

Semi-discrete Optimal Transport in Patch Space for Enriching Gaussian Textures

Bruno Galerne¹, Arthur Leclaire^{2(✉)}, and Julien Rabin³

¹ Laboratoire MAP5, Université Paris Descartes and CNRS, Sorbonne Paris Cité, Paris, France

² CMLA, ENS Cachan, CNRS, Université Paris-Saclay, 94235 Cachan, France
Arthur.Leclaire@cmla.ens-cachan.fr

³ Normandie Univ, ENSICAEN, CNRS, GREYC, 14000 Caen, France

Abstract. A bilevel texture model is proposed, based on a local transform of a Gaussian random field. The core of this method relies on the optimal transport of a continuous Gaussian distribution towards the discrete exemplar patch distribution. The synthesis then simply consists in a fast post-processing of a Gaussian texture sample, boiling down to an improved nearest-neighbor patch matching, while offering theoretical guarantees on statistical compliancy.

Keywords: Optimal transport · Texture synthesis · Patch distribution

1 Introduction

Designing models for realistic and fast rendering of structured textures is a challenging research topic. In the past, many models have been proposed for exemplar-based synthesis, which consists in synthesizing a piece of texture having the same perceptual characteristics as an observed texture sample with still some innovative content. Several authors [2, 24] have proposed non-parametric methods based on progressive sampling of the texture using a copy-paste principle. This paved the way to many other successful synthesis methods relying on patch-based sampling [3, 11–13, 16, 19] (see [23] for a detailed review). Even if these methods can be applied very efficiently [13], they do not offer much mathematical guarantees (except asymptotic results [14]) which reflects for example in the growing garbage effect described in [2].

In contrast Gaussian texture models [4] are stationary and inherently respect the frequency content of the input texture. They can be efficiently simulated on the discrete plane \mathbb{Z}^2 [6] and even generalized to the framework of procedural noises defined in the continuous domain \mathbb{R}^2 (see e.g. [7] and references therein). They also allow for dynamic texture synthesis and mixing [25] and inpainting [5]. However the Gaussian model is intrinsically limited since the color distribution of the output is always symmetric around the mean color and the local texture patterns cannot contain salient features such as contrasted contours.

The main purpose of this work is to propose a theoretically sound post-processing of the Gaussian model to cope with some of these limitations. Since the color and local pattern information is a part of the patch distribution, we propose to apply a local operation that will transform the patch distribution of the Gaussian texture into the patch distribution of the input image. This can naturally be addressed using a semi-discrete optimal transport plan.

Several tools from optimal transport (OT) have already been applied to texture synthesis: Rabin et al. [20] and Xia et al. [25] formulate texture mixing via Wasserstein barycenters, Tartavel et al. [21] use Wasserstein distances in a variational formulation of texture synthesis, and Gutierrez et al. [9] apply discrete OT in order to specify a global patch distribution. Here, we suggest to locally apply a semi-discrete transport plan which can be seen as a tweaked nearest-neighbor projection.

Indeed, semi-discrete OT corresponds to assign the centers of Laguerre cells to each point of a well-designed Laguerre partition [1, 17]. So far, deterministic methods for solving semi-discrete methods have been limited to dimensions $D = 2$ [17] and $D = 3$ [15] using explicit geometric construction of the Laguerre tessellation. Recently, Genevay et al. [8] proposed several stochastic optimization schemes for solving entropy-regularized OT problems.

In Sect. 2 we summarize the non-regularized semi-discrete OT framework and the numerical solution given in [8]. In Sect. 3 we show how to use this algorithm to transport patch distributions (in dimensions $D = 27$ or higher) thus enriching the Gaussian texture model. The resulting bilevel algorithm, can be seen as a non-iterative version of [11] with a global statistical constraint, which is confirmed by the experiments of Sect. 4.

2 Semi-discrete Optimal Transport

In this section, we recall the framework for semi-discrete optimal transport and the numerical solution given by Genevay et al. [8].

