What do regularisers do?

Tuomo Valkonen

University of Cambridge ($\rightsquigarrow$ Liverpool)

MIA'16, Paris, 2016-01-20

This work has been supported at EPN Quito by a Prometeo fellowship of the Ecuadorian ministry of Science, Education, Technology, and Innovation (Senescyt), and in Cambrige by the King Abdullah University of Science and Technology (KAUST) Award No. KUK-I1-007-43, as well as the EPSRC / Isaac Newton Trust Small Grant "Non-smooth geometric reconstruction for high resolution MRI imaging of fluid transport in bed reactors", and the EPSRC first grant Nr. EP/J009539/1 "Sparse & Higher-order Image Restoration".

Which regulariser is the best?

Is any of them any good?

Do regulariser introduce artefacts?

What other qualitative properties do they have?

TGV denoising and the jump set

TGV denoising and the jump set

For a regulariser R, suppose u solves

$$\min_{u \in \mathrm{BV}(\Omega)} \frac{1}{2}\|f - u\|^2_{L^2(\Omega)} + R(u).$$

What can we say about u? Do we have:

$$\mathcal{H}^{n-1}(J_u \setminus J_f) = 0?$$

TGV denoising and the jump set

Our studies motivated by the choice

$$\begin{aligned} R(u) &= \mathrm{TGV}^2_{(\beta,\alpha)}(u) \\ &:= \min_w \alpha \|Du - w\|_1 + \beta \|Ew\|_1. \end{aligned}$$

(Bredies, Kunisch, and Pock 2011; Bredies and T.V. 2011)

The co-area formula

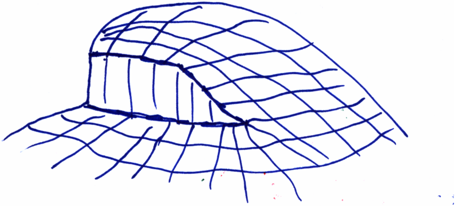
...but let us begin with

$$\mathrm{TV}(u) := \|Du\| = \int_{-\infty}^{\infty} \mathrm{Per}(\{u > t\}; \Omega)\, dt.$$

$\implies$ Minimal surface problems on level sets.
(Alter, Caselles, and Chambolle 2005; Allard 2008)

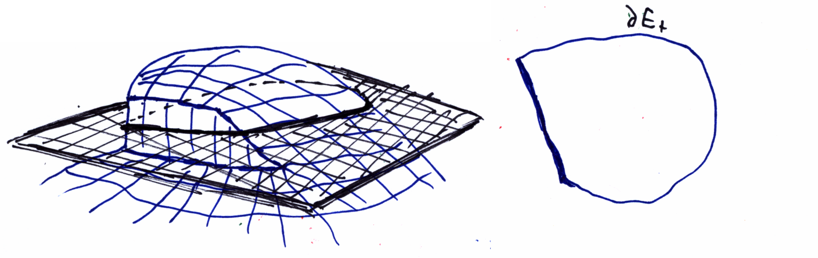
Level set approach for $R = \alpha \mathrm{TV}$

(Caselles, Chambolle, and Novaga 2008)



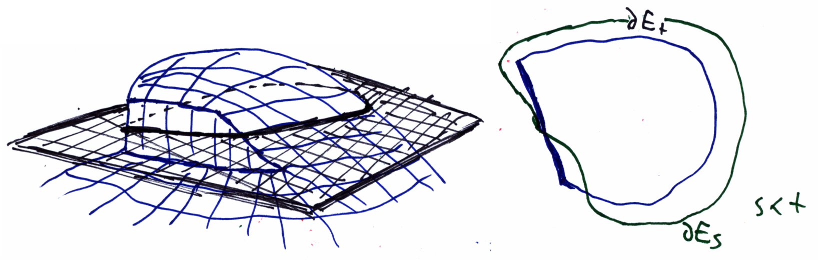
Level set approach for $R = \alpha \mathrm{TV}$

(Caselles, Chambolle, and Novaga 2008)



