesis of the low-frequency component, and then to add a local spot noise with a choice of spots driven by the low-frequency component. This would

constitute another instance of bi-level model. Again, this bi-level model inherits the stationarity from the one of its low-frequency component.

The third perspective is to design other variants of the spot noise process that are able to produce strongly non-Gaussian features. Of course, since a general central-limit theorem holds for high-intensity Poisson spot noise, one must modify the way to throw the kernel functions. One possibility is to consider a Markovian variant in which kernels are progressively thrown onto the texture with, at each step, a choice of kernel that depends on the previous texture state. We already have experimented simple examples of such models (for example throwing segments in order to reinforce the gradient magnitude) and we observed that they are undeniably able to produce non-Gaussian features. However, for now, the mathematical analysis of such Markovian spot noise processes seems difficult (for example, the model renormalization that is required to ensure convergence in distribution is not clear).

In Section 5.3, we also proposed to address macrotexture synthesis with a methodology that we called bi-level synthesis. This methodology amounts to first synthesize a low-frequency component of the texture and then add the textural details with a local refinement. In Section 5.3 we proposed an instance of bi-level model by adopting a Gaussian model for the low-frequency component, and by applying local corrections with a simple patch-based operation inspired by [Kwatra *et al.* 2005]. This bi-level model is naturally defined as a stationary texture model on $\mathbb{Z}^2$, and is able to synthesize structured textures on a very large domain with a guaranteed stability.

One could imagine several improvements of the bi-level algorithm proposed in Subsection 5.3.3. The main weakness of this algorithm is that the local functions were not chosen to answer a clear mathematical problem. A way to alleviate this problem would be to allow for non-deterministic local functions, for example by relying again on conditional simulation. Here it would be required to perform random sampling conditioned on the low-frequency component. Not only this would help to get an even clearer model, but this would also increase the innovation capacity of the algorithm. Indeed, for now, the randomness only comes from the low-resolution synthesis, and the output values are restricted to the pixel values of the exemplar, which drasticly constrains the possibility of innovation. Besides, this conditional step could rely on non-stationary operations, which are a relevant way to produce geometric features, even in the Gaussian case as demonstrated by Subsection 5.2 and [Raad *et al.* 2015].

Returning to the algorithm of Subsection 5.3.3, it is important (and actually quite surprising) to notice that so few applications of the patch-based local functions suffice to reproduce complex local geometric structures in a very convincing way (despite that all the pixels are processed independently by these local functions!). It confirms that the patch distribution is certainly a crucial information in texture modelling, certainly more important than the autocorrelation, and probably as important as the responses in a filter bank. Recall that the authors of [Varma & Zisserman 2003] and [Varma & Zisserman 2005] were led to a similar conclusion in texture classification. Despite the impressive synthesis results obtained

in [Portilla & Simoncelli 2000] (which is based on an overcomplete filter bank), we think that filter banks responses are just another way to represent the textural information, perhaps in a quite compact manner, but not very adapted to fast texture synthesis (because analyzing and synthesizing the subband correlations is a quite expensive operation).

We think that a promising avenue of research lies in the design of mathematical tools to handle patch distributions. Many questions can be addressed in this sense, for example:

1. How to compare two patch distributions? Can we build a distance on patch distributions that complies with texture perception?
2. Do we want to respect exactly the patch distribution of the exemplar?
3. Can the patch distribution be summarized only by a few representatives?

The tree-structured vector quantization of [Wei & Levoy 2000] is a first attempt to structure the patch space in a way that is relevant with respect to a particular exemplar texture. Structuring the patch space can also be addressed by using adaptive dictionary approaches [Elad & Aharon 2006]; such approaches have already come up on a successful synthesis algorithm in [Tartavel *et al.* 2014] which lies in the "texture optimization" framework. The algorithm of [Tartavel *et al.* 2014] is of particular interest because it constrains (among other things) the frequency of use of each atom of the dictionary. This dictionary constraint may certainly be translated into a constraint on the patch distribution.

The recent advances in the theory of optimal transportation may bring several clues to investigate the first question above. The second question is actually easier. Indeed, the patch distribution of a stationary random field is submitted to some simple constraints which have no reason to be satisfied by the empirical patch distribution of an exemplar (except for artificial tileable exemplar textures). Therefore, even if we want every part of the exemplar texture to be well represented in the synthesis, it is unreasonable to ask for exact matching of the patch distributions. An interesting question is thus: how can we modify the target patch distribution in order to allow for stationary synthesis? As concerns circular models, this question is equivalent to: how much are we prepared to modify the exemplar texture in order to make it tileable? It is clear that these questions are not well posed and may need additional information to be answered properly.

Once the target patch distribution is modified in order to agree for stationary synthesis, it remains to investigate how to randomly sample in the patch space. The ideal goal would be to design a random local function that is able to sample patches according to a prescribed distribution and conditioned by a low-resolution component. We would like such a local function to allow very fast computations, which seems possible at least in the case of Gaussian conditional simulation. However a non-trivial point would be to respect the compatibilities between adjacent patches in the synthesized texture. But, considering the successful results of bi-level synthesis obtained in Subsection 5.3.4 with quite simple local functions, one could hope for compatibilities to be partially inherited from low-resolution to high-resolution.

Bibliography

[Aguerrebere *et al.* 2013] C. Aguerrebere, Y. Gousseau and G. Tartavel. *Exemplar-based Texture Synthesis: the Efros-Leung Algorithm.* Image Processing On Line, vol. 2013, pp. 213–231, 2013.

[Ashikhmin 2001] M. Ashikhmin. *Synthesizing natural textures.* In Proceedings of the 2001 symposium on Interactive 3D graphics, pp. 217–226, 2001.

[Aujol *et al.* 2005] J.F. Aujol, G. Aubert, L. Blanc-Féraud and A. Chambolle. *Image decomposition into a bounded variation component and an oscillating component.* Journal of Mathematical Imaging and Vision, vol. 22, no. 1, pp. 71–88, 2005.

