Texture synthesis guided by a low-resolution image

To cite this version:

HAL Id: hal-01444096
https://hal.science/hal-01444096
Submitted on 15 Jul 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
TEXTURE SYNTHESIS GUIDED BY A LOW-RESOLUTION IMAGE

M. El Gheche\textsuperscript{1,2}, J.-F. Aujol\textsuperscript{1}, Y. Berthoumieu\textsuperscript{2}, C.-A. Deledalle\textsuperscript{1}, R. Fablet\textsuperscript{3}

\textsuperscript{1}IMB and \textsuperscript{2}IMS CNRS laboratories, IPB, Université de Bordeaux, Talence, France
\{mireille.el-gheche, yannick.berthoumieu\}@u-bordeaux.fr
\{jean-francois.aujol, charles-alban.deledalle\}@math.u-bordeaux1.fr
\textsuperscript{3}Institut Mines-Télécom, Télécom Bretagne; UMR 6285 LabSTICC
Technopôle Brest Iroise, Brest, France
ronan.fablet@telecom-bretagne.eu

ABSTRACT
In this paper, we aim at synthesizing a texture from a high-resolution patch and a low-resolution image. To do so, we solve a nonconvex optimization problem that involves a statistical prior and a Fourier spectrum constraint. The numerical analysis shows that the proposed approach achieves better results (in terms of visual quality) than state-of-the-art methods tailored to super-resolution or texture synthesis.

Index Terms— Texture synthesis, Super resolution, Wasserstein distance, Random phase texture, Proximal algorithms, Nonconvex optimization.

1. INTRODUCTION
Texture synthesis consists of creating a large, coherent, and non-periodic texture image from a given sample. In this paper, we deal with the problem of texture synthesis under the assumption that a low-resolution version of the sought texture is available in addition of the given high-resolution sample. Roughly speaking, based on the down-sampling factor $r$, the following two cases may occur:

- If $r$ is low, the resulting problem is similar to image reconstruction, with the addition of a high-resolution patch.
- If $r$ is high, the resulting problem is similar to texture synthesis, with the addition of a low-resolution image.

The present paper is placed in the second context.

1.1. Related work
The problem addressed in this paper is closely related to texture synthesis. In fact, when no low-resolution is available, the two problems are identical. Existing texture synthesis algorithms can be broadly categorized into region-growing local methods and optimization-based global methods. Local approaches grow the texture one pixel (or patch) at a time, while maintaining the spatial coherence with nearby pixels by modeling the neighborhoods with Markov fields and fractal models [1, 2]. A weakness of these methods is that the spatial coherence between pixels is enforced at a local scale. A possible approach to circumvent this limitation consists of resorting to a small patch from which to grow the texture [3], or by using a multiscale setting [4, 5]. Global methods process the entire texture as a whole, using some criteria for measuring its similarity with a small texture patch. For example, the latter can be modeled with a statistical descriptor based on histograms [6] and Fourier coefficients [7].

Recently, an approach has been proposed in [8, 9], which introduces a preliminary step of dictionary learning for exploiting the given patch, and (not least) the Wasserstein distance for comparing the histograms of the entire texture with an extended version of the small patch. While the Wasserstein distance is well-known in image processing and computer vision under the name “earth mover distance” [10], it was only recently expanded to the context of texture synthesis [11, 9, 12]. Under the assumption that the main texture characteristics are contained in their Fourier magnitude [13], many works have shown that an efficient synthesis method is achieved when the texture phase is randomized [7, 9]. However, The resulting algorithm is limited to non-structured textures [14].

Regarding the problem considered in this paper, the authors in [15] propose to recover the high-resolution image from a pair of images: a complete low-resolution image and a high-resolution but incomplete one (sample). The resulting problem is solved using an extension of the nonlocal total variation model, where a set of connections is built between the missing high-resolution pixels and a set of pixels that lies in the sample. However, to interpolate the missing data, the authors in [16] design a new nonlocal graph that provides better connections between the missing pixels and the high-resolution pixels. Additionally, they introduce a histogram-based statistical prior modeled by a sum of Wasserstein distances between the histogram of some linear transformations of the texture.

The work in [17] exploits a patch-based nonlocal regularization and a Generalized Gaussian model of the texture gradients, whose parameters are learned on the high-resolution patch. Differently from [15] and [16], the authors in [17] use a prior
on the spatial covariance of the synthesized image.

