Space-variant Generalized Gaussian Regularization for Image Restoration

Alessandro Lanza, Serena Morigi,
Monica Pragliola, Fiorella Sgallari

University of Bologna
Computational Methods for Inverse Problems in Imaging Workshop
July 16-18, 2018, Como

Outline

- Image restoration;
- Non convex regularization: the TV_p-ℓ_2 model;
- Novel contribution: the space variant $TV^{SV}_{p,\alpha}$-ℓ_q:
  - Motivation;
  - Main Idea.
- Numerical evidences;
- Conclusions and future work.

Image Restoration

Direct Problem

u	$\longrightarrow$	$u = \Phi(g)$	$\longrightarrow$	g
original image		blur and/or noise		observed image

Inverse Problem

g	$\longrightarrow$	$g = \Phi^{-1}(u)$	$\longrightarrow$	u
observed image		deblur and/or denoise		original image

Degradation model

Continuous formulation:

$$g(x) = (k * u)(x) + n(x) = \int_{\Omega} k(x-y)u(y)dy + n(x),\ x \in \Omega \subset \mathbb{R}^2,$$

where k is a known *space-invariant blur kernel*, n is the *noise* function whose distribution is known, u is the *original image* function and g is the *corrupted image* function.

Discrete formulation:

$$g = Ku + n,\ \ g, u, n \in \mathbb{R}^d,\ K \in \mathbb{R}^{d \times d},$$

where K is a large, sparse, *structured* matrix, whose structure depends on the boundary conditions. In general, K is very *ill conditioned*.

Variational Methods

$$u^* \longleftarrow \arg\min_{u \in \mathbb{R}^d} \left\{ F(u, g, K) + \mu R(u) \right\}$$

- F fidelity term, set according to the noise:

$$F(u, g, K) = \|Ku - g\|_q^q, \quad q = 1, 2;$$

$q = 2$
Gaussian noise

$q = 1$
Impulse noise

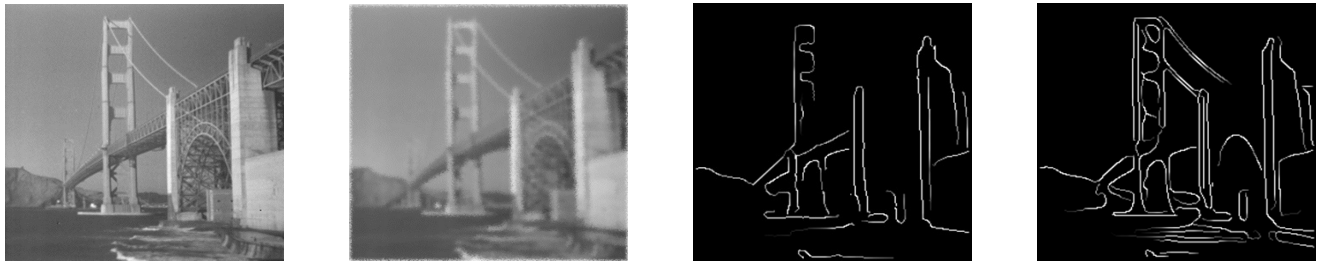
- $\mu > 0$ regularization parameter.

Variational Methods

$$u^* \longleftarrow \arg\min_{u \in \mathbb{R}^d} \left\{ F(u, g, K) + \mu R(u) \right\}$$

- R regularization term, set according to the properties of the image:

Regularizer	Tikhonov	Total Variation[1]
$R(u)$	$\sum_{i=1}^{d^2} \|(\nabla u)_i\|_2^2$	$\sum_{i=1}^{d^2} \|(\nabla u)_i\|_2$



(a) (b) (c) (d)

Figure: Original (a) and corrupted image (b), edges by Tikhonov (c) and TV (d).

[1] Rudin, L.I., Osher, S., Fatemi, E.:Nonlinear total variation based noise removal algorithms. Physica D: Nonlinear Phenomena, 1992.

The TV_p-ℓ_2 model[2]

$$u^* \longleftarrow \arg\min_{u\in\mathbb{R}^d} \left\{ \frac{\mu}{2}\|Ku-g\|_2^2 + \sum_{i=1}^{d} \|(\nabla u)_i\|_2^p \right\}$$

Regularization	$p > 1$	$p < 1$
Advantages	convexity	sparsity enhancing
Disadvantages	no sparsity enhancing	non convexity

[2] Lanza, A., Morigi, S., Sgallari, F.:Constrained $TV_p - \ell_2$ Model for Image Restoration. Journal of Scientific Computing, 2016.