[Ayer *et al.* 1955] M. Ayer, H. D. Brunk, G. M. Ewing, W. T. Reid and E. Silverman. *An empirical distribution function for sampling with incomplete information.* The Annals of Mathematical Statistics, pp. 641–647, 1955.

[Barnes *et al.* 2009] C. Barnes, E. Shechtman, A. Finkelstein and D. Goldman. *PatchMatch: a randomized correspondence algorithm for structural image editing.* ACM Transactions on Graphics, vol. 28, no. 3, 2009.

[Bauschke *et al.* 2002] H.H. Bauschke, P.L. Combettes and D.R. Luke. *Phase retrieval, error reduction algorithm, and Fienup variants: a view from convex optimization.* Journal of the Optical Society of America A, vol. 19, no. 7, pp. 1334–1345, 2002.

[Benhamou *et al.* 1994] C.l. Benhamou, E. Lespessailles, G. Jacquet, R. Harba, R. Jennane, T. Loussot, D. Tourliere and W. Ohley. *Fractal organization of trabecular bone images on calcaneus radiographs.* Journal of Bone and Mineral Research, vol. 9, no. 12, pp. 1909–1918, 1994.

[Bergen & Adelson 1988] J.R. Bergen and E.H. Adelson. *Early vision and texture perception.* Nature, vol. 333, no. 6171, pp. 363–364, 1988.

[Bertalmio *et al.* 2000] M. Bertalmio, G. Sapiro, V. Caselles and C. Ballester. *Image Inpainting.* In Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '00, pp. 417–424, 2000.

[Bertalmio *et al.* 2003] M. Bertalmio, L. Vese, G. Sapiro and S. Osher. *Simultaneous structure and texture image inpainting.* IEEE Transactions on Image Processing, vol. 12, no. 8, pp. 882–889, 2003.

[Besag & Kooperberg 1995] J. Besag and C. Kooperberg. *On conditional and intrinsic autoregressions.* Biometrika, vol. 82, no. 4, pp. 733–746, 1995.

[Besag 1974] J. Besag. *Spatial interaction and the statistical analysis of lattice systems.* Journal of the Royal Statistical Society. Series B (Methodological), pp. 192–236, 1974.

[Bickel & Levina 2008] P.J. Bickel and E. Levina. *Covariance regularization by thresholding.* The Annals of Statistics, vol. 36, no. 6, pp. 2577–2604, 2008.

[Billingsley 2012] P. Billingsley. Probability and measure. Wiley-Blackwell, 3rd edition, 2012.

[Blanchet & Moisan 2012] G. Blanchet and L. Moisan. *An explicit sharpness index related to Global Phase Coherence.* In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 1065–1068, 2012.

[Blanchet *et al.* 2005] G. Blanchet, L. Moisan and B. Rougé. *A linear prefilter for image sampling with ringing artifact control.* In Proceedings of the IEEE International Conference on Image Processing, vol. 3, pp. 577–580, 2005.

[Blanchet *et al.* 2008] G. Blanchet, L. Moisan and B. Rougé. *Measuring the global phase coherence of an image.* In Proceedings of the IEEE International Conference on Image Processing, pp. 1176–1179, 2008.

[Bolte *et al.* 2013] J. Bolte, S. Sabach and M. Teboulle. *Proximal alternating linearized minimization for nonconvex and nonsmooth problems.* Mathematical Programming, 2013.

[Bordenave *et al.* 2006] C. Bordenave, Y. Gousseau and F. Roueff. *The dead leaves model: a general tessellation modeling occlusion.* Advances in Applied Probability, pp. 31–46, 2006.

[Borgnat *et al.* 2010] P. Borgnat, P. Flandrin, P. Honeine, C. Richard and J. Xiao. *Testing Stationarity With Surrogates: A Time-Frequency Approach.* IEEE Transactions on Signal Processing, vol. 58, no. 7, pp. 3459–3470, July 2010.

[Boussidi *et al.* 2014] B. Boussidi, R. Fablet, B. Chapron and E. Autret. *Champs Gaussiens conditionnels pour la modélisation inter-échelle de textures : Application à la super-résolution en télédétection satellitaire de l'océan.* In Actes de la conférence RFIA, 2014.

[Briand *et al.* 2014] T. Briand, J. Vacher, B. Galerne and J. Rabin. *The Heeger-Bergen Pyramid Based Texture Synthesis Algorithm.* Image Processing On Line, vol. 4, pp. 276–299, 2014.

[Brémaud 1993] P. Brémaud. Signaux aléatoires pour le traitement du signal et les communications. Les Cours de l'École Polytechnique. Sciences exactes. Ellipses, 1993.

[Brémaud 1999] P. Brémaud. Markov chains : Gibbs fields, Monte Carlo simulation and queues. Texts in applied mathematics. Springer, 1999.

[Buades *et al.* 2005] A. Buades, B. Coll and J.M. Morel. *A review of image denoising algorithms, with a new one.* Multiscale Modeling & Simulation, vol. 4, no. 2, pp. 490–530, 2005.

[Cabral 2010] E. N. Cabral. *Étude spectrale des processus stationnaires multidimensionnels et analyse en composantes principales dans le domaine des fréquences.* PhD thesis, Université Toulouse III-Paul Sabatier, 2010.

[Cadzow *et al.* 1993] J. A. Cadzow, D.M. Wilkes, R.A. Peters and X. Li. *Image texture synthesis-by-analysis using moving-average models.* IEEE Transactions on Aerospace and Electronic Systems, vol. 29, no. 4, pp. 1110–1122, 1993.

[Chainais *et al.* 2011] P. Chainais, E. Kœnig, V. Delouille and J.-F. Hochedez. *Virtual Super Resolution of Scale Invariant Textured Images Using Multifractal Stochastic Processes.* Journal of Mathematical Imaging and Vision, vol. 39, no. 1, pp. 28–44, 2011.

[Chambolle 2004] A. Chambolle. *An algorithm for total variation minimization and applications.* Journal of Mathematical imaging and vision, vol. 20, no. 1-2, pp. 89–97, 2004.

[Chandler & Hemami 2007] D.M. Chandler and S.S. Hemami. *VSNR: A Wavelet-Based Visual Signal-to-Noise Ratio for Natural Images.* IEEE Transactions on Image Processing, vol. 16, no. 9, pp. 2284–2298, 2007.