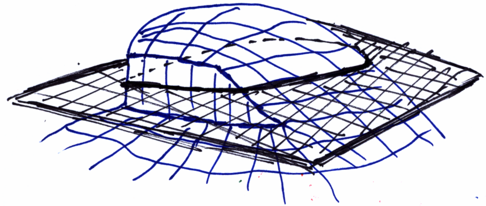
Level set approach for $R = \alpha \mathrm{TV}$

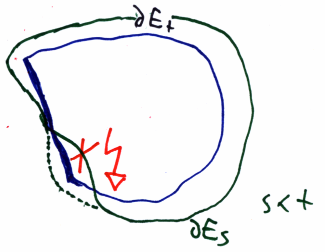
(Caselles, Chambolle, and Novaga 2008)



Level set approach for $R = \alpha \mathrm{TV}$

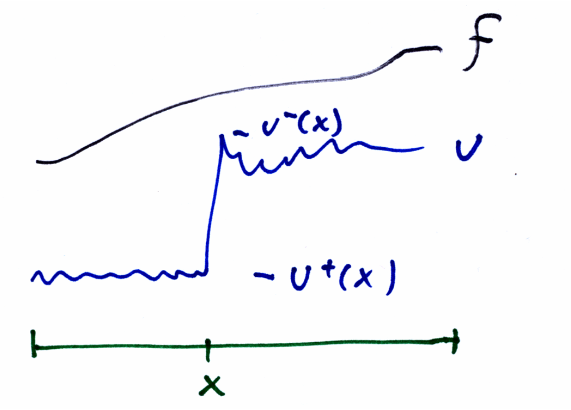
(Caselles, Chambolle, and Novaga 2008)





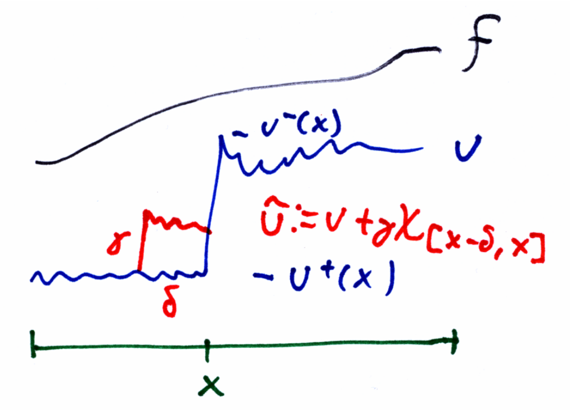
Proof for TGV in 1-D

(Bredies, Kunisch, and T.V. 2013)



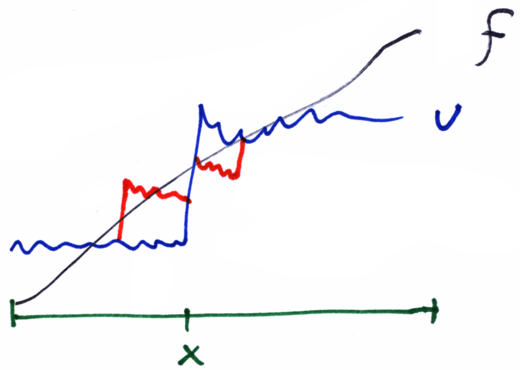
Proof for TGV in 1-D

(Bredies, Kunisch, and T.V. 2013)

