Optimal transport between GMM for multiscale texture synthesis

Julie Delon[1], Agnès Desolneux[2], Laurent Facq[3], Arthur Leclaire[3]

[1] Laboratoire MAP5, Université Paris Cité

[2] Centre Borelli, CNRS and ENS Paris-Saclay

[3] Institut de Mathématiques de Bordeaux, with support of the GdR Isis

Arthur.Leclaire@math.u-bordeaux.fr https://www.math.u-bordeaux.fr/∼aleclaire/texto/

1) Patch Optimal Transport for Texture Synthesis

Goal: Texture synthesis from an example $u : \Omega \to \mathbb{R}^d$

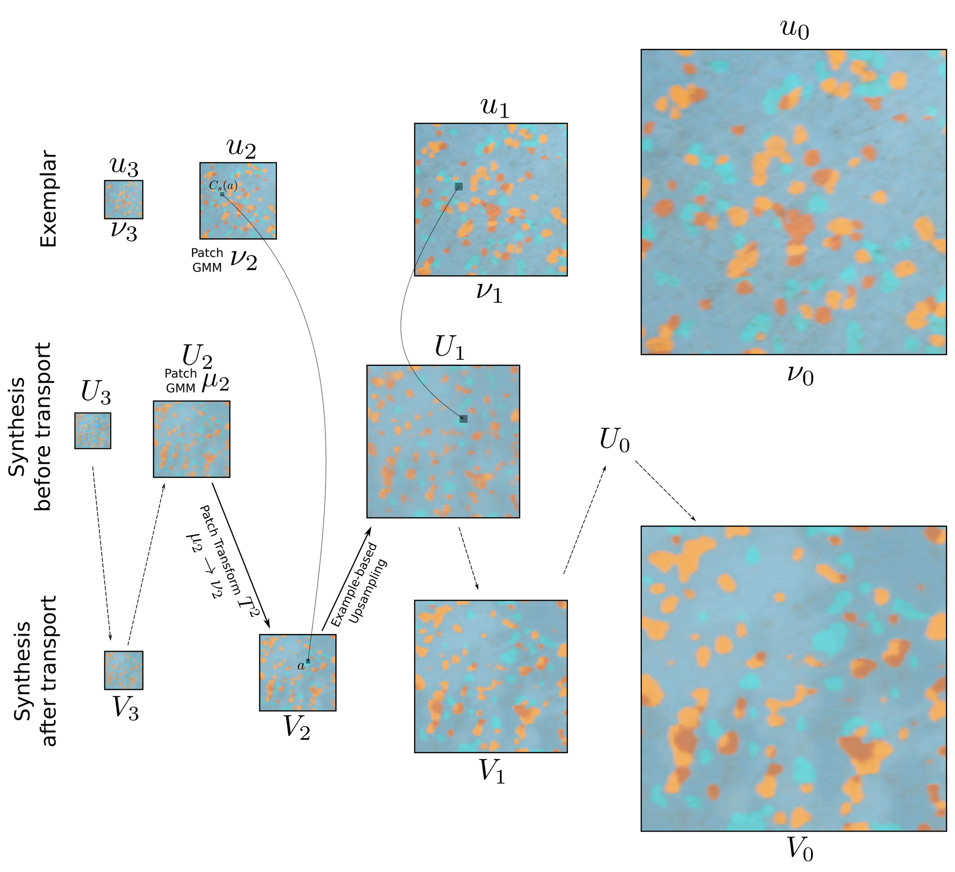
Proposed Solution: Patch-based Multiscale Synthesis:

- Compute original at scales $s = 0, \ldots, S-1$,

  $$u_s : \Omega_s \to \mathbb{R}^d \text{ with } \Omega_s \subset 2^s\mathbb{Z}^2$$

- Extract target patch distributions ν_s at scale s
- Apply patch optimal transport at each scale in order to reimpose the patch distribution ν_s on the current synthesis.

Patches are defined on a square domain ω of size $w \times w$.