[Chandler 2013] D.M. Chandler. *Seven Challenges in Image Quality Assessment: Past, Present, and Future Research.* ISRN Signal Processing, vol. 2013, pp. 1–53, 2013.

[Chandran & Elgar 1994] V. Chandran and S. Elgar. *A General Procedure for the Derivation of Principal Domains of Higher-Order Spectra.* IEEE Transactions on Signal Processing, vol. 42, no. 1, pp. 229–233, 1994.

[Chellappa & Chatterjee 1985] R. Chellappa and S. Chatterjee. *Classification of textures using Gaussian Markov random fields.* IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 33, no. 4, pp. 959–963, August 1985.

[Chellappa & Kashyap 1985] R. Chellappa and R.L. Kashyap. *Texture synthesis using 2-D noncausal autoregressive models.* IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 33, no. 1, pp. 194–203, February 1985.

[Chellappa 1985] R. Chellappa. *Two-dimensional discrete Gaussian Markov random field models for image processing.* Progress in Pattern Recognition, vol. 2, pp. 79–112, 1985.

[Collis *et al.* 1998] W.B. Collis, P.R. White and J.K. Hammond. *Higher-order spectra: the bispectrum and trispectrum.* Mechanical Systems and Signal Processing, vol. 12, no. 3, pp. 375–394, 1998.

[Cook & DeRose 2005] R.L. Cook and T. DeRose. *Wavelet noise*. In ACM Transactions on Graphics, vol. 24, pp. 803–811. ACM, 2005.

[Costantini *et al.* 2004] R. Costantini, G. Menegaz and S. Susstrunk. *A measure for spatial dependence in natural stochastic textures*. In Proceedings of the International Conference on Image Processing, vol. 3, pp. 1525–1528, 2004.

[Cramer 1940] H. Cramer. *On the Theory of Stationary Random Processes*. The Annals of Mathematics, vol. 41, no. 1, pp. 215–230, 1940.

[Criminisi *et al.* 2004] A. Criminisi, P. Pérez and K. Toyama. *Region filling and object removal by exemplar-based image inpainting*. IEEE Transactions on Image Processing, vol. 13, no. 9, pp. 1200–1212, 2004.

[Cross & Jain 1983] G.R. Cross and A.K. Jain. *Markov Random Field Texture Models*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 5, no. 1, pp. 25–39, 1983.

[Daubechies 1988] I. Daubechies. *Orthonormal bases of compactly supported wavelets*. Communications on Pure and Applied Mathematics, vol. 41, no. 7, pp. 909–996, 1988.

[De Bonet 1997] J.S. De Bonet. *Multiresolution sampling procedure for analysis and synthesis of texture images*. pp. 361–368. ACM, 1997.

[Deriugin 1956] N.G. Deriugin. *The power spectrum and the correlation function of the television signal*. Telecommunications, vol. 1, no. 7, pp. 1–12, 1956.

[Desolneux *et al.* 2008] A. Desolneux, L. Moisan and J.-M. Morel. From Gestalt Theory to Image Analysis: A Probabilistic Approach, vol. 34 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, 2008.

[Desolneux *et al.* 2012] A. Desolneux, L. Moisan and S. Ronsin. *A compact representation of random phase and Gaussian textures*. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 1381–1384, 2012.

[Desolneux *et al.* 2015] A. Desolneux, L. Moisan and S. Ronsin. *A Texton for Random Phase and Gaussian Textures*. in preparation, 2015.

[Dong *et al.* 2008] Y. Dong, S. Lefebvre, X. Tong and G. Drettakis. *Lazy solid texture synthesis*. In Computer Graphics Forum, vol. 27, pp. 1165–1174, 2008.

[Doob 1944] J. L. Doob. *The Elementary Gaussian Processes*. The Annals of Mathematical Statistics, vol. 15, no. 3, pp. 229–282, 1944.

[Doob 1949] J.L. Doob. *Time series and harmonic analysis*. In Proceedings of the Berkeley Simposium on Mathematical Statistics and Probability. University of California Press, 1949.

[Doob 1990] Joseph L. Doob. Stochastic processes. Wiley, 1990.

[Dowson & Landau 1982] D. C Dowson and B. V Landau. *The Fréchet distance between multivariate normal distributions.* Journal of Multivariate Analysis, vol. 12, no. 3, pp. 450–455, 1982.

[Dritschel & Woerdeman 2005] M. Dritschel and H. Woerdeman. *Outer factorizations in one and several variables.* Transactions of the American Mathematical Society, vol. 357, no. 11, pp. 4661–4679, 2005.

[Dritschel 2004] M.A. Dritschel. *On factorization of trigonometric polynomials.* Integral Equations and Operator Theory, vol. 49, no. 1, pp. 11–42, 2004.

[Dumitrescu 2006] B. Dumitrescu. *Trigonometric polynomials positive on frequency domains and applications to 2-D FIR filter design.* IEEE Transactions on Signal Processing, vol. 54, no. 11, pp. 4282–4292, 2006.

[Efros & Freeman 2001] A.A. Efros and W.T. Freeman. *Image Quilting for Texture Synthesis and Transfer.* In Proceedings of SIGGRAPH, pp. 341–346, 2001.

[Efros & Leung 1999] A. A. Efros and T. K. Leung. *Texture synthesis by non-parametric sampling.* In Proceedings of the IEEE International Conference on Computer Vision, vol. 2, pp. 1033–1038, 1999.

[Ekeland & Témam 1999] I. Ekeland and R. Témam. Convex analysis and variational problems. Classics in Applied Mathematics. SIAM, 1999.

[Elad & Aharon 2006] M. Elad and M. Aharon. *Image denoising via sparse and redundant representations over learned dictionaries.* IEEE Transactions on Image Processing, vol. 15, no. 12, pp. 3736–3745, 2006.

[Eom 2000] K.B. Eom. *Synthesis of Color Textures for Multimedia Applications.* Multimedia Tools and Applications, vol. 12, no. 1, pp. 81–98, 2000.

