Stochastic super-resolution for Gaussian textures[1]

Emile Pierret[a], Bruno Galerne[a,b]

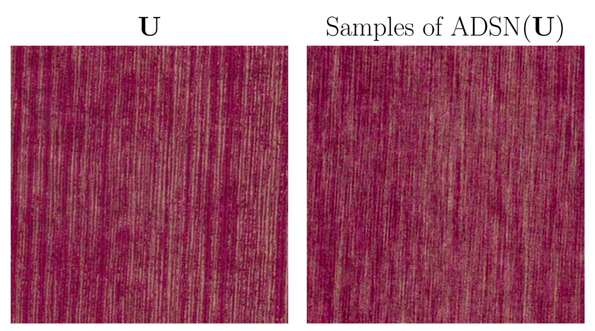
[a]Institut Denis Poisson – Université d'Orléans, Université de Tours, CNRS
[b]Institut universitaire de France (IUF)

The Asymptotic Discrete Spot Noise (ADSN) model[2]

Let $\Omega_{M,N} = [M] \times [N]$ and $\mathbf{U} \in \mathbb{R}^{\Omega_{M,N}}$ be a grayscale image, m its grayscale mean and $\mathbf{t} = \frac{1}{\sqrt{MN}}(\mathbf{U} - m)$ its associated texton. Let $\mathbf{W}$ be a white Gaussian noise,

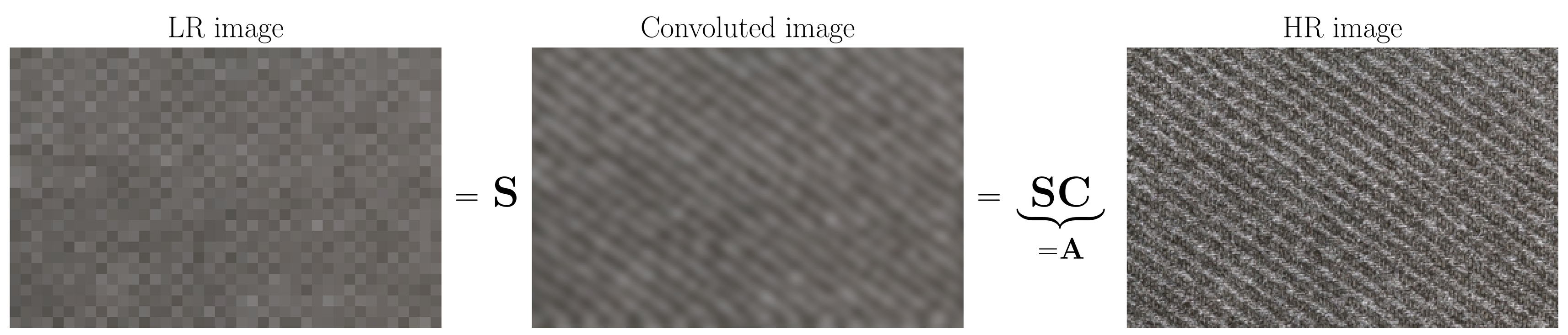
$$\mathbf{X} = \mathbf{t} \star \mathbf{W} \sim \mathrm{ADSN}(\mathbf{U}) = \mathscr{N}(\mathbf{0}, \mathbf{\Gamma})$$

which is a Gaussian **stationary** law.
$\mathbf{\Gamma}$ represents the convolution by the kernel $\boldsymbol{\gamma} = \mathbf{t} \star \check{\mathbf{t}}$.



The zoom-out operator A

Let $\mathbf{U}_{\mathrm{HR}}$ be an image of $\mathbb{R}^{\Omega_{M,N}}$, and r be an integer, we suppose that its LR version is obtained as $\mathbf{U}_{\mathrm{LR}} = \mathbf{A}\mathbf{U}_{\mathrm{HR}} \in \mathbb{R}^{\Omega_{M/r,N/r}}$ where $\mathbf{A} = \mathbf{SC}$ is a convolution $\mathbf{C}$ followed by a subsampling operator $\mathbf{S}$ with stride r.