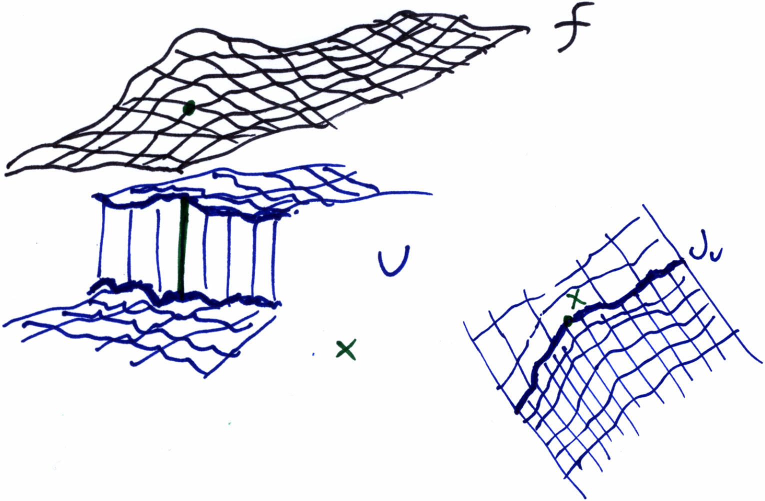


Proof for TGV in 1-D

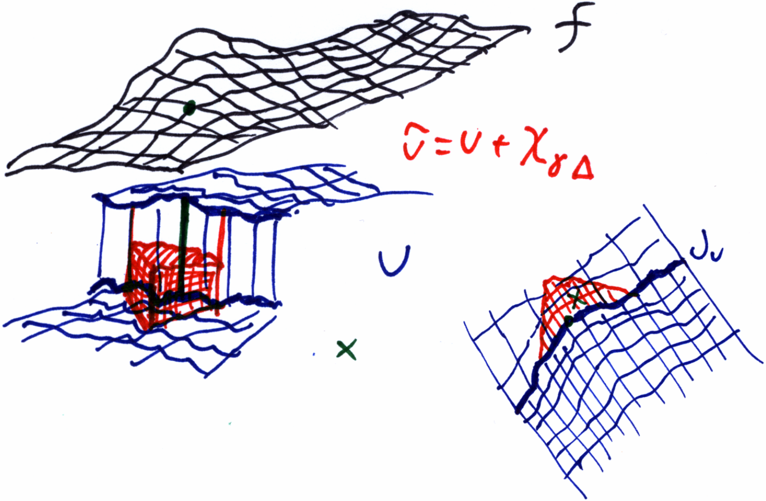
(Bredies, Kunisch, and T.V. 2013)