[Ferzli & Karam 2009] R. Ferzli and L.J. Karam. *A No-Reference Objective Image Sharpness Metric Based on the Notion of Just Noticeable Blur (JNB).* IEEE Transactions on Image Processing, vol. 18, no. 4, pp. 717–728, 2009.

[Field *et al.* 2000] D. J. Field, A. Hayes and R. F. Hess. *The roles of polarity and symmetry in the perceptual grouping of contour fragments.* Spatial Vision, vol. 13, no. 1, pp. 51–66, 2000.

[Fienup 1982] J. R. Fienup. *Phase retrieval algorithms: a comparison.* Applied Optics, vol. 21, no. 15, pp. 2758–2769, 1982.

[Fournier *et al.* 1982] A. Fournier, D. Fussell and L. Carpenter. *Computer Rendering of Stochastic Models.* Commun. ACM, vol. 25, no. 6, pp. 371–384, 1982.

[Francos & Friedlander 1998] J.M. Francos and B. Friedlander. *Parameter estimation of two-dimensional moving average random fields.* IEEE Transactions on Signal Processing, vol. 46, no. 8, pp. 2157–2165, 1998.

[Francos *et al.* 1993] J. M. Francos, A. Z. Meiri and B. Porat. *A unified texture model based on a 2-D Wold-like decomposition.* IEEE Transactions on Signal Processing, vol. 41, no. 8, pp. 2665–2678, 1993.

[Frisen 1986] M. Frisen. *Unimodal Regression.* The Statistician, vol. 35, no. 4, pp. 479–385, 1986.

[Galerne & Gousseau 2012] B. Galerne and Y. Gousseau. *The transparent dead leaves model.* Advances in Applied Probability, vol. 44, no. 1, pp. 1–20, 2012.

[Galerne *et al.* 2011a] B. Galerne, Y. Gousseau and J.-M. Morel. *Micro-Texture Synthesis by Phase Randomization.* Image Processing On Line, vol. 1, 2011.

[Galerne *et al.* 2011b] B. Galerne, Y. Gousseau and J.-M. Morel. *Random Phase Textures: Theory and Synthesis.* IEEE Transactions on Image Processing, vol. 20, no. 1, pp. 257–267, 2011.

[Galerne *et al.* 2012] B. Galerne, A. Lagae, S. Lefebvre and G. Drettakis. *Gabor noise by example.* ACM Transactions on Graphics, vol. 31, no. 4, pp. 73:1–73:9, 2012.

[Galerne *et al.* 2014] B. Galerne, A. Leclaire and L. Moisan. *A Texton for Fast and Flexible Gaussian Texture Synthesis.* In Proceedings of the 22nd European Signal Processing Conference, pp. 1686–1690. IEEE, 2014.

[Galerne 2010] B. Galerne. *Modèles d'image aléatoires et Synthèse de texture.* PhD thesis, École Normale Supérieure de Cachan, 2010.

[Gegenfurtner *et al.* 2003] K.R. Gegenfurtner, D.I. Braun and F.A. Wichmann. *The importance of phase information for recognizing natural images.* Journal of Vision, vol. 3, no. 9, 2003.

[Gelfand & Vilenkin 1967] I.M. Gelfand and N.J. Vilenkin. Les distributions: Applications de l'analyse harmonique. Dunod, 1967.

[Geman & Graffigne 1986] S. Geman and C. Graffigne. *Markov random field image models and their applications to computer vision.* In Proceedings of the International Congress of Mathematicians, vol. 1, p. 2, 1986.

[Ghazanfarpour & Dischler 1995] D. Ghazanfarpour and J. M. Dischler. *Spectral analysis for automatic 3-D texture generation.* Computers & Graphics, vol. 19, no. 3, pp. 413–422, May 1995.

[Ghiglia & Pritt 1998] D.C. Ghiglia and M.D. Pritt. Two-dimensional phase unwrapping: Theory, algorithms, and software. Wiley, 1998.

[Gilet *et al.* 2014] G. Gilet, B. Sauvage, K. Vanhoey, J.M. Dischler and D. Ghazanfarpour. *Local Random-phase Noise for Procedural Texturing.* ACM Transactions on Graphics, vol. 33, no. 6, pp. 195:1–195:11, 2014.

[Goldberg *et al.* 2008] A. Goldberg, M. Zwicker and F. Durand. *Anisotropic noise.* In ACM Transactions on Graphics, vol. 27, p. 54. ACM, 2008.

[Gousseau 2002] Y. Gousseau. *Texture synthesis through level sets.* In 2nd International Workshop on Texture Analysis and Synthesis, pp. 53–57, 2002.

[Gray *et al.* 1975] R.M. Gray, D.L. Neuhoff and P.C. Shields. *A Generalization of Ornstein's $\bar{d}$ Distance with Applications to Information Theory.* The Annals of Probability, vol. 3, no. 2, pp. 315–328, 1975.

[Grenander & Rosenblatt 1953] U. Grenander and M. Rosenblatt. *Statistical Spectral Analysis of Time Series Arising from Stationary Stochastic Processes.* The Annals of Mathematical Statistics, vol. 24, no. 4, pp. 537–558, 1953.

[Grosjean & Moisan 2009] B. Grosjean and L. Moisan. *A-contrario detectability of spots in textured backgrounds.* Journal of Mathematical Imaging and Vision, vol. 33, no. 3, pp. 313–337, 2009.

[Hall & Giannakis 1995] T. E. Hall and G. B. Giannakis. *Bispectral analysis and model validation of texture images.* IEEE Transactions on Image Processing, vol. 4, no. 7, pp. 996–1009, 1995.

[Han *et al.* 2006] J. Han, K. Zhou, L.-Y. Wei, M. Gong, H. Bao, X. Zhang and B. Guo. *Fast example-based surface texture synthesis via discrete optimization.* The Visual Computer, vol. 22, no. 9-11, pp. 918–925, 2006.

[Han *et al.* 2008] C. Han, E. Risser, R. Ramamoorthi and E. Grinspun. *Multiscale texture synthesis.* vol. 27, 2008.