Kriging reasoning for conditional sampling

For $\mathbf{U}_{\mathrm{HR}} \in \mathbb{R}^{\Omega_{M,N}}$, its associated model $\mathrm{ADSN}(\mathbf{U}) = \mathscr{N}(\mathbf{0}, \mathbf{\Gamma})$ and $\mathbf{U}_{\mathrm{LR}} = \mathbf{A}\mathbf{U}_{\mathrm{HR}}$, samples $\mathbf{X}_{\mathrm{SR}} \sim \mathrm{ADSN}(\mathbf{U})$ conditioned on $\mathbf{A}\mathbf{X}_{\mathrm{SR}} = \mathbf{U}_{\mathrm{LR}}$ have the form:

$$\mathbf{X}_{\mathrm{SR}} = \mathbf{\Lambda}^T \mathbf{U}_{\mathrm{LR}} + (\tilde{\mathbf{X}} - \mathbf{\Lambda}^T \mathbf{A} \tilde{\mathbf{X}})$$

with $\tilde{\mathbf{X}} \sim \mathrm{ADSN}(\mathbf{U})$ independent of $\mathbf{U}_{\mathrm{HR}}$
and $\mathbf{\Lambda} \in \mathbb{R}^{\Omega_{M/r,N/r} \times \Omega_{M,N}}$ verifying the **kriging equation**:

$$\mathbf{A\Gamma A}^T \mathbf{\Lambda} = \mathbf{A\Gamma}. \tag{1}$$

The convolutional form of the kriging equation

Lemma 1 (Convolution and subsampling)**.** $\mathbf{A\Gamma A}^T$ *is a convolution matrix with kernel* $\boldsymbol{\kappa} = \mathbf{S}(\mathbf{c} \star \boldsymbol{\gamma} \star \check{\mathbf{c}})$ *where* $\mathbf{c}$ *is the kernel of* $\mathbf{C}$*. Equation* (1) *becomes on each column of* $\mathbf{\Lambda}$*:*

$$\boldsymbol{\kappa} \star \boldsymbol{\lambda}(k,\ell) = \mathbf{A\Gamma}_{\Omega_{M,N} \times (k,\ell)}, \quad k,\ell \in [r] \tag{2}$$

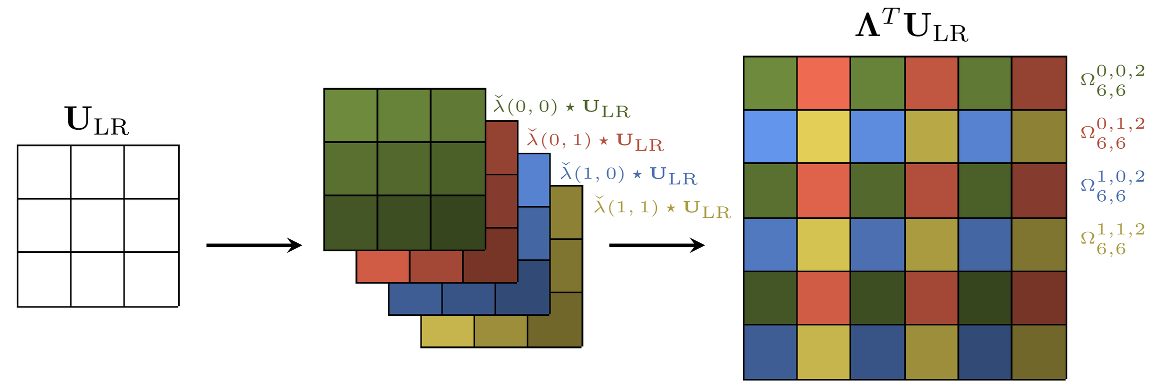
The structure of the kriging matrix

Let k, ℓ be integers in $[r]$, let $\Omega^{k,\ell,r}_{M,N} = \{(k+ir, \ell+jr),\ i,j \in [M/r] \times [N/r]\} \subset \Omega_{M,N}$ be the subgrid of $\Omega_{M,N}$ having stride r and starting at (k,ℓ).