2.1 The Optimal Transport Problem and Its Dual Formulation

Let μ, ν be two probability measures on \mathbb{R}^D . We assume that μ has a bounded probability density function ρ and that ν is a discrete measure $\nu = \sum_{y \in S} \nu_y \delta_y$ with finite support S . Let us denote by $\Pi(\mu, \nu)$ the set of probability measures on $\mathbb{R}^D \times \mathbb{R}^D$ having marginal distributions μ, ν . If T is a measurable map, we denote by $T_{\#}\mu$ the push-forward measure defined as $T_{\#}\mu(A) = \mu(T^{-1}(A))$. If $v \in \mathbb{R}^S$, we define the c -transform of v with respect to the cost $c(x, y) = \|x - y\|^2$ as $v^c(x) = \min_{y \in S} \|x - y\|^2 - v(y)$, and T_v the corresponding assignment uniquely defined almost everywhere by

$$T_v(x) = \operatorname{argmin}_{y \in S} \|x - y\|^2 - v(y). \quad (1)$$

When $v = 0$, we get the nearest neighbor (NN) projection which assigns to x the closest point in S (unique for almost all x).

The preimages of T_v define a partition of \mathbb{R}^D (up to a negligible set), called the power diagram or Laguerre tessellation, the cells of which are defined as

$$\text{Pow}_v(y) = \{x \in \mathbb{R}^D \mid \forall z \in S \setminus \{y\}, \|x - y\|^2 - v(y) < \|x - z\|^2 - v(z)\}. \quad (2)$$

The Kantorovich formulation of optimal transport consists in solving

$$\min_{\pi \in \Pi(\mu, \nu)} \int_{\mathbb{R}^D \times \mathbb{R}^D} \|x - y\|^2 d\pi(x, y). \quad (3)$$

It has been shown [1, 10, 15, 22] that this problem admits solutions of the form $(\text{Id} \times T_v)_\# \mu$ where v solves the dual problem

$$\operatorname{argmax}_{v \in \mathbb{R}^S} H(v) \quad \text{where} \quad H(v) = \int_{\mathbb{R}^D} v^c(x) \rho(x) dx + \sum_{y \in S} v(y) \nu_y. \quad (4)$$

The same authors have shown that the function H is concave, \mathcal{C}^1 -smooth, and that its gradient is given by $\frac{\partial H}{\partial v(y)} = -\mu(\text{Pow}_v(y)) + \nu_y$. Thus, v is a critical point of H if and only if $\mu(\text{Pow}_v(y)) = \nu_y$ for all y , which means that $(T_v)_\# \mu = \nu$.

2.2 Stochastic Optimization

Genevay et al. [8] have suggested to address the maximization of (4) by using a stochastic gradient ascent, which is made possible by writing

$$H(v) = \mathbb{E}[h(X, v)] \quad \text{where} \quad h(x, v) = v^c(x) + \sum_{y \in S} v(y) \nu_y \quad (5)$$

and where X is a random variable with distribution μ . Notice that for $x \in \text{Pow}_v(y)$, $v \mapsto v^c(x)$ is smooth with gradient $-e_y$ (where $\{e_y\}_{y \in S}$ is the canonical basis of \mathbb{R}^S). Therefore, for any $w \in \mathbb{R}^S$, for almost all $x \in \mathbb{R}^D$, $v \mapsto h(x, v)$ is differentiable at w and $\nabla_v h(x, w) = -e_{T_w(x)} + \nu$. (abusing notation $(\nu_y) \in \mathbb{R}^S$).

In order to minimize $-H$, Genevay et al. propose the following averaged stochastic gradient descent (ASGD) initialized with $\tilde{v}_1 = 0$

$$\begin{cases} \tilde{v}_k = \tilde{v}_{k-1} + \frac{C}{\sqrt{k}} \nabla_v h(x_k, \tilde{v}_{k-1}) & \text{where } x_k \sim \mu \\ v_k = \frac{1}{k} (\tilde{v}_1 + \dots + \tilde{v}_k). \end{cases} \quad (6)$$

Since $\nabla_v h(x, \tilde{v}_{k-1})$ exists x -a.s. and is bounded, the convergence of this algorithm is ensured by [18, Theorem 7], in the sense $\max(H) - \mathbb{E}[H(v_k)] = \mathcal{O}(\frac{\log k}{\sqrt{k}})$. The authors of [8] also proposed to address a regularized transport problem with a similar method but we do not discuss it here due to lack of space.

