école normale supérieure paris-saclay

Wasserstein Projection for texture synthesis

Numerical Image Project

BARBONI Raphaël - NAÏT S. Thiziri

Master 2 Student

MVA

ENS Paris-Saclay

February 7, 2021

Outline

1. Introduction
2. Projection in Wasserstein space
3. Mathematical insights of texture modeling
4. Steerable pyramid decomposition
5. Numerical implementation
6. Some results
7. Conclusion

Introduction

3 classes of Texture synthesis

- **Neighbored-oriented** methods
- **Example-oriented** methods
- **Statistical constraints** methods

Projection in Wasserstein space

Objective : define a projection operator

- Wasserstein distance

$$\mathcal{W}_2^2(X, Y) := \min_{\sigma \in \Sigma_N} \sum_{i,j \in I} ||X_i - Y_{\sigma(i)}||^2$$

- Wasserstein projection :

$$\sigma_* = \arg\min_{\sigma \in \Sigma_N} \mathcal{W}_\sigma^2(X, Y)$$

$\longrightarrow$ **1D case = histogram matching** = $\mathcal{O}(N \log N)$

Projection in Wasserstein space (2)

- Sliced Wasserstein distance

$$\tilde{\mathcal{W}}_2^2(X, Y) := \int_{\theta \in \Theta} \mathcal{W}(X_\theta, Y_\theta)^2 d\theta \qquad \text{where } X_\theta = \{< X_i, \theta >\}_{i \in I} \subset \mathbb{R}^N$$

- Sliced Wasserstein projection

$$X^\infty = \tilde{Proj}_{[Y]}(X)$$

  - Approximation through SGD
  - Non orthogonal projection
  - Low computation cost in high dimensional setting

Mathematics behind texture modeling

Definition (Texture)

A texture is a random field X over the discrete lattice $\mathbb{Z}^2$. Therefore, any realization of X has the same visual aspect for humans.

Theorem (Julesz conjecture (1962))

There exists a set of constraint functions $\{\phi_k\}_{1 \leqslant k \leqslant N_c}$ *such that any two realizations of any two random fields that are equal in expectations over these set of functions are visually indistinguishable for the human brain. In mathematical terms :*

$$\mathbb{E}\phi_k(X) = \mathbb{E}\phi_k(Y)\ \forall k \implies \textit{samples from } X \textit{ and } Y \textit{ are indistinguishable}$$

$\longrightarrow$ **successive statistical projection on the set of constraints.**

Steerable pyramid decomposition

Multi-resolution wavelet decomposition of a grayscale image $\Longrightarrow$ *tight frame*

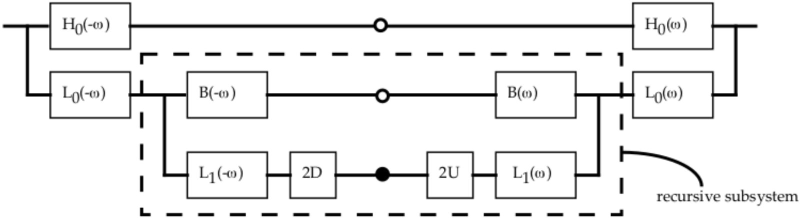

Figure: Steerable Pyramid

- Bandlimiting and essentially aliasing-free,
- Steerable (Kth order derivatives in (K+1) directions),
- Self-invertible (”tight-frame”).

Numerical Implementation : grayscale images

Algorithm 1 Grayscale texture synthesis

- Choose reference texture image X of size (N, N), order and height K, L of the decomposition, number of iterations Nit, wavelet basis w

- Define Decomposition(.) as $\texttt{SteerablePyramid}$ or $\texttt{WaveletDecomposition}$

$\mathbf{X} \leftarrow \texttt{Decomposition}(X)$

$Y \leftarrow Gaussian((N, N))$

for $0 \leqslant it < Nit$ **do**

$\quad \mathbf{Y} \leftarrow \texttt{Decomposition}(Y)$

$\quad$ **for** $0 \leqslant l < L, 0 \leqslant k < K$ **do**

$\quad\quad \mathbf{Y}[l, k] \leftarrow \texttt{HistogramMatching}(\mathbf{Y}[l, k], \mathbf{X}[k, l])$