Deriving the TV$_p$-ℓ_2 model

Let us introduce $\pi(u)$ as the prior probability density function (pdf), $\pi(g|u,K)$ as the likelihood pdf and $\pi(u|g,K)$ as the posterior pdf.

Maximum A Posteriori (MAP) Estimation

Resorting to the Bayes' formula, we have:

$$\max_{u\in\mathbb{R}^d} \pi(u|g,K) \Leftrightarrow \max_{u\in\mathbb{R}^d} log\pi(u|g,K)$$

$$\max_{u\in\mathbb{R}^d} log(\pi(u)\pi(g|u,K)) \Leftrightarrow \min_{u\in\mathbb{R}^d} \left\{ -log\pi(g|u,K) - log\pi(u) \right\}$$

Regularization term	$\longleftarrow$	Prior
Fidelity term	$\longleftarrow$	Likelihood

Gradient norm distribution

Prior

u is a Markov Random Field:

$$\pi(u) = \frac{1}{Z}\prod_{i=1}^{d}\exp\Big(-\alpha V_{c_i}(u)\Big) = \frac{1}{Z}\exp\Big(-\alpha\sum_{i=1}^{d}V_{c_i}(u)\Big),$$

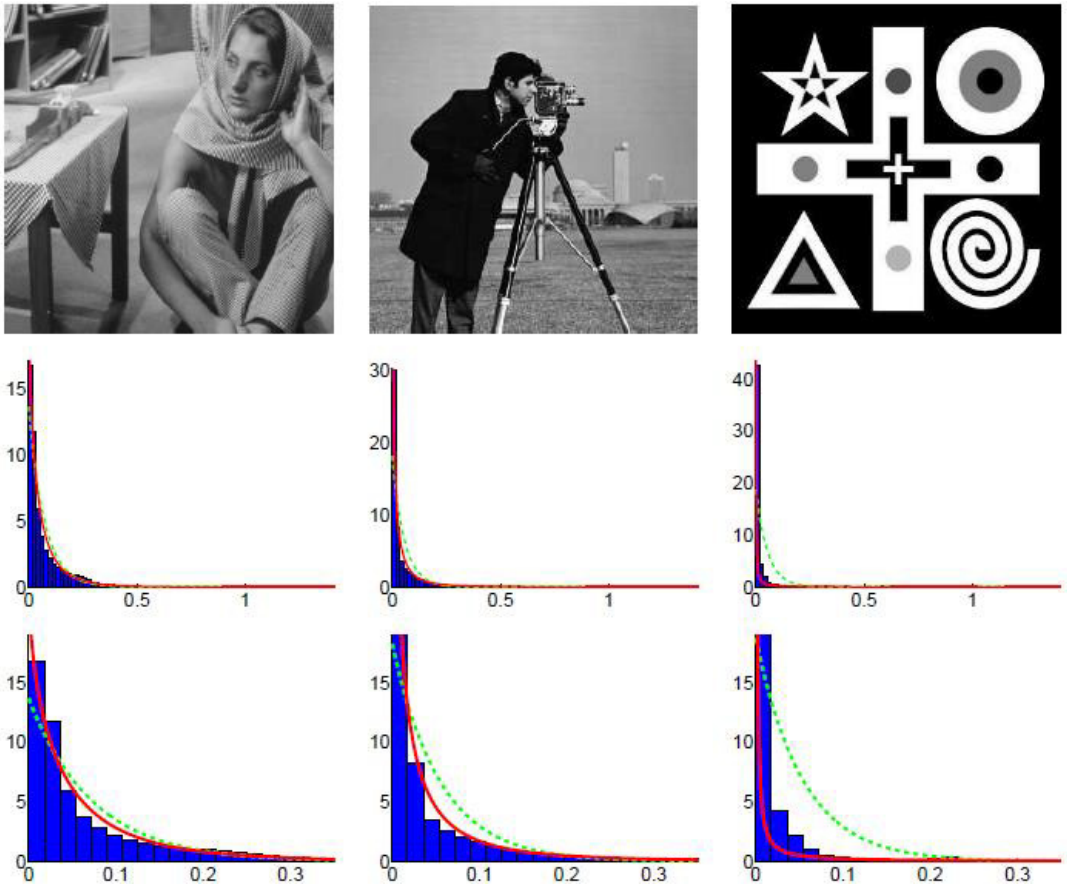
where V_{c_i} is a function depending only on the clique of neighbors of pixel i, $i = 1, ..., d$. Setting $V_{c_i} = \|(\nabla u)_i\|_2$, we have

