Optimal Transport Mixing of Gaussian Texture Models

Sira Ferradans, Gui-Song Xia, Gabriel Peyré, Jean-François Aujol

OPTIMAL TRANSPORT MIXING OF GAUSSIAN TEXTURE MODELS

Sira Ferradans, Gui-Song Xia, Gabriel Peyré

CEREMADE, Univ. Paris-Dauphine

Jean-François Aujol

Univ. Bordeaux, IMB, UMR 5251

ABSTRACT

This paper tackles the problem of mixing color texture models learned from an input dataset. We focus on stationary Gaussian texture models, also known as spot noises. We derive the barycenter and geodesic path between models according to optimal transport. This allows the user to navigate inside the set of texture models, and perform texture synthesis from the obtained interpolated models. Numerical examples on a library of exemplars show the ability of our method to generate arbitrary interpolations among unstructured natural textures.

Index Terms— Optimal transport, gaussian model, texture synthesis, texture mixing.

1. INTRODUCTION

The problem of synthesizing new textures is central in image processing and computer graphics. In order to make a scene realistic without having to simulate the response of materials to light for video games or animation films, a texture is mapped onto a given surface. Because the shape and extension of the surface may vary, the main goal of texture synthesis is to be able to generate as much texture as it is needed in a fast and realistic way. This problem has been addressed since the beginning of computer graphics, so we can find many solutions in the literature. Patch-based methods are adapted to very complicated (not even random) textures, see e.g. [1]. Statistical parametric models are generally not as good in handling complex texture patterns, but are much more flexible and fast, see for instance [2]. In this paper, we focus our attention on what could be the simplest statistical texture model, namely stationary Gaussian distribution, following and extending the methodology initiated by Galerne et al [3].

More complex textures can be obtained by texture mixing which extends the traditional texture synthesis by considering the interplay between several texture models. This is a difficult problem since it requires to average very distinct statistical features. Previous works make use of mixture models, see for instance [4]. The use of non-parameteric histogram averaging has also been proposed for grayscale [5] as well as color and wavelets features [6]. We propose here a simpler approach that makes use of a parameterization of the Gaussian texture model. Defining a geodesic path with optimal transport between the original Gaussian models, we can generate new textures sharing the characteristics of the input ones. The proposed method ensures that the new texture model stays Gaussian.

Optimal transport (OT) [7] is used intensively in computer vision as a metric between statistical features [8]. OT with non-parametric point clouds discretization is used in [6] to handle statistical constraints during texture synthesis. We propose here a radically different approach that uses a compact parametric model of textures using Gaussian distributions. This approach is well suited to handle stationary textures and leads to fast numerical schemes. Note that Gaussian optimal transport has been used for color manipulation [9], but never to achieve image modeling and synthesis.

This paper is organized as follows: In section 2, we review the Spot Noise model proposed by Galerne et al. [3], which allows us to define for each input texture a Gaussian model and extend it to color textures. In section 3, we present our first contribution, we derive the geodesic path defined by the optimal transport metric between two Spot noise models. Here, we also explain how to synthesize new textures following this geodesic path. In section 4, we present our second contribution which is the extension of this idea to several input textures, deriving the formula to compute the barycenter of a group of Gaussian models. Finally, we show some results with two and more input textures.

2. SPOT NOISE TEXTURE MODEL

2.1. Stationary Gaussian Models

Let us consider the modeling of textures $f \in \mathbb{R}^{N\times d}$, where $N = N_1 \times N_2$ is the number of pixels in the image, and d is the number of channels ($d = 1$ for grayscale and $d = 3$ for color datasets). We define a Gaussian distribution as $\mu = \mathcal{N}(m, \Sigma)$ where $m \in \mathbb{R}^{N\times d}$ is the mean of the distribution and $\Sigma \in \mathbb{R}^{Nd\times Nd}$ is a positive semi-definite covariance matrix.

We focus on stationary random fields $X \sim \mu$, which means that the distributions of X and $X(\cdot + \tau)$ are the same for any translation vector $\tau \in \mathbb{Z}^2$. For simplicity of the ex-

This work has been supported by the European Research Counsil (ERC project SIGMA-Vision) and the French National Research Agency (project NatImages).

position, we consider here periodic boundary conditions, and will detail later how to perform the learning of the parameter for a non-periodic exemplar input. Defining the model as stationary is equivalent to imposing that the mean m is a constant vector in $\mathbb{R}^d$ and that the covariance of the channel is a convolution. This is conveniently expressed using the 2-dimensional discrete Fourier transform (applied in parallel to each of the d channels)

$$\forall\, \omega = (\omega_1, \ldots, \omega_k), \quad \hat{f}(\omega) = \sum_x f(x) e^{\sum_j \frac{2i\pi}{N_j} \omega_j x_j} \in \mathbb{R}^d.$$

The transform $\hat{f} \in \mathbb{R}^{N \times d}$ is computed in $O(Nd\log(N))$ operations using the FFT, and this transform is inverted with the same complexity using the inverse FFT.

