Novel Algorithms for Vectorial Total Variation

SIAM Imaging Sciences, Albuquerque, May 2016

D. Cremers[1], B. Goldlücke[2], E. Strekalovskiy[3],

J. Duran[1,4], T. Möllenhoff[1], and M. Moeller[1]

[1]Dept. Math & Computer Science
Technical University of Munich
Garching, Germany

[2]Dept. Computer & Information Sci.
University of Konstanz
Konstanz, Germany

[3]Google
Munich, Germany

[4]Dept. Math & Computer Science
University of Balearic Islands
Palma de Mallorca, Spain

Inverse Problems and Total Variation

Consider the inverse problem

$$\min_{u \in \mathsf{BV}(\Omega;\mathbb{R}^C)} \mathsf{J}(u) + \frac{\lambda}{2}\|Au - f\|_2^2,$$

with a noisy input image $f \in L^2(\Omega, \mathbb{R}^C), \Omega \subset \mathbb{R}^M$ and a linear operator A. We focus on the design of an effective regularizer $\mathsf{J}(u)$.

Inverse Problems and Total Variation

Consider the inverse problem

$$\min_{u \in \mathsf{BV}(\Omega;\mathbb{R}^C)} \mathsf{J}(u) + \frac{\lambda}{2}\|Au - f\|_2^2,$$

with a noisy input image $f \in L^2(\Omega, \mathbb{R}^C), \Omega \subset \mathbb{R}^M$ and a linear operator A. We focus on the design of an effective regularizer $\mathsf{J}(u)$.

In the scalar-valued setting ($C = 1$), a popular convex regularizer is the total variation [Herve, Shulman '89, Rudin, Osher, Fatemi '92]:

$$J(u) = TV(u) = \int_\Omega \|\nabla u(x)\|_2 \, \mathrm{d}x = \sup_{|\xi| \leq 1} \int_\Omega u \, \mathrm{div}\, \xi \, \mathrm{d}x.$$

Inverse Problems and Total Variation

Consider the inverse problem

$$\min_{u \in \mathsf{BV}(\Omega;\mathbb{R}^C)} \mathsf{J}(u) + \frac{\lambda}{2}\|Au - f\|_2^2,$$

with a noisy input image $f \in L^2(\Omega, \mathbb{R}^C), \Omega \subset \mathbb{R}^M$ and a linear operator A. We focus on the design of an effective regularizer $\mathsf{J}(u)$.

In the scalar-valued setting ($C = 1$), a popular convex regularizer is the total variation [Herve, Shulman '89, Rudin, Osher, Fatemi '92]:

$$J(u) = TV(u) = \int_\Omega \|\nabla u(x)\|_2 \,\mathrm{d}x = \sup_{|\xi| \le 1} \int_\Omega u \operatorname{div} \xi \,\mathrm{d}x.$$

How can we generalize $TV(u)$ to vector-valued images ($C > 1$)?

Vectorial Total Variation

- Channelwise summation [Blomgren, Chan '98]:

$$TV_S(u) := \sum_{i=1}^{C} TV(u_i) \quad = \quad \sup_{\xi:\Omega\to(\mathbb{E}^d)^C} \sum_{i=1}^{C} \int_\Omega u_i \,\mathrm{div}\,\xi_i \mathrm{d}x$$

Vectorial Total Variation

- Channelwise summation [Blomgren, Chan '98]:

$$TV_S(u) := \sum_{i=1}^{C} TV(u_i) \quad = \quad \sup_{\xi:\Omega\to(\mathbb{E}^d)^C} \sum_{i=1}^{C} \int_\Omega u_i \,\mathrm{div}\, \xi_i \mathrm{d}x$$

- Global channel coupling [Sapiro, Ringach '96]:

$$TV_F(u) := \int_\Omega \|\nabla u\|_F \mathrm{d}x \quad = \quad \sup_{\xi:\Omega\to\mathbb{E}^{d\times C}} \sum_{i=1}^{C} \int_\Omega u_i \,\mathrm{div}\, \xi_i \mathrm{d}x$$

Vectorial Total Variation

- Channelwise summation [Blomgren, Chan '98]:

$$TV_S(u) := \sum_{i=1}^{C} TV(u_i) \quad = \quad \sup_{\xi:\Omega\to(\mathbb{E}^d)^C} \sum_{i=1}^{C} \int_\Omega u_i \,\operatorname{div}\xi_i \mathrm{d}x$$