Proposition 1 (Structure of the kriging matrix)**.** *There exists* $\mathbf{\Lambda} \in \mathbb{R}^{\Omega_{M/r,N/r} \times \Omega_{M,N}}$ *solution of Equation* (1) *such that* $\mathbf{Y} \in \mathbb{R}^{\Omega_{M/r,N/r}} \mapsto \mathbf{\Lambda}^T\mathbf{Y} \in \mathbb{R}^{\Omega_{M,N}}$ *corresponds to a convolution on each of the shifted subgrids* $\Omega^{k,\ell,r}_{M,N}$, $k,\ell \in [r]$*. More precisely,* $\mathbf{\Lambda}$ *is* ***fully determined by its*** $\mathbf{r^2}$ ***first columns*** $\boldsymbol{\lambda}(k,\ell) = \mathbf{\Lambda}_{\Omega_{M/r,N/r} \times (k,\ell)}$, $k,\ell \in [r]$ *and*

$$\left(\mathbf{\Lambda}^T\mathbf{Y}\right)(\Omega^{k,\ell,r}_{M,N}) = \check{\boldsymbol{\lambda}}(k,\ell) \star \mathbf{Y}.$$

Structure of $\mathbf{\Lambda}$ for $r = 2$ and $M = N = 6$.



Pseudo-code of Gaussian SR

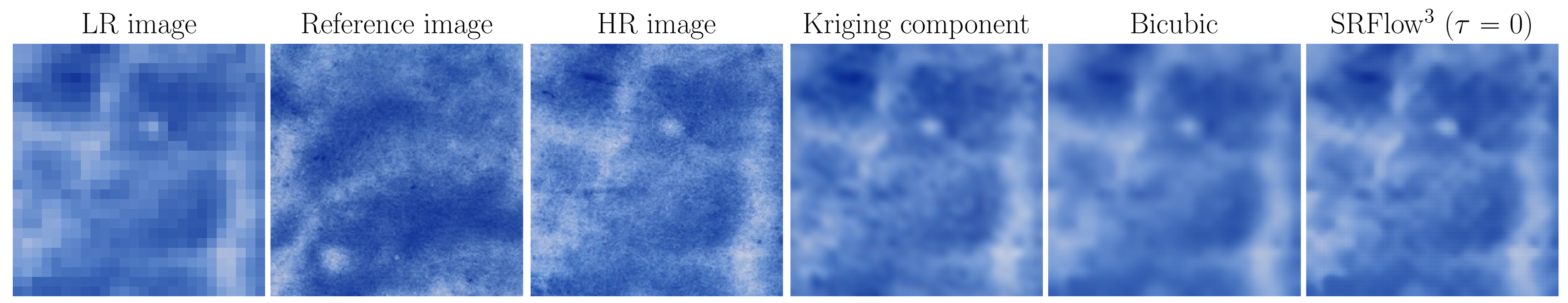
- Exact sampling using Gaussian conditional sampling.
- Efficient computations in Fourier exploiting the stationarity assumption and the form of the operator.
- The kriging matrix Λ could be stored to generate several samples.

Input: An image $\mathbf{U}_{\mathrm{LR}} \in \mathbb{R}^{\Omega_{M/r,N/r}}$, r the zoom factor, $\mathbf{t}$ the convolution kernel of the ADSN model, $\mathbf{c}$ the kernel of the convolution of the zoom-out operator $\mathbf{A} = \mathbf{SC}$