Defining our model μ as stationary is equivalent to stating that Σ is block-diagonal over the Fourier domain. This means that the covariance operator $y = \Sigma f$ can be applied over the Fourier domain as $\hat{y}(\omega) = \hat{\Sigma}(\omega)\hat{f}(\omega)$ where $\hat{\Sigma}(\omega) \in \mathbb{C}^{d \times d}$ is a positive hermitian matrix.

2.2. Spot Noise Models

Given some deterministic input exemplar $f \in \mathbb{R}^{N \times d}$, it makes sense to learn from f the parameters of a Gaussian model using the maximum likelihood estimator (MLE) which can be shown to be equivalent to the spot noise model introduced by Galerne et al. [3]. A random field $X = (X_1, \ldots, X_d)$ distributed according to the spot noise $\mu = \mu(f)$ associated to $f = (f_1, \ldots, f_d) \in \mathbb{R}^{N \times d}$ reads

$$\forall\, j = 1, \ldots, d, \quad X_j = m_j + f_j \star W_j \tag{1}$$

where $\star$ is the periodic convolution and the W_j are i.i.d. white noises $W_j \sim \mathcal{N}(0, \mathrm{Id}_N)$. Equivalently, spot noise models are the stationary Gaussian processes for which the matrices $\hat{\Sigma}(\omega)$ are rank one, and can thus be decomposed as $\hat{\Sigma}(\omega) = \hat{f}(\omega)\hat{f}(\omega)^*$ where $u^* \in \mathbb{C}^d$ is the complex conjugate transpose of $u \in \mathbb{C}^d$.

2.3. Stationary Gaussian Model Synthesis

Once the matrices $\hat{\Sigma}(\omega) \in \mathbb{C}^{d \times d}$ are computed, the synthesis of a texture $g \in \mathbb{R}^{N \times d}$ is obtained using a realization of the Gaussian process. In the general case, this is achieved by factorizing the frequency covariance $\hat{\Sigma}(\omega) = \hat{A}(\omega)\hat{A}(\omega)^*$ (for instance using the singular value decomposition) where $A(\omega) \in \mathbb{C}^{d \times d}$ and computing $\hat{g}(\omega) = \hat{A}(\omega)\hat{w}(\omega)$ for $\omega \neq 0$ where w is a realization of $\mathcal{N}(0, \mathrm{Id}_{Nd})$ and $\hat{g}(0)$ is the constant mean of the model.

In the special case where the model is a spot noise $\mu(f)$, meaning that $\hat{\Sigma}(\omega) = \hat{f}(\omega)\hat{f}(\omega)^*$, the synthesis is even faster using (for $\omega \neq 0$) $\hat{g}(\omega) = \hat{w}(\omega)\hat{f}(\omega)$, or equivalently using a realization of the convolution formula (1).

Our modeling process of the input texture is based on the FFT, thus the image is assumed to be periodic. But, symmetrizing on the boundaries introduces artificial features that may not be present in the input textures. We propose to process only the *periodic component* as defined by Moisan [10] instead of the original input texture. That is to say, substitute all f_j by its periodic component.

In our context, the process of extending the input texture of size $N_1 \times N_2$ to any arbitrary size $M_1 \times M_2$ can be done easily following the method proposed by Galerne et al. [3]: The periodic component of the original texture is located at the center of a flat new image of value m and dimensions $M_1 \times M_2$. To avoid the introduction of the high frequencies, the new borders are smoothed with a spatial weight. We process this new image instead of the original texture.

3. OPTIMAL TRANSPORT GEODESIC OF SPOT NOISE

In order to manipulate texture models, we use a geodesic distance between probability distributions, namely the L^2 optimal transport distance [7].

3.1. Optimal Transport of Gaussian Fields

OT distances are geodesic distances between arbitrary distributions. They are suitable to compare even singular distributions, which is crucial for color spot noises which have rank-deficient covariances. The L^2 OT distance between $\mu_i = \mathcal{N}(m_i, \Sigma_i)$ reads

$$d(\mu_0, \mu_1)^2 = \mathrm{tr}\left(\Sigma_0 + \Sigma_1 - 2\Sigma_{0,1}\right) + \|m_0 - m_1\|^2,$$

where $\Sigma_{0,1} = (\Sigma_1^{1/2}\Sigma_0\Sigma_1^{1/2})^{1/2}$ (see for instance [11].)