- Global channel coupling [Sapiro, Ringach '96]:

$$TV_F(u) := \int_\Omega \|\nabla u\|_F \mathrm{d}x \quad = \quad \sup_{\xi:\Omega\to\mathbb{E}^{d\times C}} \sum_{i=1}^{C} \int_\Omega u_i \,\operatorname{div}\xi_i \mathrm{d}x$$

- Spectral norm coupling [Goldlücke et al. '12]:

$$TV_J(u) := \int_\Omega \|\nabla u\|_{\sigma_1} \mathrm{d}x = \sup_{\xi:\Omega\to\mathbb{E}^d, \eta:\Omega\to\mathbb{E}^C} \sum_{i=1}^{C} \int_\Omega u_i \,\operatorname{div}(\eta_i \xi)\mathrm{d}x$$

Mixed Matrix Norms for Vectorial Total Variation

Represent an image u with N pixels and C colors by the matrix:

$$\mathbf{u} = (u_1, \ldots, u_c) \in \mathbb{R}^{N \times C} \text{ s.t. } u_k \in \mathbb{R}^N, \ \forall k \in \{1, \ldots, C\}.$$

Mixed Matrix Norms for Vectorial Total Variation

Represent an image u with N pixels and C colors by the matrix:

$$\mathbf{u} = (u_1, \ldots, u_c) \in \mathbb{R}^{N \times C} \text{ s.t. } u_k \in \mathbb{R}^N, \;\; \forall k \in \{1, \ldots, C\}.$$

The Jacobi matrix at each pixel defines a 3D tensor given by

$$K\mathbf{u} \equiv (Ku)_{i,j,k} \in \mathbb{R}^{N \times M \times C}.$$

Mixed Matrix Norms for Vectorial Total Variation

Represent an image u with N pixels and C colors by the matrix:

$$\mathbf{u} = (u_1, \dots, u_c) \in \mathbb{R}^{N \times C} \text{ s.t. } u_k \in \mathbb{R}^N, \ \forall k \in \{1, \dots, C\}.$$

The Jacobi matrix at each pixel defines a 3D tensor given by

$$K\mathbf{u} \equiv (Ku)_{i,j,k} \in \mathbb{R}^{N \times M \times C}.$$

Definition

For $A \in \mathbb{R}^{N \times M \times C}$, the mixed matrix $\ell^{p,q,r}$ norm is defined as

$$\|A\|_{p,q,r} = \left(\sum_{i=1}^{N} \left(\sum_{j=1}^{M} \left(\sum_{k=1}^{C} |A_{i,j,k}|^p \right)^{q/p} \right)^{r/q} \right)^{1/r}.$$

Mixed Matrix Norms for Vectorial Total Variation

Schatten p-norms penalize the singular values of a given matrix with an ℓ^p-norm.

Mixed Matrix Norms for Vectorial Total Variation

Schatten p-norms penalize the singular values of a given matrix with an ℓ^p-norm.

For $p = 1$, we get the nuclear norm, a convex relaxation of the rank. For $p = 2$, we get the Frobenius norm. And for $p = \infty$, we penalize the largest singular value.

Mixed Matrix Norms for Vectorial Total Variation

Schatten p-norms penalize the singular values of a given matrix with an ℓ^p-norm.

For $p = 1$, we get the nuclear norm, a convex relaxation of the rank. For $p = 2$, we get the Frobenius norm. And for $p = \infty$, we penalize the largest singular value.

Definition

For a tensor $A \in \mathbb{R}^{N \times M \times C}$, the mixed matrix Schatten (S^p, ℓ^q) norm is defined as

$$(S^p, \ell^q)(A) = \left(\sum_{i=1}^{N} \left\| \begin{pmatrix} A_{i,1,1} & \cdots & A_{i,1,C} \\ \vdots & \ddots & \vdots \\ A_{i,M,1} & \cdots & A_{i,M,C} \end{pmatrix} \right\|_{S^p}^q \right)^{1/q}$$