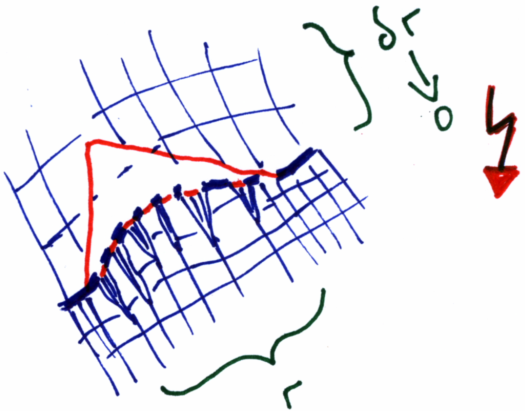
First idea of generalising to n-D



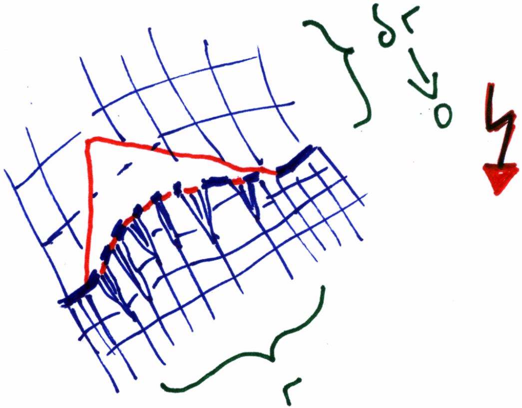
First idea of generalising to n-D

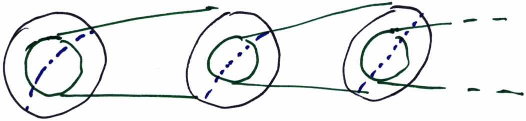


Failure



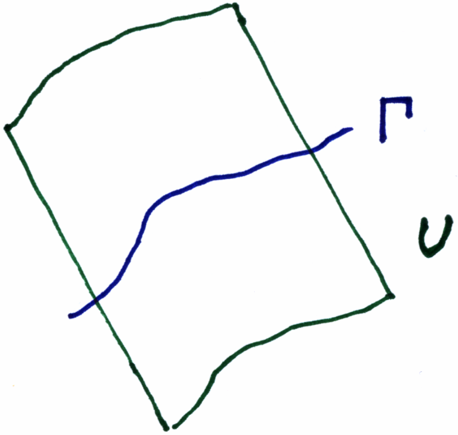
Failure





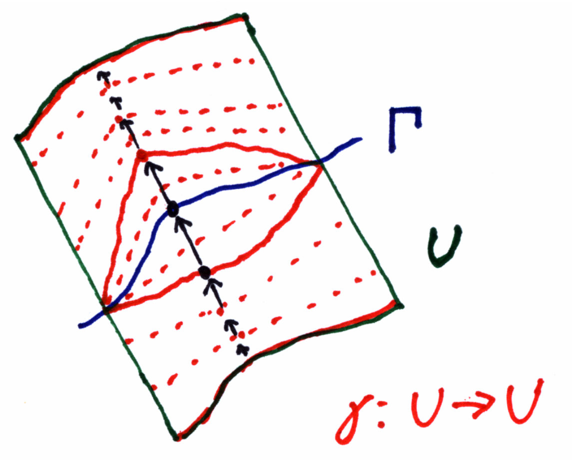
Success: *double* Lipschitz transformation

(T.V. 2015, 2014)



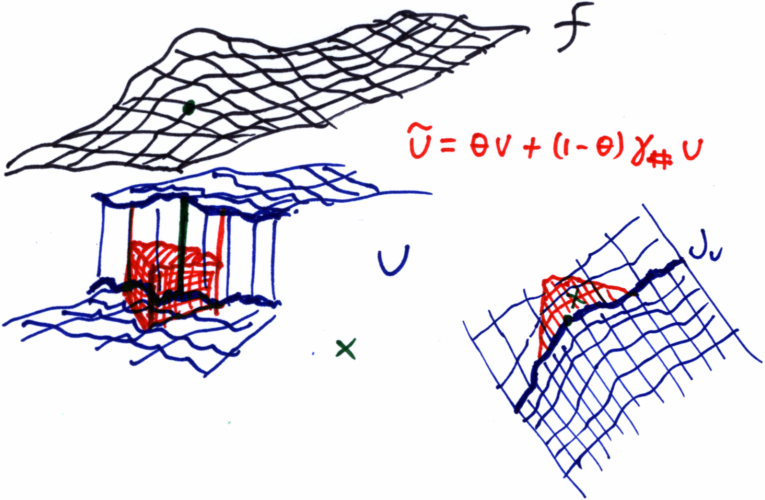
Success: *double* Lipschitz transformation

(T.V. 2015, 2014)

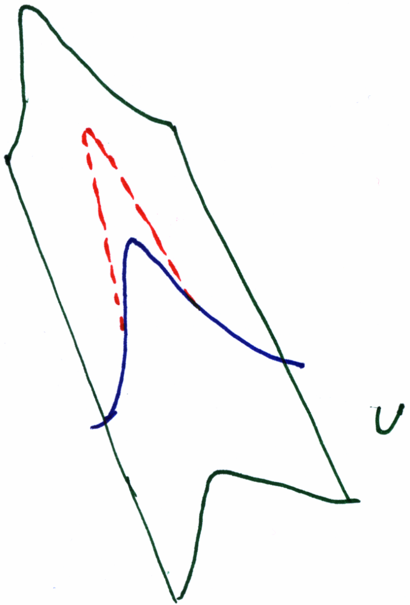


Success: *double* Lipschitz transformation

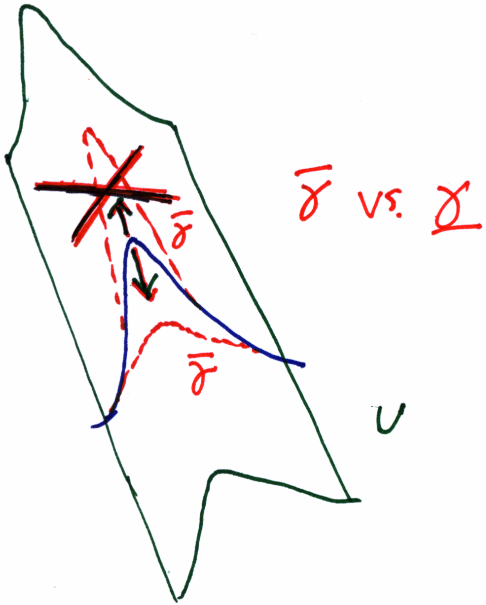
(T.V. 2015, 2014)



Success: *double* Lipschitz transformation



Success: *double* Lipschitz transformation



Success: *double* Lipschitz transformation

Definition.

$$R(\overline{\gamma}_{\#}u) + R(\underline{\gamma}_{\#}u) - 2R(u) \leq C_R T_{\overline{\gamma},\underline{\gamma}}.$$

Proof of $\mathcal{H}^{m-1}(J_u \setminus J_f) = 0$ (sketch).