3 A Bilevel Model for Texture Synthesis

In this section, we introduce a bilevel model for texture synthesis that consists in first synthesizing a Gaussian version of the texture (with long range correlations but no geometric structures) and next transforming each patch of the synthesis with an optimal assignment (in order to enforce the patch distribution of the exemplar texture).

To be more precise, let us denote by $u : \Omega \rightarrow \mathbb{R}^d$ the exemplar texture defined on a discrete domain $\Omega \subset \mathbb{Z}^2$. Let $U : \mathbb{Z}^2 \rightarrow \mathbb{R}^d$ be the asymptotic discrete spot noise (ADSN) [4, 6] associated with u , which is defined as

$$\forall x \in \mathbb{Z}^2, U(x) = \bar{u} + \sum_{y \in \mathbb{Z}^2} t_u(y)W(x-y) \quad \text{where} \quad \begin{cases} \bar{u} = \frac{1}{|\Omega|} \sum_{x \in \Omega} u(x), \\ t_u = \frac{1}{\sqrt{|\Omega|}}(u - \bar{u})\mathbf{1}_\Omega \end{cases} \quad (7)$$

where $\mathbf{1}_\Omega$ is the indicator function, and where W is a normalized Gaussian white noise on \mathbb{Z}^2 (the convolution between t_u and W is computed in Fourier domain). U is a stationary Gaussian random field whose first and second order moments are the empirical mean and covariance of the exemplar texture. In particular

$$\mathbb{E}[(U(x) - \bar{u})(U(y) - \bar{u})] = a_u(y-x) \quad \text{where} \quad a_u(z) = \sum_{x \in \mathbb{Z}^2} t_u(z)t_u(x+z)^T. \quad (8)$$

Thus, U can be considered as a ‘‘Gaussianized’’ version of u , that has the correct correlations but no salient structures.

The second step consists in a patchwise operation. Let $\omega = \{-r, \dots, r\}^2$ be the patch domain with $r \in \mathbb{N}$. Let us denote by μ the distribution of patches of U , that is $\mu = \mathcal{N}(\bar{u}, C)$ with $C(x, y) = a_u(y-x)$, $x, y \in \omega$. Let us denote by ν the empirical distribution of patches of the exemplar texture u , that is $\nu = \frac{1}{|S|} \sum_{p \in S} \delta_p$, where $S = \{u_{|x+\omega} \mid x + \omega \subset \Omega\}$. Actually, in practice we approximate it with $\nu = \frac{1}{J} \sum_{j=1}^J \delta_{p_j}$ where p_1, \dots, p_J are $J = 1000$ patches randomly drawn from the exemplar texture. Thus, μ and ν are two probability measures on \mathbb{R}^D with $D = d(2r+1)^2$. Besides, μ is a Gaussian distribution that, except in degenerate cases, admits a probability density function ρ with respect to the Lebesgue measure. Using the algorithm explained in Sect. 2.2, we compute the optimal assignment T_v that realizes the semi-discrete OT from μ to ν . We then apply this mapping T_v to each patch of the Gaussian synthesis U , and we recombine an image V by averaging the ‘‘transported’’ patches: the value at pixel x is the average of values of x in all overlapping patches. More formally,

$$\forall x \in \mathbb{Z}^2, \quad P_x = T_v(U_{|x+\omega}), \quad (9)$$

$$\forall x \in \mathbb{Z}^2, \quad V(x) = \frac{1}{|\omega|} \sum_{h \in \omega} P_{x-h}(h). \quad (10)$$

Proposition 1. *V is a stationary random field on \mathbb{Z}^2 and satisfies the following long-range independence property: if Γ denotes the finite support of the auto-correlation function a_u defined in (8), then for every $A, B \subset \mathbb{Z}^2$ such that $(A - B) \cap (\Gamma + 4\omega) = \emptyset$ the restrictions $V|_A, V|_B$ are independent.*

Proof. Since U is Gaussian, $U(x) \perp\!\!\!\perp U(y)$ as soon as $x - y \notin \Gamma$. Therefore, if $x - y \notin \Gamma + 2\omega$, then $U|_{x+\omega} \perp\!\!\!\perp U|_{y+\omega}$ and thus $P_x \perp\!\!\!\perp P_y$. After averaging we get $V(x) \perp\!\!\!\perp V(y)$ as soon as $x - y \notin \Gamma + 4\omega$. The generalization to subsets A, B is straightforward.