A Unified Framework for Vectorial Total Variation

Variant	Continuous Formulation	Our Framework
Isotropic uncoupled	$\int_\Omega \sum_{k=1}^C \sqrt{(\partial_{x_1} u_k(x))^2 + (\partial_{x_2} u_k(x))^2} \mathrm{d}x$	$\ell^{2,1,1}(der, col, pix)$
Anisotropic uncoupled	$\int_\Omega \sum_{k=1}^C \left(\lvert\partial_{x_1} u_k(x)\rvert + \lvert\partial_{x_2} u_k(x)\rvert\right) \mathrm{d}x$	$\ell^{1,1,1}(der, col, pix)$
Blomgren Chan	$\sqrt{\sum_{k=1}^{C} \left(\int_\Omega \sqrt{(\partial_{x_1} u_k(x))^2 + (\partial_{x_2} u_k(x))^2} \mathrm{d}x\right)^2}$	$\ell^{2,1,2}(der, pix, col)$
Anisotropic version	$\sqrt{\sum_{k=1}^{C} \left(\int_\Omega \left(\lvert\partial_{x_1} u_k(x)\rvert + \lvert\partial_{x_2} u_k(x)\rvert\right) \mathrm{d}x\right)^2}$	$\ell^{1,1,2}(der, pix, col)$
Bresson Chan	$\int_\Omega \sqrt{\sum_{k=1}^C \left(\partial_{x_1} u_k(x)\right)^2 + \sum_k \left(\partial_{x_2} u_k(x)\right)^2} \mathrm{d}x$	$\ell^{2,2,1}(col, der, pix)$
Anisotropic version	$\int_\Omega \sqrt{\sum_{k=1}^C \left(\lvert\partial_{x_1} u_k(x)\rvert + \lvert\partial_{x_2} u_k(x)\rvert\right)^2} \mathrm{d}x$	$\ell^{1,2,1}(der, col, pix)$

A Unified Framework for Vectorial Total Variation

Variant	Continuous Formulation	Our Framework
Anisotropic variant	$\int_\Omega \left(\sqrt{\sum_{k=1}^C (\partial_{x_1} u_k(x))^2} + \sqrt{\sum_{k=1}^C (\partial_{x_2} u_k(x))^2}\right) \mathrm{d}x$	$\ell^{2,1,1}(col, der, pix)$
Strong coupling	$\int_\Omega (\max_k \lvert\partial_{x_1} u_k(x)\rvert + \max_k \lvert\partial_{x_2} u_k(x)\rvert)\, \mathrm{d}x$	$\ell^{\infty,1,1}(col, der, pix)$
Isotropic version	$\int_\Omega \sqrt{\left(\max_k \lvert\partial_{x_1} u_k(x)\rvert\right)^2 + \left(\max_k \lvert\partial_{x_2} u_k(x)\rvert\right)^2} \mathrm{d}x$	$\ell^{\infty,2,1}(col, der, pix)$
Isotropic variant	$\int_\Omega \max_k \sqrt{(\partial_{x_1} u_k(x))^2 + (\partial_{x_2} u_k(x))^2} \mathrm{d}x$	$\ell^{2,\infty,1}(der, col, pix)$
Sapiro	$\int_\Omega \left\lVert \begin{pmatrix} (\partial_{x_1} u_k(x))_{k=1,\ldots,C} \\ (\partial_{x_2} u_k(x))_{k=1,\ldots,C} \end{pmatrix} \right\rVert_{S^1} \mathrm{d}x$	$S^1(col, der), \ell^1(pix)$
Goldluecke	$\int_\Omega \left\lVert \begin{pmatrix} (\partial_{x_1} u_k(x))_{k=1,\ldots,C} \\ (\partial_{x_2} u_k(x))_{k=1,\ldots,C} \end{pmatrix} \right\rVert_{S^\infty} \mathrm{d}x$	$S^\infty(col, der), \ell^1(pix)$

Minimization using a Primal-Dual Hybrid Gradient Method

Consider the linearly constrained convex optimization problem:

$$\min_{u,g} G(u) + F(g) \qquad \text{s.t.} \qquad Ku = g,$$

with data term G and $\ell^{p,q,r}$-norm or (S^p, ℓ^q)-norm F.

Minimization using a Primal-Dual Hybrid Gradient Method

Consider the linearly constrained convex optimization problem:

$$\min_{u,g} G(u) + F(g) \qquad \text{s.t.} \qquad Ku = g,$$

with data term G and $\ell^{p,q,r}$-norm or (S^p, ℓ^q)-norm F.

It can be solved by means of the following primal-dual algorithm [Pock, Cremers, Bischof, Chambolle '09]:

Minimization using a Primal-Dual Hybrid Gradient Method

Consider the linearly constrained convex optimization problem:

$$\min_{u,g} G(u) + F(g) \qquad \text{s.t.} \qquad Ku = g,$$

with data term G and $\ell^{p,q,r}$-norm or (S^p, ℓ^q)-norm F.