2) Gaussian Mixture Models

Definition 1 *We say that $\mu \in \mathsf{GMM}_d(K)$ if*

$$\mu = \sum_{k=1}^{K} \pi_k \mathcal{N}(m_k, \Sigma_k) \text{ with } \begin{cases} \pi \in \mathbb{R}_+^K, \ \sum \pi_k = 1 \\ m_k \in \mathbb{R}^d, \ \Sigma_k \ s.d.p. \end{cases}$$

The set of such GMM is denoted by $\mathsf{GMM}_d(K)$, and let

$$\mathsf{GMM}_d = \bigcup_{K \geq 1} \mathsf{GMM}_d(K).$$

→ Inference from samples via Expectation-Maximization (EM algorithm)

3) Gaussian Optimal Transport

Let μ_0, μ_1 be two probability measures on $\mathbb{R}^d$.

$$W_2^2(\mu_0, \mu_1) := \inf_{\gamma \in \Pi(\mu_0, \mu_1)} \int \|y_0 - y_1\|^2 d\gamma(y_0, y_1),$$

where $\Pi(\mu_0, \mu_1)$ is the set of measures on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals μ_0, μ_1.
If $\mu_0 = \mathcal{N}(m_0, \Sigma_0), \mu_1 = \mathcal{N}(m_1, \Sigma_1)$, then

$$W_2^2(\mu_0, \mu_1) = \|m_0 - m_1\|^2 + \mathrm{Tr}\left(\Sigma_0 + \Sigma_1 - 2\left(\Sigma_0^{\frac{1}{2}} \Sigma_1 \Sigma_0^{\frac{1}{2}}\right)^{\frac{1}{2}}\right),$$

If Σ_0 is non-singular, there is an optimal map

$$\forall x \in \mathbb{R}^d, \quad T(x) = m_1 + \Sigma_0^{-1}(\Sigma_0\Sigma_1)^{\frac{1}{2}}(x - m_0).$$

4) GMM Optimal Transport

Let $\mu_0 = \sum_{k=1}^{K_0} \pi_0^k \mu_0^k, \ \mu_1 = \sum_{k=1}^{K_1} \pi_1^k \mu_1^k \in \mathsf{GMM}_d.$

Definition 2 *The GMM-OT cost is defined as*

$$\inf_{\gamma \in \Pi(\mu_0, \mu_1) \cap GMM_{2d}} \int \|y_0 - y_1\|^2 d\gamma(y_0, y_1).$$

*A solution γ^* will be called a **GMMOT plan**.*

Proposition 1 *The GMM-OT cost has an equivalent discrete formulation:*

$$MW_2^2(\mu_0, \mu_1) = \min_{w \in \Pi(\pi_0, \pi_1)} \sum_{k,l} w_{kl} W_2^2(\mu_0^k, \mu_1^l).$$

The solution will be denoted by w^*.

5) GMM Transport map

The GMMOT plan between μ_0 and μ_1 is

$$\gamma(x, y) = \sum_{k,l} w_{k,l}^* g_{m_0^k, \Sigma_0^k}(x) \delta_{y = T_{k,l}(x)}.$$

where $(w_{k,l}^*)$ is a discrete OT plan for $W^2(\pi_0, \pi_1)$ and $T_{k,l}$ is the OT map between μ_0^k and μ_1^k.

PROBLEM: The GMMOT plan is not of the form $(\mathrm{Id}, T)\#\mu_0$!
In practice, we use the **barycentric projection**

$$T(x) = \frac{\sum_{k,l} w_{k,l}^* g_{m_0^k, \Sigma_0^k}(x) T_{k,l}(x)}{\sum_k \pi_0^k g_{m_0^k, \Sigma_0^k}(x)}.$$