Step 1: *Computation of kriging matrix* $\mathbf{\Lambda}$
Store $\mathbf{per}(\mathbf{t})$ the periodic component of $\mathbf{t}$
Store the convolution kernels $\boldsymbol{\gamma} = \mathbf{per}(\mathbf{t}) \star \check{\mathbf{per}}(\mathbf{t})$, $\mathbf{c} \star \boldsymbol{\gamma}$ and $\boldsymbol{\kappa} = \mathbf{c} \star \boldsymbol{\gamma} \star \check{\mathbf{c}}$
for $(k,\ell) \in [r]^2$ **do**
  $\widehat{\mathbf{b}} = \mathscr{F}_2\Big(\mathbf{S}((\mathbf{c} \star \boldsymbol{\gamma})(\cdot - k, \cdot - \ell))\Big)$
  $\widehat{\boldsymbol{\lambda}}(k,\ell)\,[\widehat{\boldsymbol{\kappa}} \neq 0] \leftarrow \frac{\widehat{\mathbf{b}}[\widehat{\boldsymbol{\kappa}} \neq 0]}{\widehat{\boldsymbol{\kappa}}[\widehat{\boldsymbol{\kappa}} \neq 0]}$
end for
Step 2: *Sampling of one SR version of* $\mathbf{U}_{\mathrm{LR}}$
Generate $\mathbf{W} \in \mathbb{R}^{\Omega_{M,N}}$ following a Gaussian standard law
$\tilde{\mathbf{X}} \leftarrow \mathbf{t} \star \mathbf{W}$
$\tilde{\mathbf{X}}_{\mathrm{LR}} \leftarrow \mathbf{A}\tilde{\mathbf{X}}$
for each shifted subgrid by $(k,\ell) \in [r]^2$ **do**
  $\mathbf{X}_{\mathrm{SR}}(\Omega^{k,\ell,r}_{M,N}) \leftarrow \mathscr{F}_2^{-1}\Big((\widehat{\mathbf{U}}_{\mathrm{LR}} - \tilde{\mathbf{X}}_{\mathrm{LR}}) \odot \overline{\widehat{\boldsymbol{\lambda}}(k,\ell)}\Big) + \tilde{\mathbf{X}}(\Omega^{k,\ell,r}_{M,N})$
end for
Output: $\mathbf{X}_{\mathrm{SR}}$

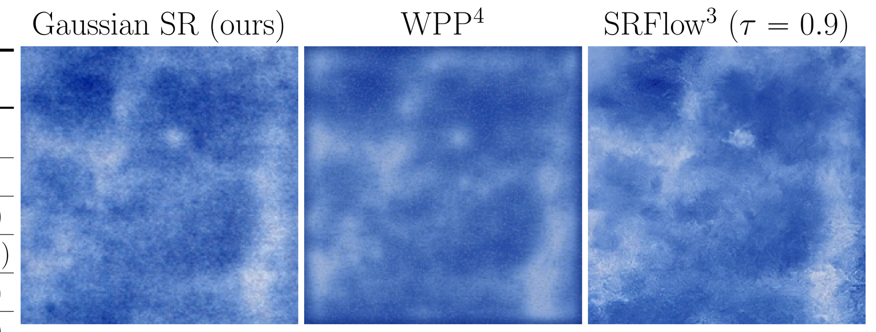
Comparaison with other methods

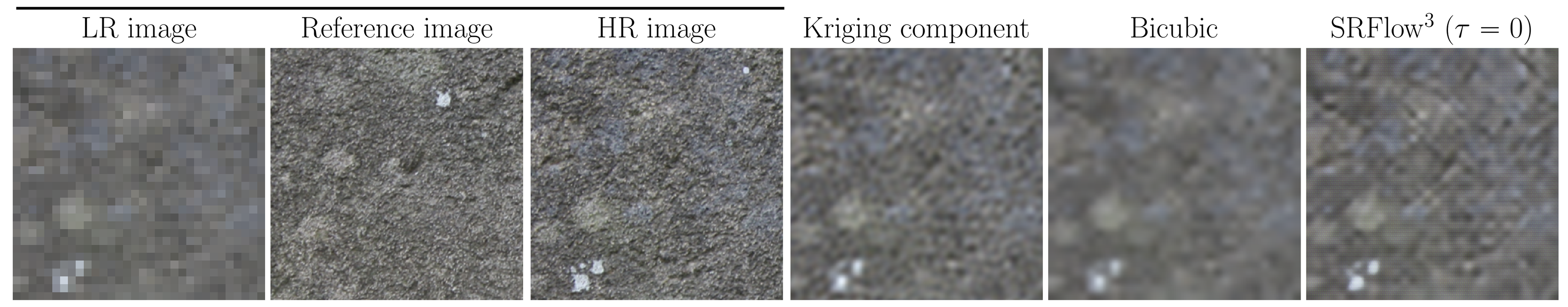
- The HR image size is respectively 208×208 and 256×256 and the factor $r = 8$.
- The ADSN model is extracted from a reference image.
- For stochastic Gaussian SR and SRFlow, the table has been realized on 200 samples.
- Our method outperforms in terms of the perceptual LPIPS metric and execution time.
- PSNR and SSIM are optimal for blurry images.