It can be solved by means of the following primal-dual algorithm [Pock, Cremers, Bischof, Chambolle '09]:

Iterate for $n \geq 0$ the following:

$$\begin{cases} \xi^{n+1} = \mathrm{prox}_{\sigma, F^*}\left(\xi^n + \sigma K \bar{u}^n\right), \\ u^{n+1} = \mathrm{prox}_{\tau, G}\left(u^n - \tau K^T \xi^{n+1}\right), \\ \bar{u}^{n+1} = u^{n+1} + \theta(u^{n+1} - u^n). \end{cases}$$

Minimization using a Primal-Dual Hybrid Gradient Method

Consider the linearly constrained convex optimization problem:

$$\min_{u,g} G(u) + F(g) \qquad \text{s.t.} \qquad Ku = g,$$

with data term G and $\ell^{p,q,r}$-norm or (S^p, ℓ^q)-norm F.

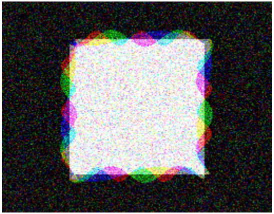
It can be solved by means of the following primal-dual algorithm [Pock, Cremers, Bischof, Chambolle '09]:

Iterate for $n \geq 0$ the following:

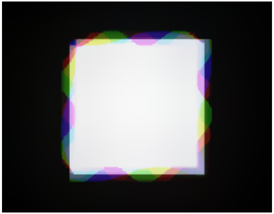
$$\begin{cases} \xi^{n+1} = \operatorname{prox}_{\sigma,F^*}\left(\xi^n + \sigma K \bar{u}^n\right), \\ u^{n+1} = \operatorname{prox}_{\tau,G}\left(u^n - \tau K^T \xi^{n+1}\right), \\ \bar{u}^{n+1} = u^{n+1} + \theta(u^{n+1} - u^n). \end{cases}$$

Converges to a saddle-point $(\hat{u}, \hat{\xi})$ for $\tau\sigma\|K\|^2 < 1$.

Which is the best channel coupling?



Noisy



$\ell^{1,1,1}(col, der, pix)$

$\ell^{2,1,1}(col, der, pix)$

$\ell^{\infty,1,1}(col, der, pix)$

Which is the best channel coupling?

Noisy

$\ell^{1,1,1}(col, der, pix)$

$\ell^{2,2,1}(col, der, pix)$

$\ell^{\infty,1,1}(col, der, pix)$

$(S^1(col, der), \ell^1(pix))$

$(S^\infty(col, der), \ell^1(pix))$

Experimental Results on Image Denoising

Figure: Noisy image with standard deviation 25. PSNR $= 20.74$.

Experimental Results on Image Denoising

Figure: $\ell^{\infty,1,1}-$regularization with $\lambda = 0.1$. PSNR $= 24.92$.

Experimental Results on Image Denoising

Figure: $\ell^{\infty,1,1}-$regularization with optimal $\lambda = 0.04$. PSNR $= 27.93$.

Experimental Results on Image Denoising

Figure: $\ell^{\infty,1,1}-$regularization with $\lambda = 0.01$. PSNR $= 24.09$.

Quantitative Evaluation on Kodak Database


Kodak 1


Kodak 2


Kodak 3


Kodak 4


Kodak 5


Kodak 6


Kodak 7


Kodak 8


Kodak 9


Kodak 10


Kodak 11


Kodak 12

Quantitative Evaluation on Kodak Database

	Noisy	$\ell^{1,1,1}$	$\ell^{2,1,1}$	$\ell^{2,2,1}$	$\ell^{\infty,1,1}$	$\ell^{\infty,2,1}$	$\ell^{2,\infty,1}$	(S^1, ℓ^1)	(S^∞, ℓ^1)
1	24.78	28.14	29.07	28.51	**29.90**	29.19	29.07	29.20	27.96
2	24.76	28.54	29.48	29.22	**30.18**	29.87	29.66	29.83	28.62
3	24.80	29.20	30.15	29.81	**30.85**	30.51	30.25	30.33	29.24
4	24.68	30.92	32.22	31.80	**32.73**	32.71	32.13	32.32	31.01
5	24.71	31.50	32.75	32.41	33.13	**33.30**	32.64	32.81	31.65
6	24.72	27.36	28.19	27.98	**29.01**	28.64	28.52	28.59	27.47
7	24.71	29.46	30.39	30.12	**30.86**	30.71	30.35	30.57	29.53
8	24.96	31.08	32.10	31.84	**32.41**	32.40	32.02	32.20	31.22
9	25.68	30.92	31.74	31.54	**32.10**	32.00	31.78	31.85	31.11
10	24.66	29.75	30.81	30.49	**31.48**	31.29	30.94	31.05	29.84
11	24.66	30.14	31.10	30.84	**31.49**	31.46	31.07	31.22	30.25
12	24.71	31.85	33.15	32.84	33.45	**33.69**	33.03	33.25	32.05
∅	24.82	29.91	30.93	30.62	**31.47**	31.31	30.96	31.10	30.00

Table: For each matrix TV method, the optimal λ in terms of PSNR was computed on the first Kodak image and then used on the others. The input noise standard deviation was 15.