$$\pi(u) = \frac{1}{Z}\exp\Big(-\alpha\sum_{i=1}^{d}\|(\nabla u)_i\|_2\Big) = \frac{1}{Z}\exp\Big(-\alpha \mathrm{TV}(u)\Big),$$

i.e. the norm of the gradients in each pixel are distributed according to an exponential distribution:

$$\pi(u;\alpha) = \begin{cases} \alpha e^{(-\alpha u)}, u \geq 0 \\ 0, u = 0. \end{cases}$$

Gradient norm distribution



Exponential

half Generalized Gaussian

$$\pi(u;\alpha) = \alpha e^{(-\alpha u)}, u > 0$$

$$\pi(u;\alpha,p) = \frac{\alpha p}{\Gamma(1/p)} e^{(-(\alpha u)^p)}, u > 0$$

The half Generalized Gaussian distribution

Hence, the prior turns into

$$\pi(u) = \frac{1}{Z} exp\Big(- \alpha \sum_{i=1}^{d} \|(\nabla u)_i\|_2^p \Big) = \frac{1}{Z} exp\Big(- \alpha \mathrm{TV}_p(u) \Big).$$

Likelihood (AWGN)

$$\pi(g|u;K) = \prod_{i=1}^{d} \frac{1}{\sqrt{2\pi}\sigma} \exp\Big(- \frac{(Ku-g)_i^2}{2\sigma^2} \Big) = \frac{1}{W} \exp\Big(- \frac{\|Ku-g\|_2^2}{2\sigma^2} \Big).$$

The half Generalized Gaussian distribution

Recalling the Bayes' formula,

$$\max_{u\in\mathbb{R}^d} \pi(u|g,K) \longrightarrow \min_{u\in\mathbb{R}^d} \left\{ -log\pi(g|u,K) - log\pi(u) \right\}$$

the following model is derived:

TV$_p$-ℓ_2

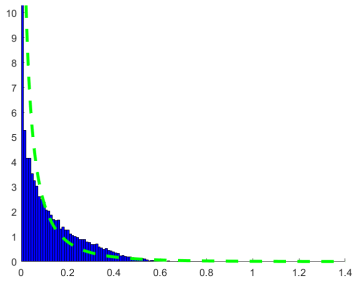
$$u^* \longleftarrow \arg\min_{u\in\mathbb{R}^d} \left\{ \frac{\mu}{2}\|Ku-g\|_2^2 + \sum_{i=1}^{d} \|(\nabla u)_i\|_2^p \right\}.$$

Note

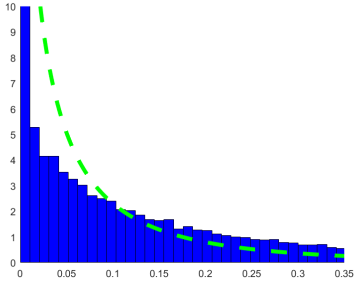
The shape parameter p of the hGGD is automatically estimated.

$\mathrm{TV}^{\mathrm{SV}}_{p,\alpha}$: Motivation

(a) test image

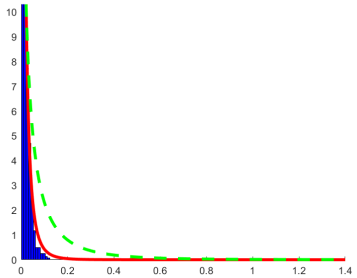


(b) global histogram

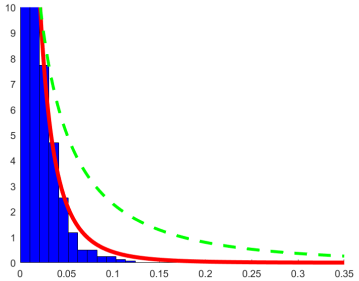


(c) zoom of (b)