6) Construction of the TextoGMM Model

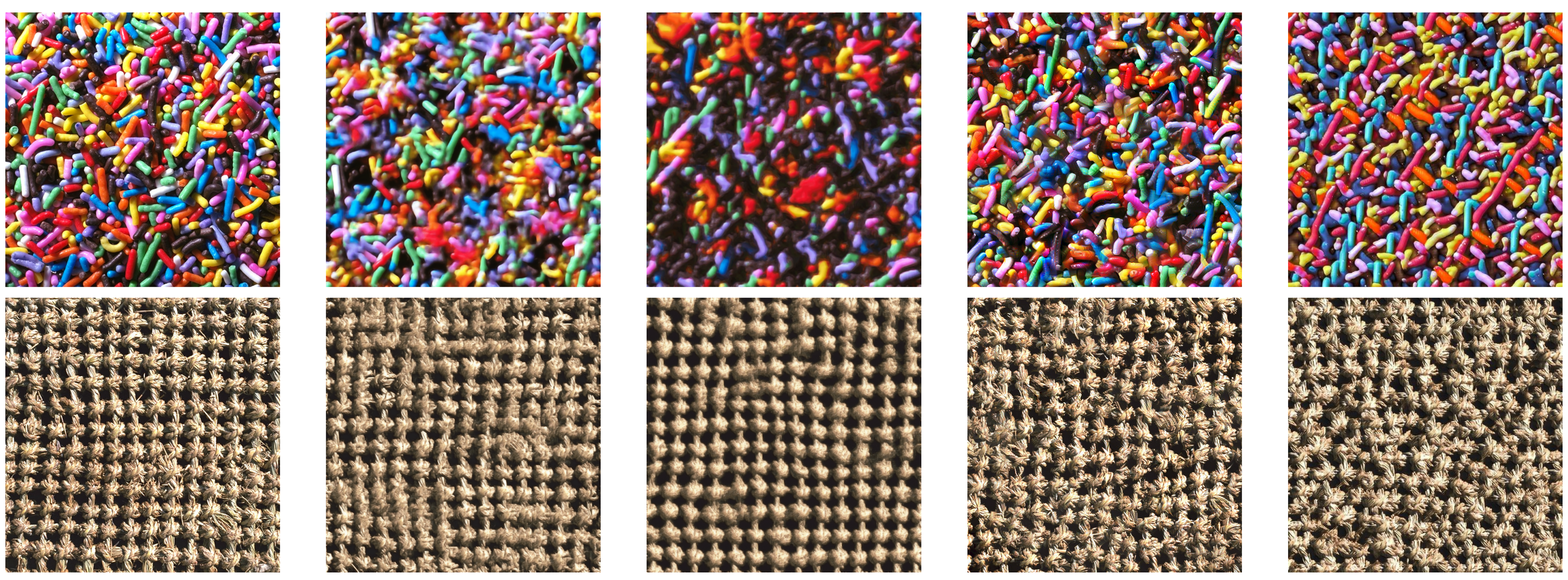
Initialize with the Gaussian model $U_{S-1} = \bar{u}_{S-1} + (u_{S-1} - \bar{u}_{S-1}) * Z$ (Z Gaussian white noise)

For $s = S-1, \ldots, 0$,

- Fit a GMM μ_s (resp. ν_s) to the patch distribution of U_s (resp. u_s)
- Compute the GMMOT plan γ_s from μ_s to ν_s, and associated map T_s
- Transport patches $T_s(U_{s|a+2^s\omega})$ and find nearest neighbor $u_{s|C_s(a)+2^s\omega}$ in u_s
- Recompose patches: $\forall a \in 2^s\mathbb{Z}^2, \quad V_s(a) = \frac{1}{|\omega|} \sum_{b \in 2^s\omega} u_s\big(C_s(a-b) + b\big)$
- Upsample: $\forall a \in 2^s\mathbb{Z}^2, \forall k \in \{0, 2^{s-1}\}^2, \quad U_{s-1}(a+k) = \frac{1}{|\omega|} \sum_{b \in 2^s\omega} u_{s-1}\big(C_s(a-b) + b + k\big)$

Output: Synthesized texture V_0

7) Results: Synthesis with TextoGMM



Original | TextoGMM $w = 5, S = 5$ | TextoML $w = 5, S = 5$ | [Gatys et al.] | [Ulyanov et al.]

Remarks - Conclusion

- One can estimate the model offline
- Patches are transformed independently
  → Parallel computations
- Costly steps: EM, NN projections
  → Random subsampling (10^4 patches)
- GMMOT compares favorably to other OT methods in terms of time. With patches:
  $1'$ to solve GMMOT (including EM step)
  $40'$ for 10^3 iterations of Sinkhorn
  $150'$ for 10^5 iterations of stochastic solver

Concluding remarks:

✓ TextoGMM can model structured textures.
✓ Model estimation is much faster than previous models based on semi-discrete OT.
✓ TextoGMM allows for mixing and inpainting.
✗ Still requires patch NN search/averaging...
✗ ... and the EM algorithm to learn GMM.

References

- Delon, Desolneux: A Wasserstein-type distance in the space of Gaussian mixture models. SIIMS, 2020.
- Leclaire, Rabin: A stochastic multi-layer algorithm for semi-discrete optimal transport with applications to texture synthesis and style transfer. JMIV, 2021.
- Gatys, Ecker, Bethge: Texture Synthesis Using Convolutional Neural Networks. NeurIPS'15.
- Ulyanov et al.: Texture networks: feed-forward synthesis of textures and stylized images. ICML'16.