In this paper, we propose a variational method that combines a histogram-based statistical prior [16] and a constraint on the power spectral density [17]. The main novelty of the proposed approach w.r.t. [16] lies in introducing a spectral constraint instead of the nonlocal regularization, and in considering the texture synthesis problem from free noise acquisitions with a high down-sampling factor. On the other hand, the originality of our technique w.r.t. [17] is twofolds: (i) the ability to consider multi-histogram priors without being constrained to a parametric model, (ii) the global optimization procedure that we carry out using a recent primal dual proximal algorithm.

The remainder of this paper is organized as follows. In Sec. 2, we present the texture synthesis as a nonconvex optimization problem. In Sec. 3, we propose to solve this problem by an iterative scheme based on proximal methods. Experimental results, showing the accuracy of the resulting images comparing with the state-of-the-art methods, are given in Sec. 4. Finally, some conclusions are drawn in Sec. 5.

1.2. Notations

Let $\| \cdot \|$ be the standard Euclidean norm, and let $\text{Id}$ be the identity matrix. The domain of a function $f : \mathbb{R}^N \to (-\infty, +\infty]$ is $\text{dom} \ f = \{ x \in \mathbb{R}^N | f(x) < +\infty \}$. $\Gamma_0(\mathbb{R}^N)$ is the class of lower semi-continuous convex functions from $\mathbb{R}^N$ to $(-\infty, +\infty]$ such that $\text{dom} \ f \neq \emptyset$. Let $f \in \Gamma_0(\mathbb{R}^N)$. The conjugate of $f$ is the function $f^* : \Gamma_0(\mathbb{R}^N)$ defined by $f^* : \mathbb{R}^N \to (-\infty, +\infty] : u \mapsto \sup_{x \in \mathbb{R}^N} x^T u - f(x)$. When $f$ is Gâteaux-differentiable at $y \in \mathbb{R}^N$, $\partial f(y) = \{ \nabla f(y) \}$ where $\nabla f(y)$ is the gradient of $f$ at $y$. A differentiable convex function has $\beta$-Lipschitz continuous gradient $\nabla f$ if $\forall (x,y) \in \mathbb{R}^N \times \mathbb{R}^N \quad \| \nabla f(x) - \nabla f(y) \| \leq \beta \| x - y \|$, where $\beta \in ]0, +\infty[$. Let $C$ be a nonempty subset of $\mathbb{R}^N$, then $1_C$ is the indicator function of $C$, equal to 0 on $C$ and $+\infty$ otherwise.

2. PROPOSED APPROACH

Let $\bar{x} \in \mathbb{R}^N$ be the unknown signal to be recovered (which generally corresponds to an image of size $N = N_1 \times N_2$), $z^{(1)} \in \mathbb{R}^Q$ the complete low-resolution image, and $z^{(2)} \in \mathbb{R}^M$ the given sample $^1$ such that such that

$$z^{(1)} = D\bar{x}, \quad z^{(2)} = M\bar{x},$$

where $D \in \mathbb{R}^{Q \times N}$ stands for spatial down-sampling by a dyadic factor in each direction yielding $Q = 2^{-r}N$, and $M \in \mathbb{R}^{M \times N}$ is a selection operator that extracts the patch from $\bar{x}$.

To recover $\bar{x}$ from the observations $z^{(1)}$ and $z^{(2)}$, we propose a variational approach that leads to solving the following optimization problem:

$$\minimize_{x \in \mathbb{R}^N} \| Dx - z^{(1)} \|^2 + t_{C_1}(x) + H(x, z^{(2)}) + t_{C_2}(x),$$

where $C_1 = \{ x \in \mathbb{R}^N | Mx = z^{(2)} \}$. Beside the data fidelity terms w.r.t. the observations $z^{(1)}$ and $z^{(2)}$, we use two additional pieces of information: a term $H(\cdot, z^{(2)})$ conveying some histogram-based statistics, and a Fourier spectrum constraint $C_2$. The histogram-based term is modeled through the Wasserstein distance discussed in Section 2.1, while the frequency constraint is presented in Section 2.2.

2.1. Statistical prior

The Wasserstein distance between the histograms$^2$ of two images $u \in \mathbb{R}^N$ and $v \in \mathbb{R}^N$ is defined as follows [11, 9]:

$$W_2^2(\nu_u, \nu_v) = \min_{\sigma \in \Sigma_N} \| u - v \circ \sigma \|^2,$$

where $\nu_u$ and $\nu_v$ are the normalized histograms of $u$ and $v$, the symbol $v \circ \sigma$ denotes a permutation of the vector $v$, and $\Sigma_N$ is the set of all the permutations of $N$-length vectors. For grayscale images, the optimal permutation $\sigma^*$ is computed as