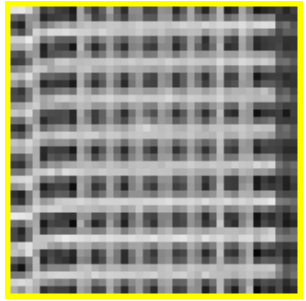
(d) smooth region



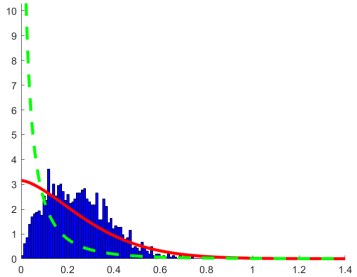
(e) local histogram for (d)



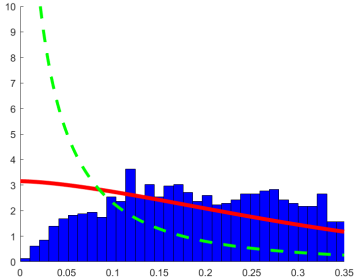
(f) zoom of (e)



(g) texture region



(h) local histogram for (g)



(i) zoom of (h)

$\mathrm{TV}^{\mathrm{SV}}_{p,\alpha}$: Main Idea

- Introduce a *space variant* prior;
- Estimate automatically the parameters characterizing the prior;
- Consider also AWGN and impulse noise, such as AWLN and Salt and Pepper Noise.

Likelihood (AWLN)

$$\pi(g|u;K) = \prod_{i=1}^{d} \frac{1}{2\beta} \exp\left(-\frac{|Ku-g|_i}{\beta} \right) = \frac{1}{W} \exp\left(-\frac{\|Ku-g\|_1}{\beta} \right).$$

Fidelity (SPN)

In order to promote sparsity of the noise, the ℓ_0 pseudo-norm of the residual $Ku - g$ is a popular choice. Nevertheless, in general the ℓ_1 is adopted:

$$\mathcal{F}(u,g,K) = \|Ku-g\|_1$$

$\mathrm{TV}^{\mathrm{SV}}_{p,\alpha}$: Main Idea

Non-stationary Markov Random Field Prior[3]

$$\pi(u) = \frac{1}{Z}\exp\left(- \sum_{i=1}^{d} \alpha_i \|(\nabla u)_i\|_2^{p_i} \right) = \frac{1}{Z}\exp\left(- \mathrm{TV}^{\mathrm{SV}}_{p,\alpha}(u) \right).$$

$\mathrm{TV}^{SV}_{p,\alpha}$-$\ell_q$, $q \in \{1,2\}$

$$u^* \longleftarrow \arg\min_{u\in\mathbb{R}^d} \left\{ \frac{\mu}{q}\|Ku - g\|_q^q + \sum_{i=1}^{d} \alpha_i \|(\nabla u)_i\|^{p_i} \right\}$$

[3] Lanza, A., Morigi, S., Pragliola, M., Sgallari, F.:Space-variant Generalized Gaussian Regularization for Image Restoration. Computer Methods in Biomechanics and Biomedical Engineering, 2018.

Automatic Parameter Estimation

The procedure adopted to estimate the global p is particularly successful when a large number of samples is available.

- $m_i = \|(\nabla u)_i\|_2,\ i = 1, ..., d$;
- N_i^s symmetric square neighborhood of pixel i of size $s \in \{3, 5, ...\}$.

We have [4]:

$$p_i \;=\; h^{-1}(\rho_i), \qquad \rho_i \;=\; \mathrm{card}\big(N_i^s\big) \left(\sum_{j \in N_i^s} m_j^2 \right) \Big/ \left(\sum_{j \in N_i^s} |m_j| \right)^2, \quad i = 1, ...d\,,$$

where

$$h(z) \;=\; \big(\Gamma(1/z)\ \Gamma(3/z)\big) \,/\, \big(\Gamma^2(2/z)\big)\,.$$

[4] Sharifi, K., Leon-Garcia, A.: Estimation of shape parameter for generalized Gaussian distributions in subband decompositions of video. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 5, Iss. 1, Feb 1995.