For our shift transformations, with max. shift ρ,

$$T_{\overline{\gamma},\underline{\gamma}} \leq C\rho^2.$$

...but fidelity improvement $\geq C'\rho$! □

Well, almost…

Works for TV and Huber-TV.

Also Perona-Malik and TV^q, if were well-posed…

But still does not work for TGV or ICTV…

Partial Lipschitz transformation

Compare *partial push-forwards* $u_{\overline{\gamma}}$ and $u_{\underline{\gamma}}$ defined for suitable v by

$$u_\gamma := \gamma_\#(u - v) + v.$$

Partial Lipschitz transformation

Compare *partial push-forwards* $u_{\overline{\gamma}}$ and $u_{\underline{\gamma}}$ defined for suitable v by

$$u_\gamma := \gamma_\#(u - v) + v.$$

Motivation: TV result (formally) usable for

$$\mathrm{ICTV}(u) := \min_v \|D(u - v)\| + \|D\nabla v\|,$$

setting $f' := f - v$ and $u' := u - v$.

(ICTV: Chambolle and Lions 1997)

Still something missing

Standard TGV: Regularity results in BD.

Generally: Local boundedness of u.

Still something missing

Standard TGV: Regularity results in BD.

Generally: Local boundedness of u.

$\Longrightarrow$ Complete proof *(T.V. 2015, 2014)*

- ICTV in dimension $m = 2$.

Crucial: $BV^2(\Omega) \hookrightarrow L^\infty(\Omega)$ *(Demengel 1984)*.

Still something missing

Standard TGV: Regularity results in BD.

Generally: Local boundedness of u.

$\Longrightarrow$ Complete proof *(T.V. 2015, 2014)*

- ICTV in dimension $m = 2$.

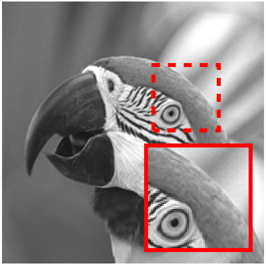
Crucial: $BV^2(\Omega) \hookrightarrow L^\infty(\Omega)$ *(Demengel 1984)*.

$\Longrightarrow$ Almost complete proof *(T.V. 2015, 2014)*

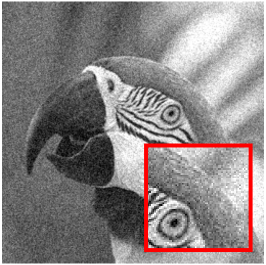
- non-symmetric TGV,
- q-norm TGV, $q > 1$.

Crucial: Korn's inequality.

$\mathrm{TGV}^{2,q}_{(\beta,\alpha)}(u) := \min_w \alpha \|Du - w\|_1 + \beta \|Ew\|_q$