$$\sigma^* = \sigma_v \circ \sigma_u^{-1},$$

where $\sigma_v$ (resp. $\sigma_u$) denotes the permutation operator that arranges the pixels of $v$ (resp. $u$) in ascending order. In our case, however, the two images have a different number of pixels ($u \in \mathbb{R}^N$ and $v \in \mathbb{R}^M$, with $M < N$). Hence, we replicate the patch $v$ so as to obtain a larger image $\tilde{v} \in \mathbb{R}^N$ such that the normalized histogram $\nu_\tilde{v}$ is equal to $\nu_v$. Although the Wasserstein distance is nonconvex (due to the histogram transformation), its gradient is Lipschitz-continuous and takes the following form [8, 12]

$$\nabla_u W_2^2(\nu_u, \nu_\tilde{v}) = 2(u - \tilde{v} \circ \sigma_\tilde{v} \circ \sigma_u^{-1}).$$

This property allows us to employ the Wasserstein distance into the optimization algorithm presented in Section 3.

In this work, inspired from [16], we define the term $H$ in (2) as

$$H(x, z^{(2)}) = \sum_{s=1}^4 \alpha_s W_2^2(\nu_{L_s x}, \nu_{z_s^{(2)}})$$

where, for every $s \in \{1, \ldots, 4\}$, $\alpha_s > 0$, $z_s^{(2)} = L_s z^{(2)}$, $z_s^{(2)}$ is the extension of $z_s^{(2)}$ (after the linear transformation), and for every $s \in \{1, \ldots, 4\}$, $L_s \in \mathbb{R}^{N_s \times N}$ is defined as follows

- $L_1$: the identity matrix ($N_1 = N$),
- $L_2$: the concatenation of the horizontal and vertical difference operators ($N_2 = 2N$),
- $L_3$: the concatenation of the diagonal difference operators ($N_3 = 2N$),
- $L_4$: the isotropic Laplacian operator ($N_4 = N$).

$^1$A sample denotes a small patch of the synthesized image.

$^2$The histogram of a signal refers to a histogram of the pixel intensity values.
whose expression is given in Sec. 3.

\[ \hat{x}(m) = \frac{1}{N_1N_2} \sum_{p \in \mathbb{R}^N} x(p) e^{-2i\pi (\frac{p1}{N_1} + \frac{p2}{N_2})}. \]

(8)

Consequently, Eq. (7) can be rewritten as a Fourier spectrum constraint defined as

\[ C_2 = \{x \in \mathbb{R}^N | \forall m, \exists \varphi(m) : \hat{x}(m) = e^{i\varphi(m)} \hat{z}^{(2)}(m) \} \]

(9)

Since \( x \) and \( \hat{z}^{(2)} \) are real images, \( \varphi(m) \) of Eq. (9) must be antisymmetric modulo \( 2\pi \). The projection onto the convex set \( C_2 \) consists in putting together the phase and the modulus, whose expression is given in Sec. 3.

Discussion We observed that Gaussian textures could be accurately synthesized using the spectral constraint introduced in Eq. (9). However, for structural textures with important oscillations, the latter hypothesis is not valid. Nonetheless, taking into account the statistical information through histogram priors and the low frequencies provided by the low-resolution acquisition, the proposed approach (see Eq. (2)) inverts the correct structures and details of the sample to the estimated texture, leading to improved results w.r.t. the classical spectral prior considered alone.

2.2. Fourier Spectrum constraint

Textures having the same second-order statistics share a common auto-covariance and, therefore, a common Fourier magnitude. So it is meaningful to introduce a constraint enforcing texture characteristics while preserving the Fourier magnitude, through a random phase textures. Hence, inspired from \([17]\), we search for an image \( x \in \mathbb{R}^N \) such that

\[ \forall m, |\hat{x}(m)| = |\hat{z}^{(2)}(m)|, \]

(7)

where \( \hat{x} \) (resp. \( \hat{z}^{(2)} \)) is the orthogonal discrete Fourier transform of \( x \) (resp. \( z^{(2)} \)), whose coefficients read

\[ \hat{z}(m) = \frac{1}{N_1N_2} \sum_{p \in \mathbb{R}^N} x(p) e^{-2i\pi (\frac{p1}{N_1} + \frac{p2}{N_2})}. \]

(8)

The solution of Eq. (2) requires an efficient algorithm for dealing with problems involving nonsmooth functions and linear operators. Recently, it has been shown experimentally that primal-dual proximal methods \([20, 21, 22, 23, 24, 25]\), which were originally designed for convex optimization, can be also applied to nonconvex problems in some circumstances \([26, 27]\). In the convex setting, the key tool of these methods is the proximity operator \([28]\) of a lower semicontinuous convex function \( f : \mathbb{R}^N \rightarrow ]-\infty; +\infty[ \), defined as

\[ (\forall y \in \mathbb{R}^N) \quad \text{prox}_f(y) = \arg\min_{z \in \mathbb{R}^N} f(z) + \frac{1}{2} \|z - y\|^2. \]

Proximy operators enjoy many properties \([29]\). In particular, they generalize the notion of projection onto a closed convex set \( C \), in the sense that \( \text{prox}_{\mathbb{R}^N} = \mathbb{R}^N \). Among the wide array of existing proximal algorithms, we employ the Forward-Backward Primal Dual method (FBPD) \([24]\) reported in Algorithm 1.