[Hansen & Hess 2006] B. C. Hansen and R. F. Hess. *The role of spatial phase in texture segmentation and contour integration.* Journal of Vision, vol. 6, no. 5, 2006.

[Haralick 1979] R.M. Haralick. *Statistical and structural approaches to texture.* Proceedings of the IEEE, vol. 67, no. 5, pp. 786–804, 1979.

[Hartman & Hössjer 2008] L. Hartman and O. Hössjer. *Fast kriging of large data sets with Gaussian Markov random fields.* Computational Statistics and Data Analysis, no. 5, pp. 2331–2349, 2008.

[Hassen *et al.* 2010] R. Hassen, Z. Wang and M. Salama. *No-reference image sharpness assessment based on local phase coherence measurement.* In Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing, pp. 2434–2437, 2010.

[Hayes 1982] M.H. Hayes. *The Reconstruction of a Multidimensional Sequence from the Phase or Magnitude of its Fourier Transform.* IEEE Transactions on acoustics, speech, and signal processing, vol. 30, no. 2, pp. 140–154, 1982.

[Heeger & Bergen 1995] D.J. Heeger and J.R. Bergen. *Pyramid-based texture analysis/synthesis.* In Proceedings of the 22nd annual conference on Computer graphics and interactive techniques, pp. 229–238, 1995.

[Helson & Lowdenslager 1961] H. Helson and D. Lowdenslager. *Prediction theory and Fourier series in several variables. II.* Acta Mathematica, vol. 106, no. 3, pp. 175–213, 1961.

[Hertzmann *et al.* 2001] A. Hertzmann, C.E. Jacobs, N. Oliver, B. Curless and D.H. Salesin. *Image analogies.* In Proceedings of the conference on Computer graphics and interactive techniques, pp. 327–340, 2001.

[Huesmann & Sturm 2013] Martin Huesmann and Karl-Theodor Sturm. *Optimal transport from Lebesgue to Poisson.* The Annals of Probability, vol. 41, no. 4, pp. 2426–2478, 2013.

[ITU 2011] ITU. *BT 601: Studio encoding parameters of digital television for standard 4: 3 and wide-screen 16: 9 aspect ratios.* International Telecommunication Union, no. BT601-7, 2011.

[Jaming & Kolountzakis 2003] P. Jaming and M. N. Kolountzakis. *Reconstruction of functions from their triple correlations.* New York Journal of Mathematics, vol. 9, pp. 149–164, 2003.

[Janson 1988] S. Janson. *Normal convergence by higher semiinvariants with applications to sums of dependent random variables and random graphs.* The Annals of Probability, vol. 16, no. 1, pp. 305–312, 1988.

[Jassim 2013] F.A. Jassim. *Image Inpainting by Kriging Interpolation Technique.* World of Computer Science and Information Technology Journal, vol. 3, no. 5, pp. 91–96, 2013.

[Julesz 1962] B. Julesz. *Visual Pattern Discrimination.* IRE Transactions on Information Theory, vol. 8, no. 2, pp. 84–92, 1962.

[Julesz 1981] B. Julesz. *Textons, the elements of texture perception, and their interactions.* Nature, vol. 290, no. 5802, pp. 91–97, 1981.

[Kim *et al.* 2012] V. G. Kim, Y. Lipman and T. Funkhouser. *Symmetry-guided texture synthesis and manipulation.* ACM Transactions on Graphics, vol. 31, no. 3, p. 22, 2012.

[Kovesi 1999] P. Kovesi. *Image features from phase congruency.* Technical report, The University of Western Australia, 1999.

[Kovesi 2000] P. Kovesi. *Phase congruency: A low-level image invariant.* Psychological research, vol. 64, no. 2, pp. 136–148, 2000.

[Kwatra *et al.* 2003] V. Kwatra, A. Schödl, I. Essa, G. Turk and A. Bobick. *Graphcut textures: image and video synthesis using graph cuts.* In ACM Transactions on Graphics, vol. 22, p. 277?286, 2003.

[Kwatra *et al.* 2005] V. Kwatra, I. Essa, A. Bobick and N. Kwatra. *Texture optimization for example-based synthesis.* In ACM Transactions on Graphics, vol. 24, pp. 795–802, 2005.

[Lagae *et al.* 2009] A. Lagae, S. Lefebvre, G. Drettakis and P. Dutré. *Procedural Noise Using Sparse Gabor Convolution.* In Proceedings of SIGGRAPH, vol. 28, pp. 54:1–54:10, 2009.

[Lagae *et al.* 2010a] A. Lagae, S. Lefebvre, R. Cook, T. DeRose, G. Drettakis, D.S. Ebert, J.P. Lewis, K. Perlin and M. Zwicker. *A survey of procedural noise functions.* In Computer Graphics Forum, vol. 29, pp. 2579–2600, 2010.

[Lagae *et al.* 2010b] A. Lagae, P. Vangorp, T. Lenaerts and P. Dutré. *Procedural isotropic stochastic textures by example.* Computers & Graphics, vol. 34, no. 4, pp. 312–321, 2010.

[Lantuéjoul 2002] C. Lantuéjoul. Geostatistical simulation: Models and algorithms. Springer, 2002.

[Leclaire & Moisan 2013a] A. Leclaire and L. Moisan. *Blind Deblurring Using a Simplified Sharpness Index.* In Proceedings of Scale Space and Variational Methods in Computer Vision, Lecture Notes in Computer Science, pp. 86–97. Springer Berlin Heidelberg, 2013.

[Leclaire & Moisan 2013b] A. Leclaire and L. Moisan. *Une Variante non Périodique du Sharpness Index.* In Actes du GRETSI, 2013.

[Lefebvre & Hoppe 2005] S. Lefebvre and H. Hoppe. *Parallel controllable texture synthesis.* ACM Transactions on Graphics, vol. 24, no. 3, pp. 777–786, 2005.

[Levi & Stark 1984] A. Levi and H. Stark. *Image restoration by the method of generalized projections with application to restoration from magnitude.* Journal of the Optical Society of America A, vol. 1, no. 9, pp. 932–943, 1984.