Quantitative Evaluation on McMaster Database


McM 1


McM 2


McM 3


McM 4


McM 5


McM 6


McM 7


McM 8


McM 9


McM 10


McM 11


McM 12

Quantitative Evaluation on McMaster Database

	$\ell^{1,1,1}$	$\ell^{2,1,1}$	$\ell^{2,2,1}$	$\ell^{\infty,1,1}$	$\ell^{\infty,2,1}$	$\ell^{\infty,\infty,1}$	$\ell^{2,\infty,1}$	(S^1, ℓ^1)	(S^∞, ℓ^1)
1	29.29	29.83	29.64	29.74	29.52	28.97	29.25	**29.98**	29.16
2	27.80	28.41	28.26	28.43	28.32	27.80	28.02	**28.60**	27.75
3	30.44	30.96	30.84	30.78	30.66	30.16	30.39	**31.17**	30.33
4	29.26	29.91	29.75	29.95	29.82	29.30	29.54	**30.13**	29.22
5	31.11	31.46	31.40	30.97	30.84	30.33	30.55	**31.64**	30.89
6	29.83	30.49	30.32	30.34	30.13	29.55	29.84	**30.74**	29.68
7	30.96	31.63	31.48	31.41	31.21	30.66	30.98	**31.80**	30.87
8	31.98	32.72	32.60	32.50	32.30	31.78	32.15	**32.88**	31.99
9	32.54	33.36	33.32	33.08	32.93	32.50	32.85	**33.53**	32.70
10	32.26	33.06	33.02	32.70	32.54	32.10	32.49	**33.20**	32.37
11	30.21	30.85	30.75	30.87	30.73	30.35	30.59	**30.98**	30.29
12	30.58	31.18	30.99	31.11	30.87	30.36	30.69	**31.30**	30.50
∅	30.52	31.16	31.03	30.99	30.82	30.32	30.61	**31.33**	30.48

Table: For each matrix TV method, the optimal λ in terms of RMSE was computed on the first McMaster image and then used on the others.

Nonlocal Vectorial Total Variation

Variant	Continuous Formulation	Our Framework
Isotropic uncoupled	$\int_\Omega \left(\sum_{k=1}^{C} \sqrt{\int_\Omega (u_k(y) - u_k(x))^2 \, \omega(x,y) \, \mathrm{d}y} \right) \mathrm{d}x$	$\ell^{2,1,1}(der, col, pix)$
Anisotropic uncoupled	$\int_\Omega \left(\sum_{k=1}^{C} \int_\Omega \lvert u(y) - u(x) \rvert \sqrt{\omega(x,y)} \, \mathrm{d}y \right) \mathrm{d}x$	$\ell^{1,1,1}(der, col, pix)$
Duan Pan, Tai	$\sqrt{\sum_{k=1}^{C} \left(\int_\Omega \sqrt{\int_\Omega (u_k(y) - u_k(x))^2 \, \omega(x,y) \, \mathrm{d}y} \, \mathrm{d}x \right)^2}$	$\ell^{2,1,2}(der, pix, col)$
Anisotropic coupled	$\int_\Omega \int_\Omega \sqrt{\sum_{k=1}^{C} (u_k(y) - u_k(x))^2 \, \omega(x,y)} \, \mathrm{d}y \, \mathrm{d}x$	$\ell^{2,1,1}(col, der, pix)$
Isotropic coupled	$\int_\Omega \sqrt{\int_\Omega \sum_{k=1}^{C} (u_k(y) - u_k(x))^2 \, \omega(x,y) \, \mathrm{d}y} \, \mathrm{d}x$	$\ell^{2,2,1}(col, der, pix)$
Strong coupling	$\int_\Omega \int_\Omega \max_k \left((u_k(y) - u_k(x))^2 \, \omega(x,y) \right) \mathrm{d}y \, \mathrm{d}x$	$\ell^{\infty,1,1}(col, der, pix)$

Local versus Nonlocal Color Total Variation

Clean image

Local versus Nonlocal Color Total Variation

Noisy image with noise s.d. 12.75. PSNR $= 26.10$.

Local versus Nonlocal Color Total Variation

$\ell^{1,1,1}-$TV regularization. PSNR $= 33.60$.

Local versus Nonlocal Color Total Variation

$\ell^{1,1,1}$–NLTV regularization. PSNR $= 35.41$.