Evaluation metrics

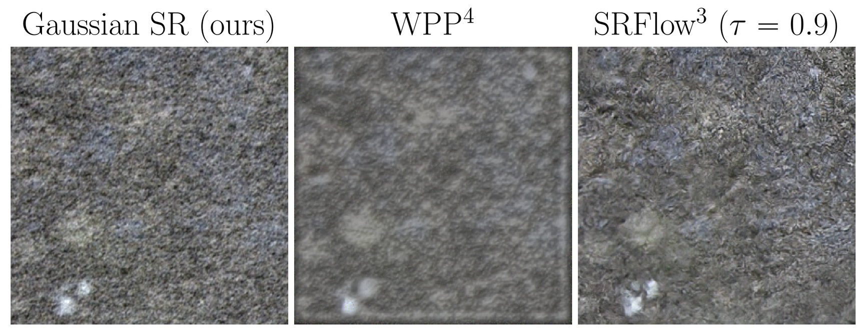
	PSNR (dB) ↑	SSIM ↑	LPIPS ↓	TIME (s)
Kriging component	28.42	0.56	0.69	
Bicubic	**30.21**	**0.65**	0.54	
SRFlow[3]($\tau = 0$)	28.55	0.54	0.63	0.47 (GPU)
Gaussian SR (ours)	26.25 ± 0.05	0.42 ± 0.00	**0.12 ± 0.01**	**0.01** (CPU)
WPP[4]	24.70	0.39	0.22	64.0 (GPU)
SRFlow[3]($\tau = 0.9$)	27.33 ± 0.34	0.48 ± 0.02	0.20 ± 0.03	0.47 (GPU)





Evaluation metrics

	PSNR (dB) ↑	SSIM ↑	LPIPS ↓	TIME (s)
Kriging component	21.78	0.24	0.87	
Bicubic	**23.52**	**0.45**	0.70	
SRFlow[3]($\tau = 0$)	21.84	0.24	0.87	0.55 (GPU)
Gaussian SR (ours)	18.99 ± 0.05	0.14 ± 0.01	**0.25 ± 0.01**	**0.02** (CPU)
WPP[4]	21.12	0.21	0.42	77.0 (GPU)
SRFlow[3]($\tau = 0.9$)	18.99 ± 0.38	0.14 ± 0.01	0.39 ± 0.04	0.55 (GPU)



Conclusion

- Super-resolution is performed in a well-based mathematical framework.
- An efficient sampler can be computed due to the stationarity assumption.
- The same routine could be used for other operators of the form convolution followed by subsampling.

References

[1]Pierret, É., & Galerne, B. (2023). Stochastic super-resolution for gaussian textures. *ICASSP 2023*
[2]Galerne, B., Gousseau, Y., & Morel, J.-M. (2011). Random Phase Textures: Theory and Synthesis. *IEEE Transactions on Image Processing*
[3]Lugmayr, A., Danelljan, M., Van Gool, L., & Timofte, R. (2020). SRFlow: Learning the Super-Resolution Space with Normalizing Flow. *ECCV 2020*
[4]Hertrich, J., Houdard, A., & Redenbach, C. (2022). Wasserstein Patch Prior for Image Superresolution. *IEEE Transactions on Computational Imaging*