This property is a guarantee of spatial stability for synthesis, meaning that the synthesis algorithm will not start to “grow garbage” as may do the method of [2]. We also have a guarantee on the patch distribution. Indeed, if T_ν is the true solution to the semi-discrete optimal-transport problem, then any patch P_x is exactly distributed according to ν . After recomposition, the distribution of $V|_\omega$ may not be exactly ν but is expected to be not too far away (as will be confirmed in Fig. 4). In this sense, V still respects the long range correlations (inherited from U) while better preserving the local structures.

4 Results and Discussion

In this section, we discuss experimental results of texture synthesis obtained with the proposed bilevel model. Figure 1 compares the synthesis results obtained with the Gaussian model before and after local transformation. One clearly observes the benefit of applying the optimal transport in the patch domain: it restores the color distribution of the exemplar and also creates salient features from the Gaussian content. In particular (Rows 1 and 3), it is able to break the symmetry of the color distribution which is a strong restriction of the Gaussian model.

Figures 1 and 2 demonstrate that the optimal assignment (OT) is better suited than the simple NN projection, as illustrated in Fig. 3(a). While these two operators project on the exemplar patches, the assignment is optimized to globally respect a statistical constraint. On Fig. 2 we also question the use of different patch sizes. The simple 1×1 case (which only amounts to apply a color map to the Gaussian field) is already interesting because it precisely respects the color distribution of the exemplar. Thus the bilevel model with 1×1 patch can be seen as an extension of the ADSN model with prescribed marginal distributions. For a larger patch size, we observe only minor improvements of the model with slightly cleaner geometric structures.

One possible explanation could be the slow convergence speed in very high dimensions of the stochastic gradient descent scheme. In this non-smooth setting ($\nabla v(x, \cdot)$ is not Lipschitz continuous), the convergence rate given in [18] is only $\mathcal{O}(\frac{\log k}{\sqrt{k}})$. Besides, in our setting, we apply this algorithm for 10^6 iterations with gradient step parameter $C = 1$ in a very high-dimensional space \mathbb{R}^D ($D = 27$ for 3×3 RGB patches), and with discrete target distributions ν having large