3.2. Optimal Transport Geodesics

If d is a geodesic distance (as this is the case for the OT distance), then the geodesic mixing of two distributions μ_0 and μ_1 is defined as

$$\forall\, t \in [0,1], \quad \mu_t = \operatorname*{argmin}_{\mu} (1-t)d(\mu_0, \mu)^2 + td(\mu_1, \mu)^2,$$

which defines a geodesic path $t \mapsto \mu_t$ linking μ_0 to μ_1. The following proposition details how to compute this geodesic path for arbitrary Gaussian distributions.

Proposition 1. *If* $\ker(\Sigma_0) \cap \mathrm{Im}(\Sigma_1) = \{0\}$*, the OT geodesic of Gaussian distributions* $\mu_i = \mathcal{N}(m_i, \Sigma_i)$ *(for* $i = 0, 1$*) is a Gaussian distribution* $\mathcal{N}(m_t, \Sigma_t)$ *where* $m_t = (1-t)m_0 + tm_1$ *and*

$$\Sigma_t = [(1-t)\mathrm{Id} + tT]\Sigma_0[(1-t)\mathrm{Id} + tT] \tag{2}$$

where $T = \Sigma_1^{1/2}\Sigma_{0,1}^{+}\Sigma_1^{1/2}$ *and where* A^+ *is the Moore-Penrose pseudo-inverse and* $A^{1/2}$ *is the unique positive square root of a symmetric semi-definite matrix.*

Fig. 1. $f^{[0]}, f^{[1]}, f^{[2]}$ original textures, see text.

Proof. The proof follows the one in [12] with the extra care that the covariance can be rank-deficient, hence requiring a pseudo-inverse. □

Note that for rank-1 covariances, the conditions $\ker(\Sigma_0) \cap \mathrm{Im}(\Sigma_1) = \{0\}$ means that the leading eigenvectors of the covariances should not be orthogonal. If this condition holds, there exists an infinite number of geodesic paths. The following theorem shows that the geodesic between two spot noise models is a spot noise computed by a suitable linear blending of the Fourier transforms.

Theorem 1. *For $i = 0, 1$, let $\mu_i = \mu(f^{[i]})$ be spot noise distributions associated with $f^{[0]}, f^{[1]} \in \mathbb{R}^{N \times d}$. The OT geodesic path μ_t defined in equation (2) is a spot noise model $\mu_t = \mu(f^{[t]})$ where $f^{[t]} = (1-t)f^{[0]} + tg^{[1]}$ with*

$$\forall\, \omega, \quad \hat{g}^{[1]}(\omega) = \hat{f}^{[1]}(\omega) \frac{\hat{f}^{[1]}(\omega)^* \hat{f}^{[0]}(\omega)}{|\hat{f}^{[1]}(\omega)^* \hat{f}^{[0]}(\omega)|}. \tag{3}$$

Proof. Denoting $\hat{\Sigma}_i(\omega) = \hat{f}^{[i]}(\omega)\hat{f}^{[i]}(\omega)^* \in \mathbb{C}^{d \times d}$ the Fourier covariances, one shows that T is block diagonal over the Fourier domain and that

$$\forall\, \omega, \quad \hat{T}(\omega) = \frac{\hat{f}^{[1]}(\omega)\hat{f}^{[1]}(\omega)^*}{|\hat{f}^{[1]}(\omega)\hat{f}^{[0]}(\omega)^*|}.$$

Equation (2) implies that Σ_t is block diagonal over the Fourier domain with $\hat{\Sigma}_t(\omega) = \hat{f}^{[t]}(\omega)\hat{f}^{[t]}(\omega)^*$ where $\hat{f}^{[t]}(\omega) = [(1-t)\mathrm{Id} + t\hat{T}(\omega)]\hat{f}^{[0]}(\omega)$ which corresponds to the expression (3). □

3.3. Numerical Examples

On the columns of Fig. 1, we can see four sets of three input textures. Let us focus on the first two input textures: $f^{[0]}$ and $f^{[1]}$. Along this section, we showed how to compute the in-between models $f^{[t]}$, defined along the geodesic path between $f^{[0]}$ and $f^{[1]}$. In Fig. 2, we can see a realization of two of these $f^{[t]}$ models, tagged as $1, 2$ respectively. Note that both images maintain features of the two original textures. We would also like to note that the synthesized $f^{[0]}$ and $f^{[1]}$ tagged as 0,3 respectively, reproduce the perceptual features of the original textures.