Local versus Nonlocal Color Total Variation

$\ell^{\infty,1,1}$−TV regularization. PSNR $= 34.88$.

Local versus Nonlocal Color Total Variation

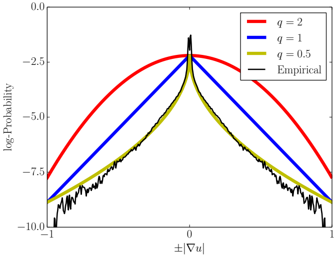
$\ell^{\infty,1,1}$–NLTV regularization. PSNR $= 35.65$.

Quantitative Evaluation on Kodak Database

Kodak	Noisy	$\ell^{1,1,1}$	$\ell^{2,1,1}$	$\ell^{2,2,1}$	$\ell^{\infty,1,1}$
1	26.15	31.01	31.14	31.07	**31.20**
2	26.14	31.23	31.36	31.21	**31.44**
3	26.17	31.78	31.88	31.76	**31.99**
4	26.08	34.38	**35.06**	34.66	35.03
5	26.10	35.02	35.69	35.35	**35.73**
6	26.11	29.28	29.37	29.30	**29.60**
7	26.08	31.64	31.70	31.58	**31.77**
8	26.31	33.88	34.24	34.02	**34.29**
9	26.98	34.40	34.74	34.67	**34.78**
10	26.06	32.21	32.50	32.36	**32.61**
11	26.06	32.31	32.39	32.27	**32.45**
12	26.09	35.17	35.93	35.33	**35.94**
$\varnothing$	26.19	32.69	33.00	32.80	**33.07**

Nonconvex Extension

The statistics of natural images [Huang, Mumford '99] suggest the use of nonconvex regularizers.



The nuclear norm is a convex relaxation of rank minimization. Respective non-convex formulations should more directly penalize the rank of the Jacobian thereby favoring parallel color gradients (rank 1) and piecewise constant regions (rank 0).

Nonconvex Extension

We propose the following generalizations:

- Vectorial TV^q based on Frobenius norm:

$$TV_F^q(u) = \int_\Omega \|\nabla u\|_F^q \,\mathrm{d}x, \qquad q \geq 0.$$

Nonconvex Extension

We propose the following generalizations:

- Vectorial TV^q based on Frobenius norm:

$$TV_F^q(u) = \int_\Omega \|\nabla u\|_F^q \, \mathrm{d}x, \qquad q \geq 0.$$

- Schatten-q TV:

$$TV_{S^q}^q(u) = \int_\Omega \|\nabla u\|_{S^q}^q \, \mathrm{d}x, \qquad q \geq 0,$$

Nonconvex Extension

We propose the following generalizations:

- Vectorial TV^q based on Frobenius norm:

$$TV_F^q(u) = \int_\Omega \|\nabla u\|_F^q \, \mathrm{d}x, \qquad q \geq 0.$$

- Schatten-q TV:

$$TV_{S^q}^q(u) = \int_\Omega \|\nabla u\|_{S^q}^q \, \mathrm{d}x, \qquad q \geq 0,$$

where the Schatten-q norm is defined as

$$\|A\|_{S_q} = (\sigma_1^q + \ldots + \sigma_n^q)^{1/q}.$$

Nonsmooth and nonconvex optimization

- Majorization-minimization methods for non-convex problems

Nonsmooth and nonconvex optimization

- Majorization-minimization methods for non-convex problems
- Iteratively reweighted L_1 minimization [Ochs et al. '12]

Nonsmooth and nonconvex optimization

- Majorization-minimization methods for non-convex problems
- Iteratively reweighted L_1 minimization [Ochs et al. '12]
- Iteratively reweighted nuclear norm min. [Gu et al. '14]

Nonsmooth and nonconvex optimization

- Majorization-minimization methods for non-convex problems
- Iteratively reweighted L_1 minimization [Ochs et al. '12]
- Iteratively reweighted nuclear norm min. [Gu et al. '14]
- Extend primal-dual algorithms to the nonconvex setting

Nonsmooth and nonconvex optimization

- Majorization-minimization methods for non-convex problems
- Iteratively reweighted L_1 minimization [Ochs et al. '12]
- Iteratively reweighted nuclear norm min. [Gu et al. '14]
- Extend primal-dual algorithms to the nonconvex setting

Proposition

Let $F(g) = |g|^q$ and $0 \leq q < 1$. The Fenchel conjugate is given by

$$F^*(\xi) = \begin{cases} 0, & |\xi| = 0, \\ \infty, & |\xi| \neq 0, \end{cases}$$

and the biconjugate/convex envelope $(F^*)^*$ is zero everywhere.

