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 $TV_p-\ell_2$ model;
- Novel contribution: the space variant $TV_{p,\alpha}^{SV}-\ell_q$:
  - Motivation;
  - Main Idea.
- Numerical evidences;
- Conclusions and future work.
Image Restoration

Direct Problem

\[ u \rightarrow u = \Phi(g) \rightarrow g \]
original image  blur and/or noise  observed image

Inverse Problem

\[ g \rightarrow g = \Phi^{-1}(u) \rightarrow u \]
observed image  deblur and/or denoise  original image
Degradation model

Continuous formulation:

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

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:

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

where \( K \) is a large, sparse, structured matrix, whose structure depends on the boundary conditions. In general, \( K \) is very ill conditioned.
Variational Methods

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

- \( F \) 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

\[ u^* \leftarrow \arg \min_{u \in \mathbb{R}^d} \{ F(u, g, K) + \mu R(u) \} \]

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

<table>
<thead>
<tr>
<th>Regularizer</th>
<th>Tikhonov</th>
<th>Total Variation(^1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R(u) )</td>
<td>[ \sum_{i=1}^{d^2} | \nabla u_i |_2^2 ]</td>
<td>[ \sum_{i=1}^{d^2} | \nabla u_i |_2 ]</td>
</tr>
</tbody>
</table>

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

The $\text{TV}_{p-\ell_2}$ model\textsuperscript{2}

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

<table>
<thead>
<tr>
<th>Regularization</th>
<th>$p &gt; 1$</th>
<th>$p &lt; 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advantages</td>
<td>convexity</td>
<td>sparsity enhancing</td>
</tr>
<tr>
<td>Disadvantages</td>
<td>no sparsity enhancing</td>
<td>non convexity</td>
</tr>
</tbody>
</table>

Deriving the TV\textsubscript{$p\,-\,\ell_2$} model

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

**Maximum A Posteriori (MAP) Estimation**

Resorting to the Bayes’ formula, we have:

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

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

| Regularization term | ←— Prior |
| Fidelity term       | ←— Likelihood |
**Gradient norm distribution**

**Prior**

*u* is a Markov Random Field:

\[
\pi(u) = \frac{1}{Z} \prod_{i=1}^{d} \exp \left( -\alpha V_{c_{i}}(u) \right) = \frac{1}{Z} \exp \left( -\alpha \sum_{i=1}^{d} V_{c_{i}}(u) \right),
\]

where \( V_{c_{i}} \) is a function depending only on the clique of neighbors of pixel \( i \), \( i = 1, \ldots, d \). Setting \( V_{c_{i}} = \| (\nabla u)_{i} \|_{2} \), we have

\[
\pi(u) = \frac{1}{Z} \exp \left( -\alpha \sum_{i=1}^{d} \| (\nabla u)_{i} \|_{2} \right) = \frac{1}{Z} \exp \left( -\alpha TV(u) \right),
\]

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

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

half Generalized Gaussian

\[ \pi(u; \alpha, p) = \frac{\alpha p}{\Gamma(1/p)} e^{-(\alpha u)^p}, \quad u > 0 \]
The half Generalized Gaussian distribution

Hence, the prior turns into

\[
\pi(u) = \frac{1}{Z} \exp \left( -\alpha \sum_{i=1}^{d} \| (\nabla u)_i \|_2^p \right) = \frac{1}{Z} \exp \left( -\alpha \text{TV}_p(u) \right).
\]

Likelihood (AWGN)

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

\[
TV_{p-\ell_2}
\]

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

Note

The shape parameter \( p \) of the hGGD is automatically estimated.
$TV_{p,\alpha}^{SV}$: 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)
**TV^{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 $\ell_0$ pseudo-norm of the residual $Ku - g$ is a popular choice. Nevertheless, in general the $\ell_1$ is adopted:

\[
\mathcal{F}(u, g, K) = \|Ku - g\|_1
\]
**TV\(^{SV}_{p,\alpha}\): Main Idea**

Non-stationary Markov Random Field Prior\(^3\)

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

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

\[
u^* \leftarrow \arg \min_{u \in \mathbb{R}^d} \left\{ \frac{\mu}{q} \| Ku - g \|_q^q + \sum_{i=1}^{d} \alpha_i \| (\nabla u)_i \|_2^{p_i} \right\}
\]

---

Automatic Parameter Estimation

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

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

We have:\textsuperscript{4}

$$p_i = h^{-1}(\rho_i), \quad \rho_i = \text{card}(N_i^s) \left( \sum_{j \in N_i^s} m_j^2 \right) / \left( \sum_{j \in N_i^s} |m_j| \right)^2, \quad i = 1, \ldots, d,$$

where

$$h(z) = \left( \Gamma(1/z) \Gamma(3/z) \right) / \left( \Gamma^2(2/z) \right).$$