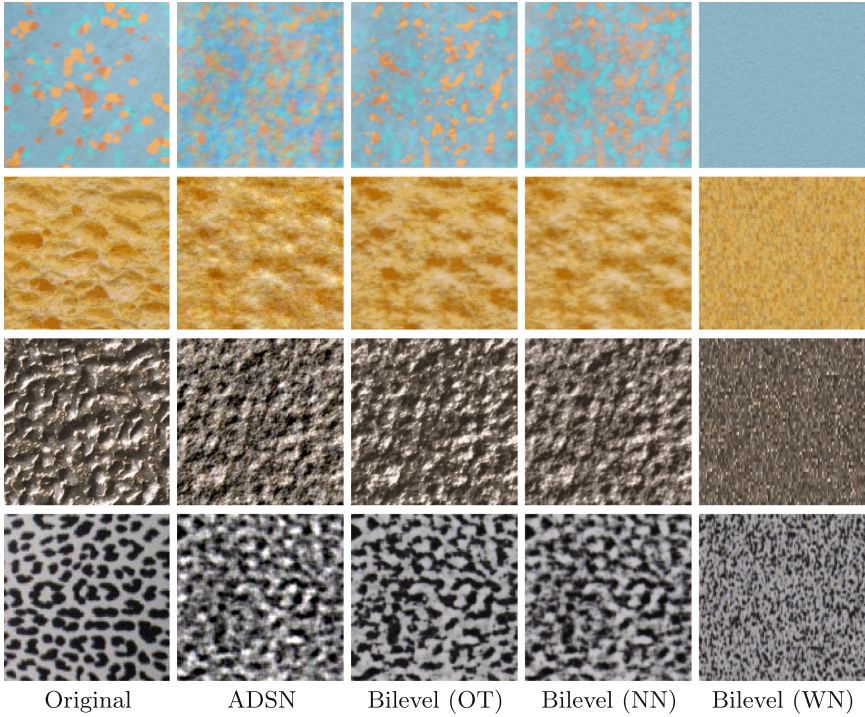


Fig. 1. Bilevel synthesis. Each row displays an exemplar texture, a sample of the associated ADSN model, and samples of the bilevel models obtained with 3×3 patch optimal assignment (OT) or patch nearest neighbor projection (NN), and the one with white noise initialization (WN). The OT assignment better preserves patch statistics than the NN projection. Besides, the last column illustrates the importance to start from a spatially correlated Gaussian model at the first level.

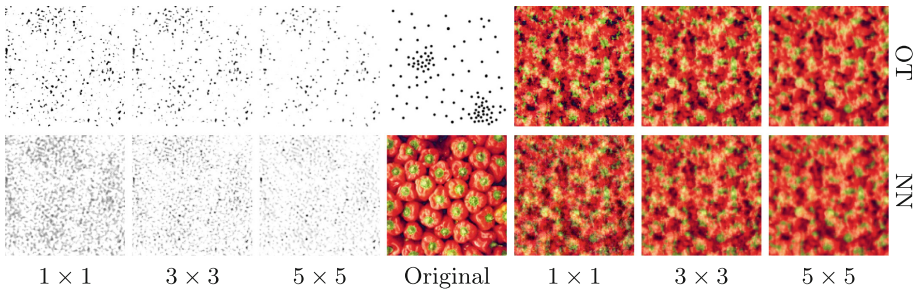


Fig. 2. Influence of the patch size. On the middle column we display two original textures and on the other columns we display samples of the bilevel models with varying patch size using the patch optimal assignment (OT, first row) and the patch nearest neighbor projection (NN, second row). The OT performs in general better than a NN (it better preserves the color/patch statistics) but fails to reproduce complex geometry (like in the right example). (Color figure online)

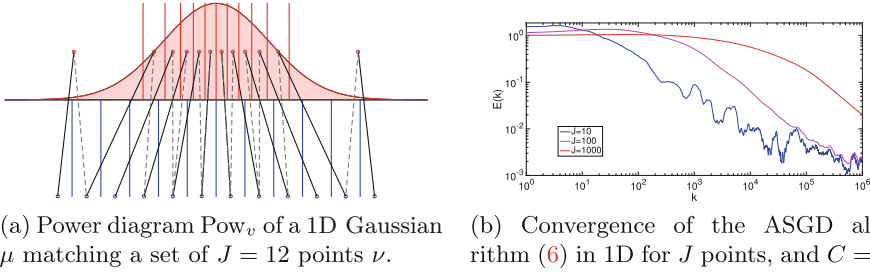


Fig. 3. Illustration of the semi-discrete OT and the convergence of ASGD (6) in 1D. (a) Semi-discrete Transport of a Gaussian distribution μ (red curve) towards a set of points ν (blue dots with $J = 12$). The corresponding power diagram Pow_ν (in red lines) is compared to the Voronoi diagram Pow_0 (blue lines). The optimal transport plan T_ν (black lines) is compared to the nearest-neighbor matching T_0 (grey dotted lines). (b) ASGD in 1D. Evolution of the relative error, defined as $E(k) = \frac{\|v_k - v^*\|}{\|v^*\|}$ where v^* is the (closed-form) optimal solution and k the number of iterations. The curves are shown for $J = 10, 10^2$, and 10^3 points, using the same random sequence. (Color figure online)

support ($J = |S| = 10^3$). Updating all coordinates of v during the gradient descent requires to visit enough all the power cells. As a consequence, even in the 1D case, practical convergence can be very slow (see on Fig. 3(b) for $C = 10$).