$\quad Y \leftarrow \texttt{DecompositionRec}(\mathbf{Y})$

$\quad Y \leftarrow \texttt{HistogramMatching}(Y, X)$

Output: Y

Some results : grayscale images

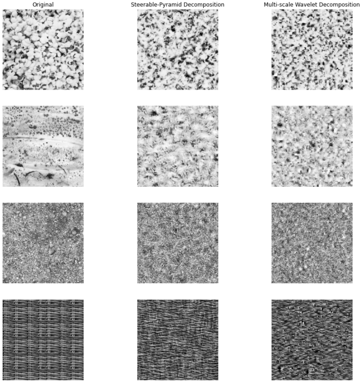

Figure: Results on grayscale images

Numerical implementation : RGB images (1)

- Treat color channels as decorrelated
  $\longrightarrow$ Apply previous algo on each channel
  $\longrightarrow$ Reconstruct the image with previous channels
  $\longrightarrow$ Mistaken assumption

- Treat color channels as correlated Briand et al., 2014
  $\longrightarrow$ Apply PCA to get a space where axis uncorrelated
  $\longrightarrow$ Apply previous algo on each new uncorrelated channel
  $\longrightarrow$ Reconstruct the image with previous channels

Numerical implementation : RGB images (2)

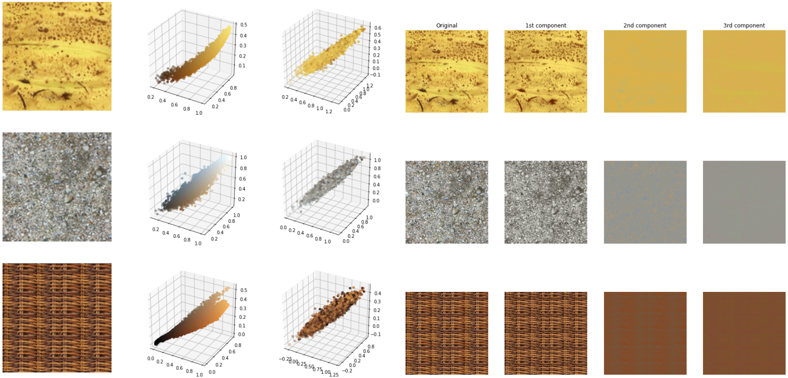

Figure: Correlation between channels when applying PCA

Numerical implementation : RGB images (3)

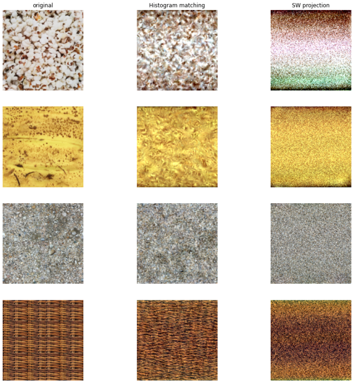

Figure: Results on RGB images using steerable-pyramid decomposition and histogram matching on each channel (center) or Sliced-Wasserstein projection (right)

Conclusion

- Steerable Pyramid : mother of Neural networks ?
- Alternative method : dictionary learning (see Tartavel et al. 2015)

References

[1] T. Briand, J. Vacher, B. Galerne, and J. Rabin, The Heeger and Bergen Pyramid Based Texture Synthesis Algorithm, Image Processing On Line, 4 (2014), pp. 276–299.

[2] E. P. Simoncelli and W. T. Freeman, The steerable pyramid: a flexible architecture for multi-scale derivative computation, in Proceedings., International Conference on Image Processing, vol. 3, 1995, pp. 444–447 vol.3.

[3] J. Rabin, G. Peyré, J. Delon, and M. Bernot, Wasserstein Barycenter and Its Application to Texture Mixing, in Scale Space and Variational Methods in Computer Vision.

Thank You

for Listening.