[Levin *et al.* 2011] A. Levin, Y. Weiss, F. Durand and W. T. Freeman. *Efficient marginal likelihood optimization in blind deconvolution.* In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2657–2664, 2011.

[Levina & Bickel 2006] E. Levina and P.J. Bickel. *Texture synthesis and nonparametric resampling of random fields.* The Annals of Statistics, vol. 34, no. 4, pp. 1751–1773, 2006.

[Lewis 1984] J.P. Lewis. *Texture Synthesis for Digital Painting.* In Proceedings of the Conference on Computer Graphics and Interactive Techniques, SIGGRAPH, pp. 245–252. ACM, 1984.

[Lewis 1987] J.P. Lewis. *Generalized stochastic subdivision.* ACM Transactions on Graphics, vol. 6, no. 3, pp. 167–190, 1987.

[Lewis 1989] J.P. Lewis. *Algorithms for Solid Noise Synthesis.* In Proceedings of the Conference on Computer Graphics and Interactive Techniques, SIGGRAPH, pp. 263–270. ACM, 1989.

[Li 2009] S.Z. Li. Markov random field modeling in image analysis, vol. 26. Springer, 2009.

[Liang *et al.* 2001] L. Liang, C. Liu, Y.Q. Xu, B. Guo and H.Y. Shum. *Real-time texture synthesis by patch-based sampling.* ACM Transactions on Graphics, vol. 20, no. 3, pp. 127–150, 2001.

[Liu *et al.* 2013] Y. Liu, J. Wang, S. Cho, A. Finkelstein and S. Rusinkiewicz. *A No-Reference Metric for Evaluating the Quality of Motion Deblurring.* ACM Transactions on Graphics, vol. 32, no. 6, 2013.

[Louchet & Moisan 2014] C. Louchet and L. Moisan. *Total Variation denoising using iterated conditional expectation.* In Proceedings of the 22nd European Signal Processing Conference, pp. 1592–1596, 2014.

[Mairal *et al.* 2008] J. Mairal, M. Elad and G. Sapiro. *Sparse Representation for Color Image Restoration.* IEEE Transactions on Image Processing, vol. 17, no. 1, pp. 53–69, 2008.

[Malik & Perona 1990] J. Malik and P. Perona. *Preattentive texture discrimination with early vision mechanisms.* Journal of the Optical Society of America A, vol. 7, no. 5, pp. 923–932, 1990.

[Malik *et al.* 1999] J. Malik, S. Belongie, J. Shi and T. Leung. *Textons, contours and regions: Cue integration in image segmentation.* In Proceedings of the International Conference on Computer Vision, 1999.

[Mandelbrot & Van Ness 1968] B.B. Mandelbrot and J.W. Van Ness. *Fractional Brownian motions, fractional noises and applications.* SIAM review, vol. 10, no. 4, pp. 422–437, 1968.

[Marziliano *et al.* 2004] P. Marziliano, F. Dufaux, S. Winkler and T. Ebrahimi. *Perceptual blur and ringing metrics: Application to JPEG2000.* Signal Processing: Image Communication, pp. 163–172, 2004.

[Masnou & Morel 1998] S. Masnou and J.-M. Morel. *Level lines based disocclusion.* In Proceedings of the International Conference on Image Processing., vol. 3, pp. 259–263, 1998.

[Matheron 1968] G. Matheron. *Modèle séquentiel de partition aléatoire*. Technical report, 1968.

[Matsubara 2007] T. Matsubara. *Statistics of Fourier Modes in Non-Gaussian Fields*. The Astrophysical Journal Supplement Series, vol. 170, p. 1, 2007.

[Mead 1971] R. Mead. *Models for interplant competition in irregularly distributed populations.* In Proceedings of the International Symposium on Statistical Ecology, 1971.

[Meyer 2001] Y. Meyer. Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures. AMS, 2001.

[Moisan 2007] L. Moisan. *How to discretize the Total Variation of an image?* In Proceedings of ICIAM, vol. 7, pp. 1041907–1041908, 2007.

[Moisan 2011] L. Moisan. *Periodic plus smooth image decomposition*. Journal of Mathematical Imaging and Vision, vol. 39, no. 2, pp. 161–179, 2011.

[Moreno & Calderero 2013] P. Moreno and F. Calderero. *Evaluation of sharpness measures and proposal of a stop criterion for reverse diffusion in the context of image deblurring*. In Proceedings of the International Conference on Computer Vision Theory and Applications, 2013.

[Morrone & Burr 1988] M.C. Morrone and D.C. Burr. *Feature detection in human vision: a phase-dependent energy model*. Proceedings of the Royal Society of London, pp. 221–245, 1988.

[Myaskouvskey *et al.* 2013] A. Myaskouvskey, Y. Gousseau and M. Lindenbaum. *Beyond Independence: An Extension of the A Contrario Decision Procedure*. International Journal of Computer Vision, vol. 101, no. 1, pp. 22–44, 2013.

[Newson *et al.* 2014] A. Newson, A. Almansa, M. Fradet, Y. Gousseau and P. Pérez. *Video Inpainting of Complex Scenes*. Journal of Mathematical Imaging and Vision, 2014.

[Newson 2014] A. Newson. *Sur la restauration et l'édition de vidéo : détection de rayures et inpainting de scènes complexes*. PhD thesis, Telecom ParisTech, 2014.

[Nikias & Mendel 1993] C.L. Nikias and J.M. Mendel. *Signal processing with higher-order spectra*. IEEE Signal Processing Magazine, vol. 10, no. 3, pp. 10–37, 1993.

[Oppenheim & Lim 1981] A.V. Oppenheim and J.S. Lim. *The importance of phase in signals*. Proceedings of the IEEE, vol. 69, no. 5, pp. 529–541, 1981.

[Ord 1975] K. Ord. *Estimation Methods for Models of Spatial Interaction*. Journal of the American Statistical Association, vol. 70, no. 349, pp. 120–126, 1975.

[Paget & Longstaff 1998] R. Paget and I.D. Longstaff. *Texture synthesis via a non-causal nonparametric multiscale Markov random field.* IEEE Transactions on Image Processing, vol. 7, no. 6, pp. 925–931, 1998.