Automatic Parameter Estimation

Once p_i for a pixel is estimated, the local likelihood function is given by

$$\begin{aligned}\mathcal{L}(\alpha, p_i; x_1, ..., x_n) &= \prod_{i=1}^{n} \left(\frac{\alpha p_i}{\Gamma(1/p_i)} \right) \exp(-(\alpha x_i)^{p_i}) \\ &= \left(\frac{\alpha p_i}{\Gamma(1/p_i)} \right)^n \exp\left(-\sum_{i=1}^{n} (\alpha x_i)^{p_i} \right).\end{aligned}$$

The value of the local scale parameter is obtained by solving the following optimization problem:

$$\alpha_i = \arg\max_{\alpha} \log \mathcal{L}(\alpha, p_i; x_1, ..., x_n).$$

By imposing the first order optimality, we have:

$$\alpha_i = \left(\frac{p_i}{n} \sum_{i=1}^{n} x_i^{p_i} \right)^{-\frac{1}{p_i}}.$$

Alternating Directions Method of Multipliers[5]

Original problem:

$$u^* \leftarrow \min_u \{f_1(u) + f_2(Du)\}$$

Variable splitting:

$$\{u^*, z^*\} \leftarrow \min_{u,z} \{f_1(u) + f_2(z)\}, \ s.t. \quad z = Du$$

Augmented Lagrangian:

$$\mathcal{L}(u, z; \lambda) = f_1(u) + f_2(z) + \langle \lambda, z - Du \rangle + \frac{\beta}{2} \|z - Du\|_2^2$$

Saddle point problem: Find $(u^*, z^*; \lambda^*)$ such that:

$$\mathcal{L}(u^*, z^*; \lambda) \leq \mathcal{L}(u^*, z^*; \lambda^*) \leq \mathcal{L}(u, z; \lambda^*), \ \ \forall (u, z; \lambda).$$

Subproblems:

$$\begin{aligned} (u^{(k+1)}, z^{k+1}) &\leftarrow \arg\min_{u,z} \ \mathcal{L}(u, z; \lambda^{(k)}), \\ \lambda^{(k+1)} &\leftarrow \lambda^{(k)} - \beta_r \left(z^{(k+1)} - Du^{(k+1)} \right). \end{aligned}$$

[5] Boyd, S. et al.: Distributed optimization and statistical learning via the admm. 2011.

Alternating Directions Method of Multipliers

Constrained form:

$$\{u^*, r^*, t^*\} \leftarrow \min_{u,r,t} \left\{ (\mu/q)\|r\|_q^q + \sum_{i=1}^{d} \alpha_i \|t_i\|_2^{p_i} \right\}, \quad q = 1, 2$$
$$s.t. \quad r = Ku - g, \quad t = Du$$

Augmented Lagrangian:

$$\mathcal{L}(u, r, t; \lambda_r, \lambda_t) = \sum_{i=1}^{d} \alpha_i \|t_i\|_2^{p_i} + (\mu/q)\|r\|_q^q - \langle \lambda_t, t - Du \rangle + (\beta_t/2)\|t - Du\|_2^2 +$$
$$- \langle \lambda_r, r - (Ku - g) \rangle + (\beta_r/2)\|r - (Ku - g)\|_2^2$$

Equivalent problems: The restored image u^* is such that

$$(u^*, r^*, t^*; \lambda_r^*, \lambda_t^*) \quad \longleftarrow \quad \overset{\textit{dual ascent}}{\arg\max_{\lambda_r, \lambda_t}} \quad \overset{\textit{primal descent}}{\min_{u,r,t}} \quad \mathcal{L}(u, r, t; \lambda_r, \lambda_t)$$

Subproblems for ADMM

One variable at a time is updated, while the other ones are fixed as in the previous iteration step.

$$
\begin{aligned}
r^{(k+1)} &\leftarrow \arg\min_{r\in V} \mathcal{L}(u^{(k)}, r, t^{(k)}; \lambda_r^{(k)}, \lambda_t^{(k)}), && \text{closed form} \\
t^{(k+1)} &\leftarrow \arg\min_{t\in Q} \mathcal{L}(u^{(k)}, r^{(k+1)}, t; \lambda_r^{(k)}, \lambda_t^{(k)}), && \text{closed form} \\
u^{(k+1)} &\leftarrow \arg\min_{u\in V} \mathcal{L}(u, r^{(k+1)}, t^{(k+1)}; \lambda_r^{(k)}, \lambda_t^{(k)}), && \text{linear system} \\
\lambda_r^{(k+1)} &\leftarrow \lambda_r^{(k)} - \beta_r \left(r^{(k+1)} - (Ku^{(k+1)} - g) \right), \\
\lambda_t^{(k+1)} &\leftarrow \lambda_t^{(k)} - \beta_t \left(t^{(k+1)} - Du^{(k+1)} \right).
\end{aligned}
$$