Direct application of the PDHG doesn't impose any regularization!

A Primal-Dual Algorithm for Nonconvex Regularizers

In [Strekalovskiy, Cremers '14], we consider the problem

$$\min_{u,g} G(u) + F(g) \quad \text{s.t. } g = Ku, \quad \text{with nonconvex } F.$$

A Primal-Dual Algorithm for Nonconvex Regularizers

In [Strekalovskiy, Cremers '14], we consider the problem

$$\min_{u,g} G(u) + F(g) \quad \text{s.t. } g = Ku, \quad \text{with nonconvex } F.$$

Introducing a Lagrange multiplier y leads to

$$\max_{y} \min_{u,g} G(u) + F(g) + \langle y, Ku - g \rangle,$$

A Primal-Dual Algorithm for Nonconvex Regularizers

In [Strekalovskiy, Cremers '14], we consider the problem

$$\min_{u,g} G(u) + F(g) \quad \text{s.t. } g = Ku, \quad \text{with nonconvex } F.$$

Introducing a Lagrange multiplier y leads to

$$\max_{y} \min_{u,g} G(u) + F(g) + \langle y, Ku - g \rangle,$$

which is solved with primal-dual algorithm

$$g^{n+1} = \arg\min_{g} \frac{\sigma}{2} \|g - K\bar{u}^n\|^2 - \langle g, y^n \rangle + F(g),$$

$$y^{n+1} = y^n + \sigma(K\bar{u}^n - g^{n+1}),$$

$$u^{n+1} = \arg\min_{u} \frac{1}{2\tau} \|u - u^n\|^2 + \langle Ku, y^{n+1} \rangle + G(u),$$

$$\bar{u}^{n+1} = u^{n+1} + \theta(u^{n+1} - u^n).$$

A Primal-Dual Algorithm for Nonconvex Regularizers

Proposition (Strekalovskiy, Cremers ECCV '14)

For convex problems, the above algorithm is equivalent to the primal-dual algorithm of [Pock, Cremers, Chambolle, Bischof '09].

A Primal-Dual Algorithm for Nonconvex Regularizers

Proposition (Strekalovskiy, Cremers ECCV '14)

For convex problems, the above algorithm is equivalent to the primal-dual algorithm of [Pock, Cremers, Chambolle, Bischof '09].

Proposition (Strekalovskiy, Cremers ECCV '14)

For nonconvex regularizers F, the above algorithm still incorporates the regularizer in a non-trivial manner.

A Primal-Dual Algorithm for Nonconvex Regularizers

Proposition (Strekalovskiy, Cremers ECCV '14)

For convex problems, the above algorithm is equivalent to the primal-dual algorithm of [Pock, Cremers, Chambolle, Bischof '09].

Proposition (Strekalovskiy, Cremers ECCV '14)

For nonconvex regularizers F, the above algorithm still incorporates the regularizer in a non-trivial manner.

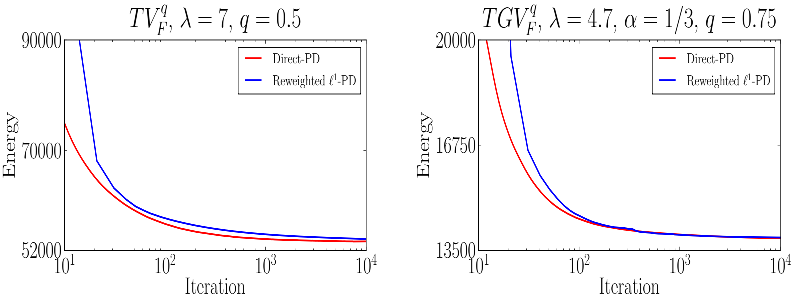
Proposition (Möllenhoff, Strekalovskiy, Möller, Cremers SIIMS '15)

Let $G - \frac{c}{2}\|\cdot\|_2^2$ and $F + \frac{\omega}{2}\|\cdot\|_2^2$ be convex with $c > \omega\|K\|_2^2$. Then the latter algorithm converges to the (unique) minimizer of

$$G(u) + F(Ku)$$

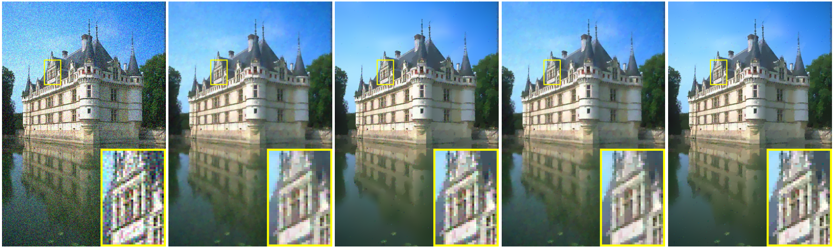
for $0 < \sigma = 2\omega$, $\tau\sigma\|K\|_2^2 \leq 1$, and any $\theta \in [0,1]$ with rate $1/N$.