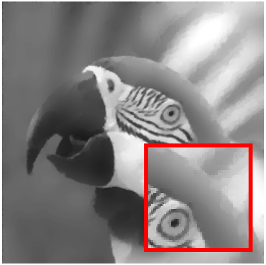


(a) Original

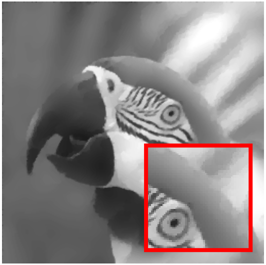


(b) Noisy image

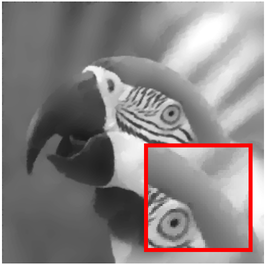
(c) TV, $\alpha = 25$



(d) $q = 1$, $\beta = 250$

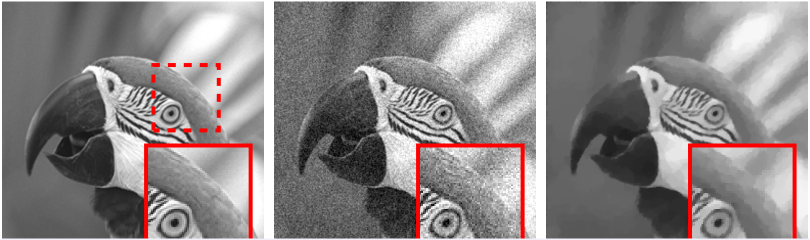


(e) $q = 1\frac{1}{2}$, $\beta = 10079$



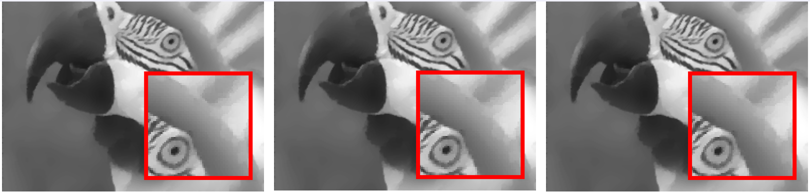
(f) $q = 2$, $\beta = 64000$

$$\mathrm{TGV}^{2,q}_{(\beta,\alpha)}(u) := \min_w \alpha\|Du - w\|_1 + \beta\|Ew\|_q$$



β factor from norm equivalence / Cauchy-Schwarz.

$\mathrm{TGV}^{2,q}$ PSNR 29.2 $\forall q$; TV PSNR 28.0.



(d) $q = 1$, $\beta = 250$ (e) $q = 1\frac{1}{2}$, $\beta = 10079$ (f) $q = 2$, $\beta = 64000$

Do regularisers improve images?

Optimal TV parameter

Bi-level optimisation

$$\hat{\alpha} := \arg\min_{\alpha \geq 0} F(u_\alpha)$$

subject to

$$u_\alpha \in \arg\min_u \ \frac{1}{2}\|f - u\|^2 + \alpha \mathrm{TV}(u)$$

Optimal TV parameter

Theorem. *For* $F(u) = \frac{1}{2}\|f_0 - u\|^2$*, we have*

$$\hat{\alpha} > 0$$

if

$$\mathrm{TV}(f) > \mathrm{TV}(f_0).$$

(de Los Reyes, Schönlieb, and T.V. 2015)

Optimal TV parameter

Theorem. *For* $F(u) = \|Df_0 - Du\|_{discrete}$, *we have*

$$\hat{\alpha} > 0$$

if for some $t > 0$ *and* $-\operatorname{div}\xi \in \partial F(f)$ *holds*

$$\mathrm{TV}(f) > \mathrm{TV}(f + t\operatorname{div}\xi).$$

(de Los Reyes, Schönlieb, and T.V. 2015)

Other regularisers

General problem

$$\hat{\vec{\alpha}} := \arg\min_{\vec{\alpha} \geq 0} F(u_{\vec{\alpha}})$$

with

$$u_{\vec{\alpha}} \in \arg\min_{u} \ \frac{1}{2}\|Ku - f\|^2 + \sum_{j=1}^{N} \alpha_j \|A_j u\|_{\mathcal{M}}$$

	$u =$	$Ku =$	$A_1 u =$	$A_2 u =$
TGV^2	(v, w)	v	$Dv - w$	Ew
ICTV	(v_1, v_2)	$v_1 + v_2$	Dv_1	$D^2 v_2$

Optimal TGV^2 parameter

Theorem. *For* $F(u) = \frac{1}{2}\|f_0 - u\|^2$*, we have*

$$\hat{\alpha}, \hat{\beta} > 0$$

if $\exists\ \alpha_0 > 0$ *with*

$$\mathrm{TGV}^2_{1,\alpha_0}(f) > \mathrm{TGV}^2_{1,\alpha_0}(f_0).$$

(de Los Reyes, Schönlieb, and T.V. 2015)