[Papoulis 1971] A. Papoulis. *High Density Shot Noise and Gaussianity.* Journal of Applied Probability, vol. 8, no. 1, pp. 118–127, 1971.

[Perlin 1985] K. Perlin. *An image synthesizer.* ACM SIGGRAPH Computer Graphics, vol. 19, no. 3, pp. 287–296, 1985.

[Peters & Itti 2008] R. Peters and L. Itti. *The role of Fourier phase information in predicting saliency.* Journal of Vision, vol. 8, no. 6, pp. 879–879, 2008.

[Peyré 2009] G. Peyré. *Sparse modeling of textures.* Journal of Mathematical Imaging and Vision, vol. 34, no. 1, pp. 17–31, 2009.

[Picinbono 1998] B. Picinbono. *Géométrie des polyspectres.* Traitement du signal, vol. 15, no. 6, pp. 447–456, 1998.

[Popat & Picard 1993] K. Popat and R.W. Picard. *Novel Cluster-Based Probability Model for Texture Synthesis, Classification, and Compression.* In Visual Communications and Image Processing, pp. 756–768, 1993.

[Portilla & Simoncelli 2000] J. Portilla and E.P. Simoncelli. *A parametric texture model based on joint statistics of complex wavelet coefficients.* International Journal of Computer Vision, vol. 40, no. 1, pp. 49–70, 2000.

[Prum 1997] B. Prum. Processus sur un réseau et mesures de Gibbs. Dunod, 1997.

[Raad *et al.* 2014] L. Raad, A. Desolneux and J.-M. Morel. *Locally Gaussian exemplar-based texture synthesis.* In Proceedings of the IEEE International Conference on Image Processing, 2014.

[Raad *et al.* 2015] L. Raad, A. Desolneux and J.-M. Morel. *Conditional Gaussian Models for Texture Synthesis.* In Proceedings of Scale Space and Variational Methods in Computer Vision, 2015.

[Ramanarayanan & Bala 2007] G. Ramanarayanan and K. Bala. *Constrained Texture Synthesis via Energy Minimization.* IEEE Transactions on Visualization and Computer Graphics, vol. 13, no. 1, pp. 167–178, 2007.

[Rice 1977] J. Rice. *On generalized shot noise.* Advances in Applied Probability, pp. 553–565, 1977.

[Ronsin *et al.* 2013] S. Ronsin, H. Biermé and L. Moisan. *The Billard Theorem for Multiple Random Fourier Series.* HAL preprint:hal-00850112, 2013.

[Rosenblum & Rovnyak 1997] M. Rosenblum and J. Rovnyak. Hardy classes and operator theory. Courier Dover Publications, 1997.

[Rozanov 1967] Y. A. Rozanov. Stationary random processes. Holden-Day, 1967.

[Ruderman 1994] D. L Ruderman. *The statistics of natural images.* Network: Computation in Neural Systems, vol. 5, no. 4, pp. 517–548, 1994.

[Rudin *et al.* 1992] L.I. Rudin, S. Osher and E. Fatemi. *Nonlinear Total Variation Based Noise Removal Algorithms.* Physica D, vol. 60, pp. 259–268, 1992.

[Sampat & Bovik 2003] M.P. Sampat and A.C. Bovik. *Detection of spiculated lesions in mammograms.* In Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, vol. 1, pp. 810–813, 2003.

[Schreiber & Schmitz 2000] T. Schreiber and A. Schmitz. *Surrogate time series.* Physica D: Nonlinear Phenomena, vol. 142, no. 3, pp. 346–382, 2000.

[Shevtsova 2010] I. G. Shevtsova. *An improvement of convergence rate estimates in the Lyapunov theorem.* In Doklady Mathematics, vol. 82, pp. 862–864. Springer, 2010.

[Simpson & Weiner 1989] J.A. Simpson and E.S.C. Weiner. The Oxford English dictionary. Clarendon press, 2nd edition, 1989.

[Stout 2008] Q.F. Stout. *Unimodal Regression via Prefix Isotonic Regression.* Computational Statistics and Data Analysis, vol. 53, pp. 289–297, 2008.

[Takeda *et al.* 2007] H. Takeda, S. Farsiu and P. Milanfar. *Kernel Regression for Image Processing and Reconstruction.* IEEE Transactions on Image Processing, vol. 16, no. 2, pp. 349–366, 2007.

[Tartavel *et al.* 2014] G. Tartavel, Y. Gousseau and G. Peyré. *Variational Texture Synthesis with Sparsity and Spectrum Constraints.* Journal of Mathematical Imaging and Vision, 2014.

[Theiler *et al.* 1992] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian and J.D. Farmer. *Testing for nonlinearity in time series: the method of surrogate data.* Physica D: Nonlinear Phenomena, vol. 58, no. 1, pp. 77–94, 1992.

[Van Wijk 1991] J. J. Van Wijk. *Spot noise texture synthesis for data visualization.* In ACM SIGGRAPH Computer Graphics, vol. 25, pp. 309–318, 1991.

[Varma & Zisserman 2003] M. Varma and A. Zisserman. *Texture classification: are filter banks necessary?* In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, 2003.

[Varma & Zisserman 2005] M. Varma and A. Zisserman. *A statistical approach to texture classification from single images.* International Journal of Computer Vision, vol. 62, no. 1-2, pp. 61–81, 2005.

[Villani 2003] C. Villani. Topics in Optimal Transportation. American Mathematical Society, 2003.

[Vu & Chandler 2009] C.T. Vu and D.M. Chandler. *S3: A spectral and spatial sharpness measure*. In Proceedings of the First International Conference on Advances in Multimedia, pp. 37–43, 2009.

[Wang & Bovik 2006] Z. Wang and A.C. Bovik. Modern image quality assessment (synthesis lectures on image, video, and multimedia processing). Morgan & Claypool Publishers, 2006.

[Wang & Simoncelli 2004] Z. Wang and E. P. Simoncelli. *Local phase coherence and the perception of blur*. In Advances in Neural Information Processing Systems, vol. 16, pp. 786–792, 2004.