This means that in our experiments for texture synthesis, the convergence is not reached, and yet the OT solution provides better synthesis results than a simple NN approach, which is confirmed by examining the output patch distribution (Fig. 4). Interestingly, it is also true after patch recomposition (which remains to be properly justified).

The results obtained with this bilevel model raise several questions. First, one may very well question the use of the ℓ^2 -distance for patch comparison, both in the transport problem and the recomposition step. It is already well known [11] that using other distances for recomposition may improve the visual quality of the results (less blur than with the ℓ^2 average). Also, it would be interesting to analyze more precisely the effect of the recomposition step on the patch distribution.

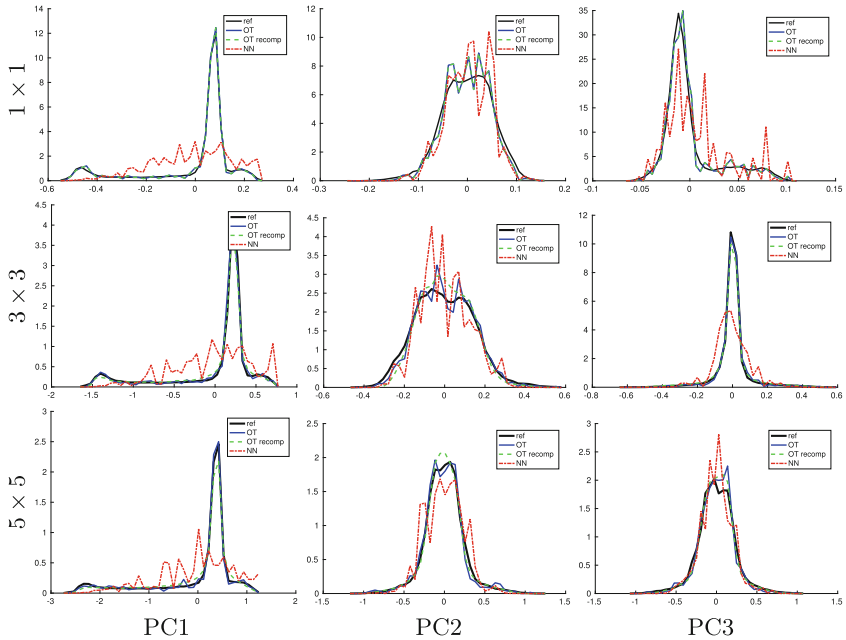


Fig. 4. Patch distribution (three first principal components). For the first image of Fig. 1 we plot the estimated distribution of patches in the three first principal components (columns) for different patch sizes (rows). The PCA transform is obtained on the exemplar patch distribution. We compare the patch distributions of the exemplar image (legend “ref”), of the synthesized image before patch recombination (legend “OT”) and after (legend “OT recomp”), and of the transformed patch with nearest-neighbor projection (legend “NN”). Even if we only approximate the optimal transport mapping, it suffices to reproduce the reference patch distribution better than the NN projection.

Acknowledgments. This work has been partially funded by Project Texto (Projet Jeunes Chercheurs du GdR Isis).

References

1. Aurenhammer, F., Hoffmann, F., Aronov, B.: Minkowski-type theorems and least-squares clustering. *Algorithmica* **20**(1), 61–76 (1998)
2. Efros, A.A., Leung, T.K.: Texture synthesis by non-parametric sampling. In: *Proceedings of the ICCV 1999*, p. 1033 (1999)
3. Efros, A., Freeman, W.: Image quilting for texture synthesis and transfer. In: *ACM TOG*, pp. 341–346, August 2001
4. Galerne, B., Gousseau, Y., Morel, J.M.: Random phase textures: theory and synthesis. *IEEE Trans. Image Process.* **20**(1), 257–267 (2011)
5. Galerne, B., Leclaire, A.: Texture inpainting using efficient Gaussian conditional simulation. *SIIMS* **10**(3), 1446–1474 (2017)