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

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

$$
= \left( \frac{\alpha p_i}{\Gamma(1/p_i)} \right)^n \exp \left( - \sum_{i=1}^{n} (\alpha x_i)^{p_i} \right).
$$

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

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

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\textsuperscript{5}

Original problem:

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

Variable splitting:

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

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:

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

Subproblems:

\[
(u^{(k+1)}, z^{k+1}) \leftarrow \arg \min_{u,z} \mathcal{L}(u, z; \lambda^{(k)}), \\
\lambda^{(k+1)} \leftarrow \lambda^{(k)} - \beta_r \left( z^{(k+1)} - Du^{(k+1)} \right). \]

\textsuperscript{5}Boyd, S. et al.: Distributed optimization and statistical learning via the admm. 2011.
Alternating Directions Method of Multipliers

Constrained form:

\[
\{u^*, r^*, t^*\} \leftarrow \min_{u,r,t} \left\{ (\mu/q)\|r\|_q^q + \sum_{i=1}^{d} \alpha_i \|t_i\|_p^p \right\}, \quad q = 1, 2
\]

s.t. \quad r = Ku - g, \quad t = Du

Augmented Lagrangian:

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

\[
- \langle \lambda_r, r - (Ku - g) \rangle + (\beta_r/2)\|r - (Ku - g)\|_2^2
\]

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

\[
(u^*, r^*, t^*; \lambda_r^*, \lambda_t^*) \leftarrow \arg \max_{\lambda_r, \lambda_t} \quad \text{dual ascent}
\]

\[
\min_{u,r,t} \quad \text{primal descent}
\]

\[
\mathcal{L}(u, r, t; \lambda_r, \lambda_t)
\]
Subproblems for ADMM

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

\[ r^{(k+1)} \leftarrow \arg \min_{r \in V} \mathcal{L}(u^{(k)}, r, t^{(k)}; \lambda_r^{(k)}, \lambda_t^{(k)}), \quad \text{closed form} \]

\[ t^{(k+1)} \leftarrow \arg \min_{t \in Q} \mathcal{L}(u^{(k)}, r^{(k+1)}, t; \lambda_r^{(k)}, \lambda_t^{(k)}), \quad \text{closed form} \]

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

\[ \lambda_r^{(k+1)} \leftarrow \lambda_r^{(k)} - \beta_r \left( r^{(k+1)} - (Ku^{(k+1)} - g) \right), \]

\[ \lambda_t^{(k+1)} \leftarrow \lambda_t^{(k)} - \beta_t \left( t^{(k+1)} - Du^{(k+1)} \right). \]
Subproblems for ADMM

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

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

The coefficient matrix has full rank, in fact:

\[
\ker (D^T D) \cap \ker (K^T K) = \{0\} ,
\]

<table>
<thead>
<tr>
<th>Boundary Condition</th>
<th>Transformation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Periodic BC</td>
<td>→ DFT</td>
</tr>
<tr>
<td>Reflective BC</td>
<td>→ DCT</td>
</tr>
<tr>
<td>Anti-Reflective BC</td>
<td>→ DST</td>
</tr>
</tbody>
</table>
Numerical tests: AWGN

(a) original  (b) corrupted (BSNR = 20)  (c) $TV_{p,\alpha}^S-L_2$

(a) $p$-map  (b) $\alpha$-map

<table>
<thead>
<tr>
<th>geometric</th>
<th>TV-L2</th>
<th>TV$_{p}$-L2</th>
<th>TV$_{p,\alpha}^S$-L2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISNR</td>
<td>7.77</td>
<td>7.92</td>
<td>8.60</td>
</tr>
</tbody>
</table>
Numerical tests: SPN

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) $TV_{p,\alpha}^S$ - $L_1$

(d) original (e) corrupted (BSNR = 10) (f) $TV_{p,\alpha}^S$ - $L_1$

<table>
<thead>
<tr>
<th>lungs</th>
<th>TV-$L_2$</th>
<th>TV$_p$-$L_2$</th>
<th>TV$_{p,\alpha}^S$-$L_2$</th>
<th>TV-$L_2$</th>
<th>TV$_p$-$L_2$</th>
<th>TV$_{p,\alpha}^S$-$L_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISNR</td>
<td>6.20</td>
<td>6.80</td>
<td>8.32</td>
<td>5.93</td>
<td>6.40</td>
<td>8.32</td>
</tr>
</tbody>
</table>
Conclusions and future work

- We proposed a local version of the $TV_p-\ell_2$ model with an automatic procedure for the estimation of the local parameters. Furthermore, we introduced the $TV_{p,\alpha}^S\ell_1$ 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 problem (e.g. Majorize-Minimize).
Thank you for your attention!