Subproblems for ADMM

$$u^{(k+1)} \quad \leftarrow \quad \arg\min_{u \in V} \mathcal{L}(u, r^{(k+1)}, t^{(k+1)}; \lambda_r^{(k)}, \lambda_t^{(k)}), \quad \text{linear system}$$

$$\left(D^T D + \frac{\beta_r}{\beta_t} K^T K\right) u = D^T \left(t^{(k+1)} - \frac{1}{\beta_t}\lambda_t^{(k)}\right) + \frac{\beta_r}{\beta_t} K^T \left(r^{(k+1)} - \frac{1}{\beta_r}\lambda_r^{(k)} + g\right).$$

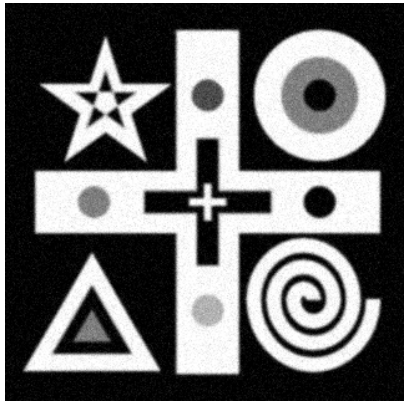
The coefficient matrix has full rank, in fact:

$$\mathrm{Ker}\left(D^T D\right) \cap \mathrm{Ker}\left(K^T K\right) = \{0\},$$

Periodic BC	$\longrightarrow$	DFT
Reflective BC	$\longrightarrow$	DCT
Anti-Reflective BC	$\longrightarrow$	DST

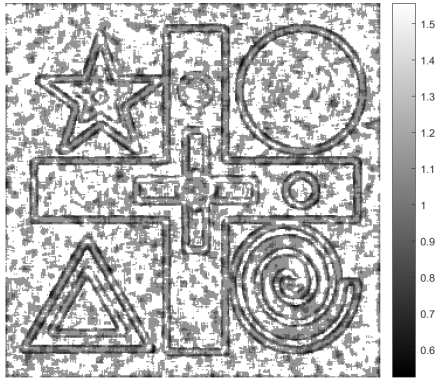
Numerical tests: AWGN

(a) original

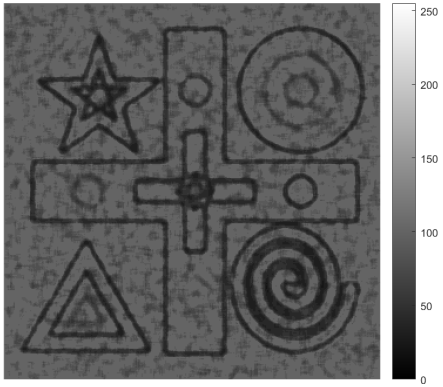


(b) corrupted (BSNR = 20)

(c) $\mathrm{TV}^{\mathrm{sv}}_{p,\alpha}$-$\mathrm{L}_2$



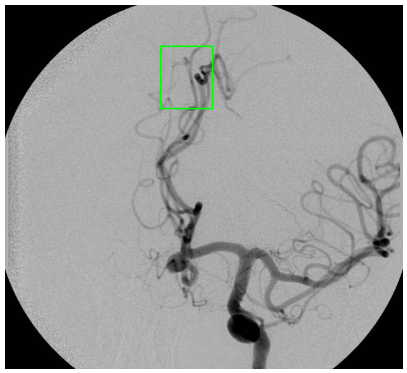
(a) p-map



(b) α-map

geometric	TV-L_2	TV$_p$-L_2	$\mathrm{TV}^{\mathrm{sv}}_{p,\alpha}$-$\mathrm{L}_2$
ISNR	7.77	7.92	8.60

Numerical tests: SPN



(a) original

(b) TV-L_1 (18.55)

(c) TV_p-L_1 (19.10)

(d) TV_p^{sv}-L_1 (21.14)

(d) corrupted

(e) zoom of (b)