4. OPTIMAL TRANSPORT BARYCENTER OF SPOT NOISE

4.1. Optimal Transport Barycenter

Given a family of Gaussian process $(\mu_i)_{i \in I}$ and weights ρ_i with $\sum_i \rho_i = 1$, the OT barycenter is defined as

$$\mu^\star = \operatorname*{argmin}_{\mu} \sum_{i \in I} \rho_i d(\mu_i, \mu)^2. \tag{4}$$

For the special case of a Gaussian distribution $\mu_i = \mathcal{N}(m_i, \Sigma_i)$, there is no close form solution if $|I| > 2$, but the barycenter can be shown to be Gaussian $\mu^\star = \mathcal{N}(m^\star, \Sigma^\star)$ where $m^\star = \sum_{i \in I} \rho_i m_i$ and the covariance matrix is solution of the following fixed point equation

$$\Sigma^\star = \sum_{i \in I} \rho_i \left(\Sigma^{1/2} \Sigma_i \Sigma^{1/2} \right)^{1/2}. \tag{5}$$

This barycenter can be shown to be unique if one of the Σ_i is full rank [13]. We leave for future work the theoretical analysis of the uniqueness when all the covariances are rank-deficient.

4.2. Spot Noise Barycenter

When $\mu_i = \mu(f^{[i]})$ are spot noise, the covariance $\Sigma^\star$ of the barycenter is block diagonal over the Fourier domain, and the blocks $\hat{\Sigma}^\star(\omega)$ satisfies the fixed point equation $\hat{\Sigma}^\star(\omega) = \Phi_\omega(\hat{\Sigma}^\star(\omega))$ with

$$\Phi_\omega(\Sigma) = \sum_{i \in I} \rho_i \left(\Sigma^{1/2} \hat{\Sigma}_i(\omega) \Sigma^{1/2} \right)^{1/2}.$$

We note that in general, $\mu^\star$ is not a spot noise because $\hat{\Sigma}^\star(\omega)$ is not necessarily rank one. Following [14], we propose to compute $\hat{\Sigma}^\star(\omega)$ by iterating the mapping Φ_ω, i.e. compute the sequence $\hat{\Sigma}^{(k+1)}(\omega) = \Phi_\omega(\hat{\Sigma}^{(k)}(\omega))$. Although the mapping Φ_ω is not strictly contracting, we observe numerically the convergence $\hat{\Sigma}^{(k)}(\omega) \to \hat{\Sigma}^\star(\omega)$ when $k \to +\infty$. The numerical computation of Φ_ω in the case $d = 3$ requires the computation of the square root of 3×3 matrices, which is performed explicitly by computing the eigenvalue of the symmetric matrix as the root of a third order polynomial.

4.3. Numerical Examples

Given three input textures, $f^{[0]}, f^{[1]}, f^{[2]}$, and the path defined in Fig. 3 by the red numbers in increasing order, we generated the Gaussian models associated to each point. A