Optimal TGV^2 parameter

Theorem. *For* $F(u) = \|Df_0 - Du\|_{discrete}$, *we have*

$$\hat{\alpha}, \hat{\beta} > 0$$

if $\exists\ \alpha_0 > 0,\ t > 0,$ *and* $-\operatorname{div}\xi \in \partial F(f)$ *with*

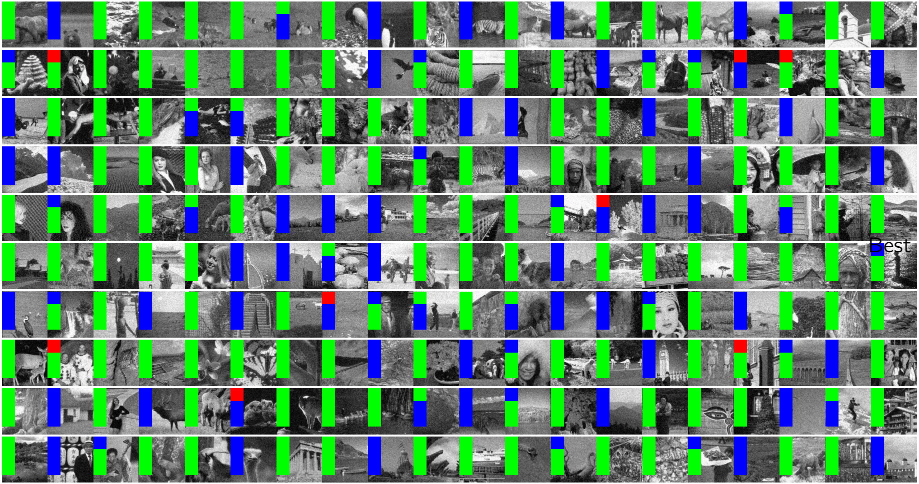
$$\mathrm{TGV}^2_{1,\alpha_0}(f) > \mathrm{TGV}^2_{1,\alpha_0}(f + t\operatorname{div}\xi).$$

(de Los Reyes, Schönlieb, and T.V. 2015)

Which regulariser is the best?

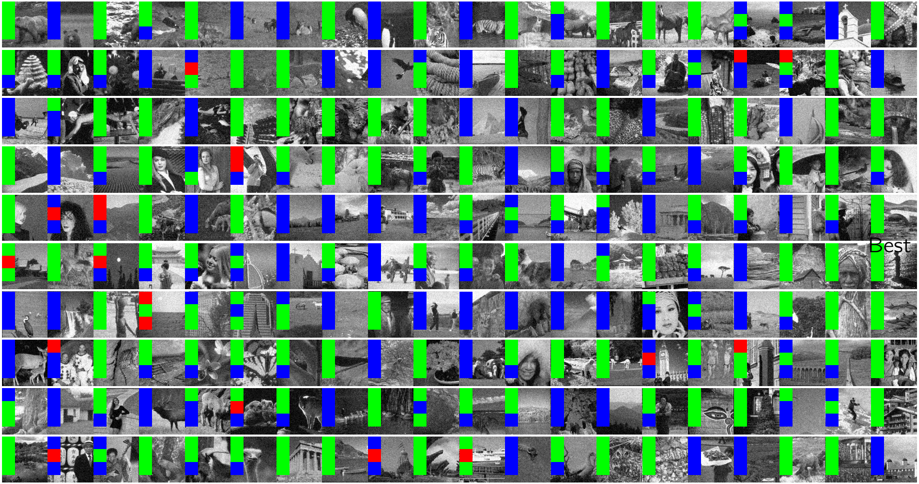
BSDS300 data set

Results $\sigma = 20$, $F(u) = \frac{1}{2}\|f_0 - u\|^2$



regulariser: TV, ICTV, TGV^2; top=SSIM, middle=PSNR, bottom=F.

Results $\sigma = 20$, $F(u) = \|Df_0 - Du\|_{\text{discrete}}$



regulariser: TV, ICTV, TGV^2; top=SSIM, middle=PSNR, bottom=F.

According to 95% t-test, ICTV works best.

On piecewise smooth images, TGV^2 is visually most pleasing.

(de Los Reyes, Schönlieb, and T.V. 2014)


The end