Numerical results - convergence



Numerical results - natural image denoising ($q = 0.75$)

Extending the Total Generalized Variation (TGV) [Bredies, Kunisch, Pock '10] and the multichannel version TGV_F [Bredies '14], we proposed a nuclear-norm vectorial version TGV_{S^1} and respective non-convex formulations TGV_F^q and $TGV_{S^q}^q$.



Noisy, $\sigma = 0.1$	TGV_F, PSNR=28.5	TGV_F^q, PSNR=28.9	TGV_{S^1}, PSNR=29.0	$TGV_{S^q}^q$, PSNR=**29.4**

Conclusion

- We introduced a unified framework for Vectorial Total Variation based on the mixed matrix norms $\ell^{p,q,r}$ and (S^p, ℓ^q).

Conclusion

- We introduced a unified framework for Vectorial Total Variation based on the mixed matrix norms $\ell^{p,q,r}$ and (S^p, ℓ^q).
- Depending on the amount of inter-channel correlation, different matrix norms are suited.

Conclusion

- We introduced a unified framework for Vectorial Total Variation based on the mixed matrix norms $\ell^{p,q,r}$ and (S^p, ℓ^q).
- Depending on the amount of inter-channel correlation, different matrix norms are suited.
- $\ell^{\infty,1,1}$ and (S^1, ℓ^1) best exploit color-channel correlations.

Conclusion

- We introduced a unified framework for Vectorial Total Variation based on the mixed matrix norms $\ell^{p,q,r}$ and (S^p, ℓ^q).
- Depending on the amount of inter-channel correlation, different matrix norms are suited.
- $\ell^{\infty,1,1}$ and (S^1, ℓ^1) best exploit color-channel correlations.
- We proposed respective Nonlocal Vectorial Total Variations.

Conclusion

- We introduced a unified framework for Vectorial Total Variation based on the mixed matrix norms $\ell^{p,q,r}$ and (S^p, ℓ^q).
- Depending on the amount of inter-channel correlation, different matrix norms are suited.
- $\ell^{\infty,1,1}$ and (S^1, ℓ^1) best exploit color-channel correlations.
- We proposed respective Nonlocal Vectorial Total Variations.
- We proposed non-convex formulations of respective Vectorial TV and Vectorial TGV. In particular, $TGV^q_{S^q}$ enables a more direct rank penalization enforcing color channel alignment.

Conclusion

- We introduced a unified framework for Vectorial Total Variation based on the mixed matrix norms $\ell^{p,q,r}$ and (S^p, ℓ^q).
- Depending on the amount of inter-channel correlation, different matrix norms are suited.
- $\ell^{\infty,1,1}$ and (S^1, ℓ^1) best exploit color-channel correlations.
- We proposed respective Nonlocal Vectorial Total Variations.
- We proposed non-convex formulations of respective Vectorial TV and Vectorial TGV. In particular, $TGV^q_{S^q}$ enables a more direct rank penalization enforcing color channel alignment.
- We proposed two primal-dual algorithms for convex and non-convex regularizers F which coincide for convex F.

Publications

- Duran, Moeller, Sbert, Cremers, “Collaborative Total Variation: A General Framework for Vectorial TV Models”, SIAM-IS 2015.
- Duran, Moeller, Sbert, Cremers, “On the Implementation of Collaborative TV Regularization”, IPOL 2016.
- Duran, Moeller, Sbert, Cremers, “A Novel Framework for Nonlocal Vectorial Total Variation Based on $\ell^{p,q,r}$-norms”, EMMCVPR 2015.
- Moellenhoff, Strekalovskiy, Moeller, Cremers, “Low Rank Priors for Color Image Regularization”, EMMCVPR 2015.
- Pock, Cremers, Chambolle, Bischof, “An Algorithm for Minimizing the Mumford-Shah Functional”, ICCV 2009.
- Strekalovskiy, Cremers, “Real-Time Minimization of the Piecewise Smooth Mumford-Shah Functional”, ECCV 2014.
- Moellenhoff, Strekalovskiy, Moeller, Cremers, “The Primal-Dual Hybrid Gradient Method for Semiconvex Splittings”, SIAM-IS 2015.