Algorithm 1 FBPD \([24]\)

**Initialization**

\[
\begin{align*}
\text{Choose } & (x^{[0]}, y^{[0]}) \in \mathbb{R}^n \times \mathbb{R}^{Kn}, \\
\text{set } & \tau > 0 \text{ and } \omega > 0 \text{ such that } \tau (\beta/2 + \omega) < 1.
\end{align*}
\]

**For** \( l = 0, 1, \ldots \)

\[
\begin{align*}
\tilde{z}^{[l]} &= \nabla f(x^{[l]}) + y^{[l]}, \\
x^{[l+1]} &= P_{\{x \rightarrow z^{(2)}} (x^{[l]} - \tau \tilde{z}^{[l]}), \\
\tilde{y}^{[l]} &= (2x^{[l+1]} - x^{[l]}), \\
y^{[l+1]} &= y^{[l]} + \omega \tilde{y}^{[l]} - P_{C_2} (y^{[l]} + \omega \tilde{y}^{[l]}).
\end{align*}
\]

The operators required by this algorithm are detailed below.
We address the problem by considering a hierarchical approach, which consists of initializing our algorithm with the low-resolution image and iterating three consecutive steps: (i) up-sample the initialization by a factor 2 using a bicubic interpolation, (ii) apply a patch-based approach [30], and (iii) run the FBPD algorithm. These three steps are repeated in loop until the size of the output signal is equal to the sought one (512x512 in our case).

4. EXPERIMENTAL RESULTS

4.1. Algorithm and Initialization

In this paper, we have investigated the application of texture synthesis guided by a low resolution image. The proposed approach is adequate for various texture images. It is also able to exploit the potentials offered by multicore/GPU parallel architectures. However, one of the current limitations of the proposed method is that it is nonconvex, thus requiring a good initialization. In our future work, we therefore plan to consider a convex relaxation of the histogram prior in order to have a convex minimization problem [32].

4.2. Results

This section provides numerical results of our synthesis algorithm. We illustrate the effect of the Fourier constraint on the outcome of Eq. (2). We compare our results to the most classical synthesis method [31] and to that of super-resolution method [18]. Fig. 1 gives two texture examples where the down-sampling factor is equal to 4, and the sample amounts to 25% of the total image. As illustrated by these experiments, we assess the performance achieved with and without the spectral constraint. Using the proposed formulation (2), we obtain the best visual results while the state-of-the-art super-resolution method [18] tends to smooth the textures and the state-of-the-art texture synthesis approaches [19] produce an image (by replicating the patch) without exploiting information provided by the low-resolution image. Hence, the results may be far from the ground truth. Fig. 2 gives a similar example where the down-sampling factor is equal to 8 and the sample amounts to 25% of the total image. As illustrated by these experiments, we assess the performance achieved with and without the spectral constraint. The obtained results are better than the ones achieved without the spectral prior. Formulation (2) leads to a better texture reconstruction in the synthesized images compared to the state-of-the-art methods. This is confirmed by looking at the reconstructed textures and the low-resolution image. Even if the synthesized image of [19] is close to the sample but it is far from the low-resolution guide.

5. CONCLUSIONS AND FUTURE WORK

The gradient of the sum of the remaining terms, that is \( f(x) = \|Dx - z^{(1)}\|_2^2 + \sum_{s=1}^{4} \alpha_s W_s^2 (\nu L_s x, \nu \bar{z}(2)) \), reads

\[
\nabla f = 2D^\top (Dx - z^{(1)}) + 2 \sum_{s=1}^{4} \alpha_s L_s^\top (L_s x - \bar{z}(2) \circ \sigma^{-1}_{L_s x}),
\]

where \( \nabla f \) is \( \beta \)-Lipschitz with \( \beta = 2(1 + \sum_{s=1}^{4} \alpha_s \| L_s \|^2) \).

Although there is no theoretical guarantee about the estimate produced by Algorithm 1, in our experiments we observed that it always converges to a stable solution.
6. REFERENCES