[Wang *et al.* 2004] Z. Wang, A. C Bovik, H. R Sheikh and E. P Simoncelli. *Image quality assessment: From error visibility to structural similarity*. IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600–612, 2004.

[Wei & Levoy 2000] L.Y. Wei and M. Levoy. *Fast texture synthesis using tree-structured vector quantization*. In Proceedings of the 27th annual conference on Computer graphics and interactive techniques, pp. 479–488, 2000.

[Wei & Levoy 2002] L.-Y. Wei and M. Levoy. *Order-independent texture synthesis*. Technical report TR-2002-01, Computer Science Department, Stanford University, 2002.

[Wei *et al.* 2008] L.-Y. Wei, J. Han, K. Zhou, H. Bao, B. Guo and H.-Y. Shum. *Inverse texture synthesis*. In ACM Transactions on Graphics, vol. 27, 2008.

[Wei *et al.* 2009] L.-Y. Wei, S. Lefebvre, V. Kwatra and G. Turk. *State of the art in example-based texture synthesis*. In Eurographics 2009, State of the Art Report, EG-STAR, pp. 93–117, 2009.

[Wexler *et al.* 2004] Y. Wexler, E. Shechtman and M. Irani. *Space-time video completion*. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 120–127, 2004.

[Whittle 1954] P. Whittle. *On Stationary Processes in the Plane*. Biometrika, vol. 41, no. 3/4, pp. 434–449, 1954.

[Winkler 2006] Gerhard Winkler. Image analysis, random fields and markov chain monte carlo methods: A mathematical introduction. Springer, 2006.

[Wold 1938] H. Wold. *A study in the analysis of stationary time series*. PhD thesis, Stockholm, 1938.

[Wood & Chan 1994] A. T. A. Wood and G. Chan. *Simulation of Stationary Gaussian Processes in* $[0, 1]$. Journal of Computational and Graphical Statistics, vol. 3, no. 4, pp. 409–432, 1994.

[Woods 1972] J.W. Woods. *Two-dimensional discrete Markovian fields*. IEEE Transactions on Information Theory, vol. 18, no. 2, pp. 232–240, 1972.

[Xia *et al.* 2014] G. Xia, S. Ferradans, G. Peyré and J. Aujol. *Synthesizing and Mixing Stationary Gaussian Texture Models.* SIAM Journal on Imaging Sciences, vol. 7, no. 1, pp. 476–508, 2014.

[Xiao & Wu 2012] H. Xiao and W.B. Wu. *Covariance matrix estimation for stationary time series.* The Annals of Statistics, vol. 40, no. 1, pp. 466–493, 2012.

[Yaglom 1987] A. M. Yaglom. Correlation theory of stationary and related random functions: Basic results. Springer, 1987.

[Yule 1921] G.U. Yule. *On the Time-Correlation Problem, with Especial Reference to the Variate- Difference Correlation Method.* Journal of the Royal Statistical Society, vol. 84, no. 4, 1921.

[Zhu & Milanfar 2010] X. Zhu and P. Milanfar. *Automatic Parameter Selection for Denoising Algorithms Using a No-Reference Measure of Image Content.* IEEE Transactions on Image Processing, vol. 19, no. 12, pp. 3116–3132, 2010.

[Zhu *et al.* 1998] S.C. Zhu, Y. Wu and D. Mumford. *Filters, random fields and maximum entropy (FRAME): Towards a unified theory for texture modeling.* International Journal of Computer Vision, vol. 27, no. 2, pp. 107–126, 1998.

[Zhu *et al.* 2005] S.-C. Zhu, C.-E. Guo, Y. Wang and Z. Xu. *What are Textons?* International Journal of Computer Vision, vol. 62, no. 1-2, pp. 121–143, 2005.

Résumé

Dans cette thèse, on étudie la structuration des phases de la transformée de Fourier d'images naturelles, ce qui, du point de vue applicatif, débouche sur plusieurs mesures de netteté ainsi que sur des algorithmes rapides pour la synthèse de texture par l'exemple.

Le Chapitre 2 présente dans un cadre unifié plusieurs modèles de champs aléatoires, notamment les champs *spot noise* et champs gaussiens, en prêtant une attention particulière aux représentations fréquentielles de ces champs aléatoires.

Le Chapitre 3 détaille l'utilisation des champs à phase aléatoire à la synthèse de textures peu structurées (microtextures). On montre qu'une microtexture peut être résumée en une image de petite taille s'intégrant à un algorithme de synthèse très rapide et flexible via le modèle *spot noise*. Aussi on propose un algorithme de désocclusion de zones texturales uniformes basé sur la simulation gaussienne conditionnelle.

Le Chapitre 4 présente trois mesures de cohérence globale des phases de la transformée de Fourier. Après une étude théorique et pratique établissant leur lien avec la netteté d'image, on propose un algorithme de déflouage aveugle basé sur l'optimisation stochastique de ces indices.

Enfin, dans le Chapitre 5, après une discussion sur l'analyse et la synthèse directe de l'information de phase, on propose deux modèles de textures à phases cohérentes qui permettent la synthèse de textures plus structurées tout en conservant quelques garanties mathématiques simples.

Abstract

This thesis deals with the Fourier phase structure of natural images, and addresses no-reference sharpness assessment and fast texture synthesis by example.

In Chapter 2, we present several models of random fields in a unified framework, like the spot noise model and the Gaussian model, with particular attention to the spectral representation of these random fields.

In Chapter 3, random phase models are used to perform by-example synthesis of microtextures (textures with no salient features). We show that a microtexture can be summarized by a small image that can be used for fast and flexible synthesis based on the spot noise model. Besides, we address microtexture inpainting through the use of Gaussian conditional simulation.

In Chapter 4, we present three measures of the global Fourier phase coherence. Their link with the image sharpness is established based on a theoretical and practical study. We then derive a stochastic optimization scheme for these indices, which leads to a blind deblurring algorithm.

Finally, in Chapter 5, after discussing the possibility of direct phase analysis or synthesis, we propose two non random phase texture models which allow for synthesis of more structured textures and still have simple mathematical guarantees.