6. Galerne, B., Leclaire, A., Moisan, L.: A texton for fast and flexible Gaussian texture synthesis. In: Proceedings of the EUSIPCO, pp. 1686–1690 (2014)
7. Galerne, B., Leclaire, A., Moisan, L.: Texton noise. *Comput. Graph. Forum* (2017). doi:[10.1111/cgf.13073](https://doi.org/10.1111/cgf.13073)
8. Genevay, A., Cuturi, M., Peyré, G., Bach, F.: Stochastic optimization for large-scale optimal transport. In: Proceedings of the NIPS, pp. 3432–3440 (2016)
9. Gutierrez, J., Rabin, J., Galerne, B., Hurtut, T.: Optimal patch assignment for statistically constrained texture synthesis. In: Lauze, F., Dong, Y., Dahl, A.B. (eds.) *SSVM 2017. LNCS*, vol. 10302, pp. 172–183. Springer, Cham (2017). doi:[10.1007/978-3-319-58771-4_14](https://doi.org/10.1007/978-3-319-58771-4_14)
10. Kitagawa, J., Mérigot, Q., Thibert, B.: A Newton algorithm for semi-discrete optimal transport. *J. Eur. Math Soc.* (2017)
11. Kwatra, V., Essa, I., Bobick, A., Kwatra, N.: Texture optimization for example-based synthesis. *ACM TOG* **24**(3), 795–802 (2005)
12. Kwatra, V., Schödl, A., Essa, I., Turk, G., Bobick, A.: Graphcut textures: image and video synthesis using graph cuts. *ACM TOG* **22**(3), 277–286 (2003)
13. Lefebvre, S., Hoppe, H.: Parallel controllable texture synthesis. *ACM TOG* **24**(3), 777–786 (2005)
14. Levina, E., Bickel, P.: Texture synthesis and nonparametric resampling of random fields. *Ann. Stat.* **34**(4), 1751–1773 (2006)
15. Lévy, B.: A numerical algorithm for L2 semi-discrete optimal transport in 3D. *ESAIM: M2AN* **49**(6), 1693–1715 (2015)
16. Liang, L., Liu, C., Xu, Y.Q., Guo, B., Shum, H.Y.: Real-time texture synthesis by patch-based sampling. *ACM TOG* **20**(3), 127–150 (2001)
17. Mérigot, Q.: A multiscale approach to optimal transport. *Comput. Graph. Forum* **30**(5), 1583–1592 (2011)
18. Moulines, E., Bach, F.: Non-asymptotic analysis of stochastic approximation algorithms for machine learning. In: Proceedings of the NIPS, pp. 451–459 (2011)
19. Raad, L., Desolneux, A., Morel, J.: A conditional multiscale locally Gaussian texture synthesis algorithm. *J. Math. Imaging Vis.* **56**(2), 260–279 (2016)
20. Rabin, J., Peyré, G., Delon, J., Bernot, M.: Wasserstein barycenter and its application to texture mixing. In: Bruckstein, A.M., ter Haar Romeny, B.M., Bronstein, A.M., Bronstein, M.M. (eds.) *SSVM 2011. LNCS*, vol. 6667, pp. 435–446. Springer, Heidelberg (2012). doi:[10.1007/978-3-642-24785-9_37](https://doi.org/10.1007/978-3-642-24785-9_37)
21. Tartavel, G., Peyré, G., Gousseau, Y.: Wasserstein loss for image synthesis and restoration. *SIAM J. Imaging Sci.* **9**(4), 1726–1755 (2016)
22. Villani, C.: *Topics in Optimal Transportation*. American Mathematical Society, Providence (2003)
23. Wei, L.Y., Lefebvre, S., Kwatra, V., Turk, G.: State of the art in example-based texture synthesis. In: Eurographics, State of the Art Reports, pp. 93–117 (2009)
24. Wei, L.Y., Levoy, M.: Fast texture synthesis using tree-structured vector quantization. In: Proceedings of the SIGGRAPH 2000, pp. 479–488 (2000)
25. Xia, G., Ferradans, S., Peyré, G., Aujol, J.: Synthesizing and mixing stationary gaussian texture models. *SIAM J. Imaging Sci.* **7**(1), 476–508 (2014)