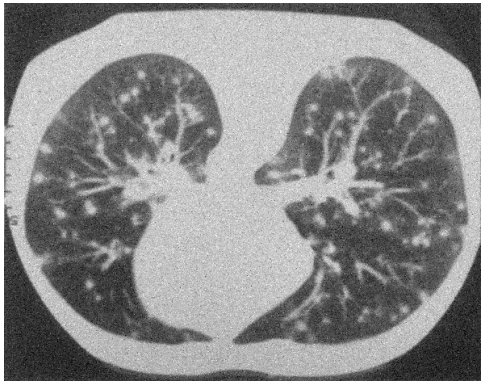
(f) zoom of (c)

(g) zoom of (d)

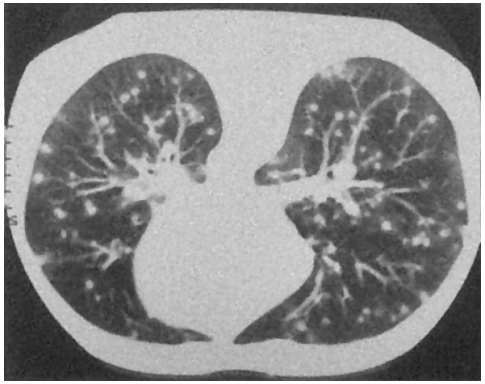
Figure: Example (SPN): restoration (ISNR in brackets) of the test image `aneurism` (https://medpix.nlm.nih.gov) corrupted by a $\gamma = 0.1$ level noise.

Numerical tests: AWLN

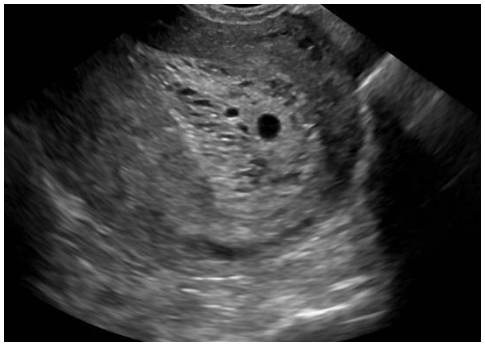
(a) original



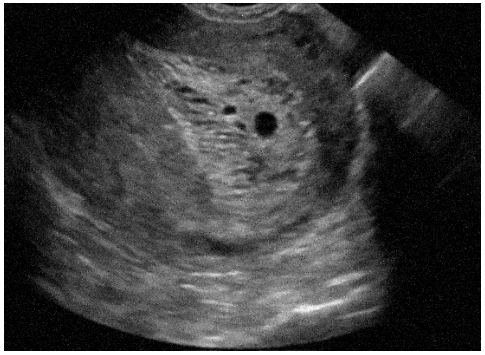
(b) corrupted



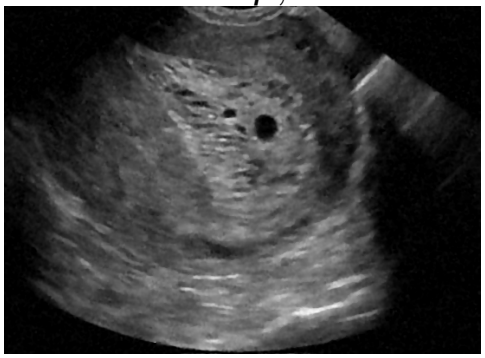
(c) $\mathrm{TV}^{\mathrm{sv}}_{p,\alpha}$-$\mathrm{L}_1$



(d) original



(e) corrupted (BSNR = 10)



(f) $\mathrm{TV}^{\mathrm{sv}}_{p,\alpha}$-$\mathrm{L}_1$

	lungs			**ecography**		
	TV-L_2	TV$_p$-L_2	$\mathrm{TV}^{\mathrm{sv}}_{p,\alpha}$-$\mathrm{L}_2$	TV-L_2	TV$_p$-L_2	$\mathrm{TV}^{\mathrm{sv}}_{p,\alpha}$-$\mathrm{L}_2$
ISNR	6.20	6.80	8.32	5.93	6.40	8.32

Conclusions and future work

- We proposed a local version of the TV_p-ℓ_2 model with an automatic procedure for the estimation of the local parameters. Furthermore, we introduced the $TV^{SV}_{p,\alpha}$-ℓ_1 model, more suitable when dealing with impulse noise.

As a future work, we are going to:

- Substitute the norm of the gradient in each pixel with the gradient itself, resorting to a Multivariate Generalized Gaussian distribution, in order to take into account in the restoration process the orientation of textures and edges.
- Adopt other algorithms for solving the optimization probem (e.g. Majorize-Minimize).

Thank you for your attention!