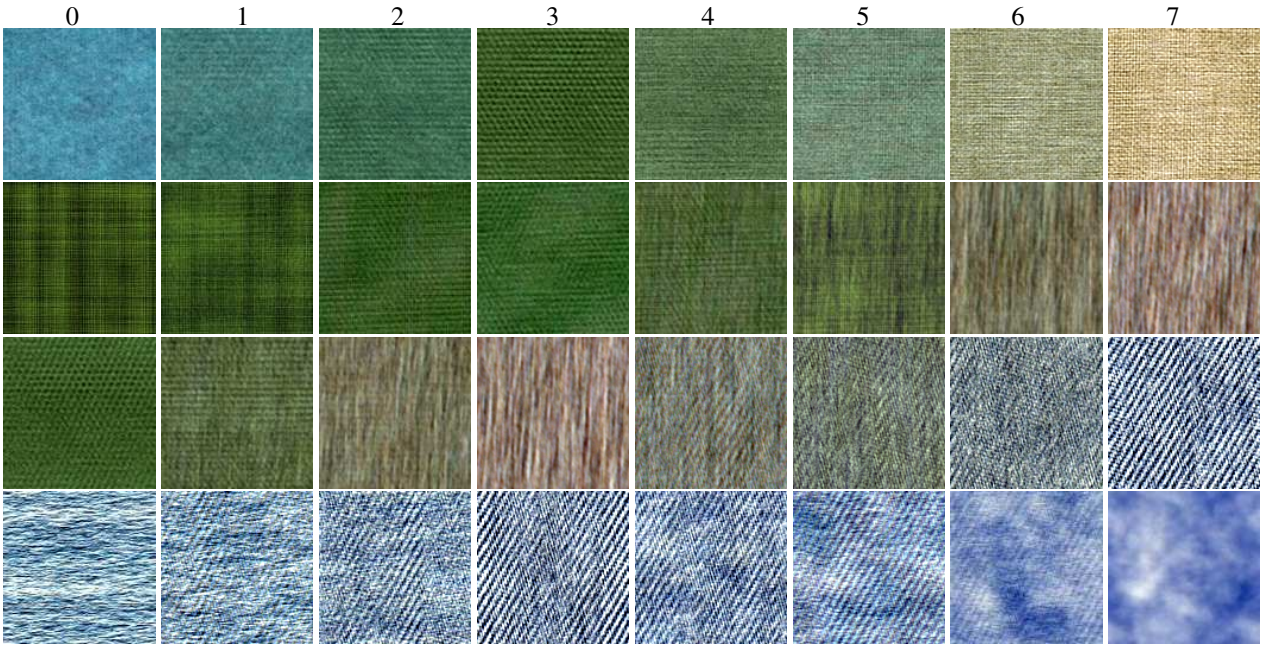

Fig. 2. Each row corresponds to a single experiment whose input textures are shown as the columns of Fig. 1, respectively. The parameter ρ_i of eq. 4 has been defined according to the triangle coordinates of the points in Fig. 3.

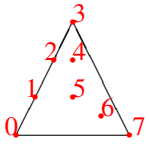

Fig. 3. Spatial location scheme.

realization of each of these models can be observed in Fig. 2 using as input textures the columns of Fig. 1. Again, as we approach in the path to an original model, the features of it tend to predominate in the synthesized texture. Note that this method is also able to reproduce small periodic patterns. More examples can be checked in http://www.enst.fr/~xia/mixingGaussianTextures.html

5. CONCLUSION

We have presented a new method for texture synthesis and mixing based on the geodesic path defined by the OT metric on Gaussian models. We showed that using this geodesic path between two Gaussian models, we can generate new textures with visual features of the original images. Using the barycenter of a set of Gaussian models we generalized the result to more that two input textures. The instances of these new in-between models produced natural results while reproducing the features of the original textures.

6. REFERENCES

[1] A. A. Efros and T. K. Leung, "Texture synthesis by non-parametric sampling," in *Proc. of ICCV '99*, 1999, p. 1033.

[2] J. Portilla and E. P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," *Int. Journal of Computer Vision*, vol. 40, no. 1, pp. 49–70, 2000.

[3] 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.

[4] Z. Bar-Joseph, R. El-Yaniv, D. Lischinski, and M. Werman, "Texture mixing and texture movie synthesis using statistical learning," *IEEE Tr. on Vis. and Comp. Graph.*, vol. 7, no. 2, pp. 120–135, 2001.

[5] W. Matusik, M. Zwicker, and F. Durand, "Texture design using a simplicial complex of morphable textures," *ACM Transactions on Graphics*, vol. 24, no. 3, pp. 787–794, 2005.

[6] J. Rabin, G. Peyré, J. Delon, and M. Bernot, "Wasserstein barycenter and its application to texture mixing," *Proc. SSVM'11*, 2011.

[7] C. Villani, *Topics in Optimal Transportation*, American Mathematical Society, 2003.

[8] Y. Rubner, C. Tomasi, and L. J. Guibas, "The earth mover's distance as a metric for image retrieval," *International Journal of Computer Vision*, vol. 40, no. 2, pp. 99–121, Nov. 2000.

[9] F. Pitié, A. Kokaram, and R. Dahyot, "Automated colour grading using colour distribution transfer," *Computer Vision and Image Understanding*, February 2007.

[10] L. Moisan, "Periodic plus smooth image decomposition," *J. Math. Imaging Vis.*, vol. 39, no. 2, pp. 161–179, February 2011.

[11] D. C. Dowson and B. V. Landau, "The fréchet distance between multivariate normal distributions," *J. Multivariate Anal.*, vol. 3, no. 12, pp. 450–455, 1982.

[12] A. Takatsu, "Wasserstein geometry of gaussian measures," *Osaka J. Math.*, , no. 45, 2011.

[13] M. Agueh and G. Carlier, "Barycenters in the wasserstein space," *SIAM J. on Mathematical Analysis*, vol. 43, no. 2, pp. 904–924, 2011.

[14] M. Knott and C. S. Smith, "On a generalization of cyclic monotonicity and distances among random vectors," *Linear Algebra and its Applications*, vol. 199, pp. 363–371, 1994.