What if image self-similarity can be better exploited in data fidelity terms?

Ivan Pereira-Sánchez, Julia Navarro and Joan Duran

Dept. Mathematics and Computer Science, IAC3. Universitat de les Illes Balears, E-07122, Palma, Spain

(i.pereira@uib.es, julia.navarro@uib.es and joan.duran@uib.es)

Universitat de les Illes Balears

IAC3

nº 3299

1. Introduction

We propose a variational model for image restoration that exploits self-similarity of natural images in the fidelity term instead of the regularization term. Nonlocal regularization transfers similarity between patches from the degraded image to the restored one. In contrast, the assumption behind a nonlocal fidelity term is that two similar patches in the degraded image must resemble after going through the degradation process. The performance of the proposed model is evaluated on super-resolution, denoising and depth filtering.

2. Proposed model

Let us denote the degraded image as $f \in \mathbb{R}^{C\times N}$, where N is the number of pixels and C is the number of spectral bands, the restored image as $u \in \mathbb{R}^{C\times M} := X$, with $M \geq N$, and $A : \mathbb{R}^{C\times M} \to \mathbb{R}^{C\times N}$ that model the problem.

Then we define the nonlocal fidelity term as:

$$\|(Au-f)_\omega\|_{2,2,1} := \sum_{k=1}^{C} \sqrt{\sum_{i,j=1}^{N} \omega_{i,j} \left((Au)_{k,i} - f_{k,j}\right)^2},$$

where $(Au-f)_\omega \in \mathbb{R}^{C\times N\times N}$ is defined as

$$((Au-f)_\omega)_{k,i,j} := \sqrt{\omega_{i,j}}\left((Au)_{k,i} - f_{k,j}\right).$$

The weights $\{\omega_{i,j}\}$ are computed on the observed data f, taking into account both the spatial closeness and the similarity in f. In practice, the weights are defined as

$$\omega_{i,j} = \frac{1}{\Gamma_i}\exp\left(-\frac{\|i-j\|^2}{h_{\text{spt}}^2} - \frac{\|f(P_i)-f(P_j)\|^2}{h_{\text{sim}}^2}\right)$$

if $\|i-j\|_\infty \leq \nu$ and zero otherwise and Γ_i is a normalitzation factor. The filtering parameters $h_{\text{spt}} > 0$ and $h_{\text{sim}} > 0$ measure how fast the weights decay with increasing spatial distance or dissimilarity between patches.

Therefore, we propose the following variational model for image restoration:

$$\min_{u\in X} \|\nabla u\|_{2,2,1} + \lambda\|(Au-f)_\omega\|_{2,2,1}, \tag{1}$$

The minimization problem (1) is convex but non smooth. To find a fast, global optimal solution we use the first-order primal-dual algorithm proposed by Chambolle and Pock [2].

3. Experimental results

• Super-Resolution experimentation

	BIC	TV	NLTV	IRCNN	BSRGAN	DRCNN	Ours
Building	10.38	10.55	9.63	9.97	9.68	9.85	9.54
Computer	9.90	9.82	8.73	8.61	9.88	8.76	8.57
Dice	4.44	2.35	2.20	6.20	3.34	5.29	2.60
Flowers	11.10	12.21	10.92	10.27	11.13	10.28	10.31
Hall	5.87	4.53	4.19	6.79	4.49	6.07	4.33
Leaves	5.03	3.86	3.70	5.88	4.44	5.47	3.85
Traffic	11.67	11.82	10.81	10.57	11.48	10.75	10.40
Trees	20.31	24.69	22.71	18.76	19.70	19.08	19.75
Valldemossa	11.21	12.10	13.62	10.52	9.30	10.46	10.62
Yard	9.15	8.76	8.15	8.45	7.44	8.45	7.84
Avg.	9.90	10.06	9.46	9.60	9.08	9.44	**8.78**

Table: RSME evaluation on the dataset from IPOL for $s = 2$, $\sigma_{blur} = 0.75$ and $\sigma_{noise} = 5$.

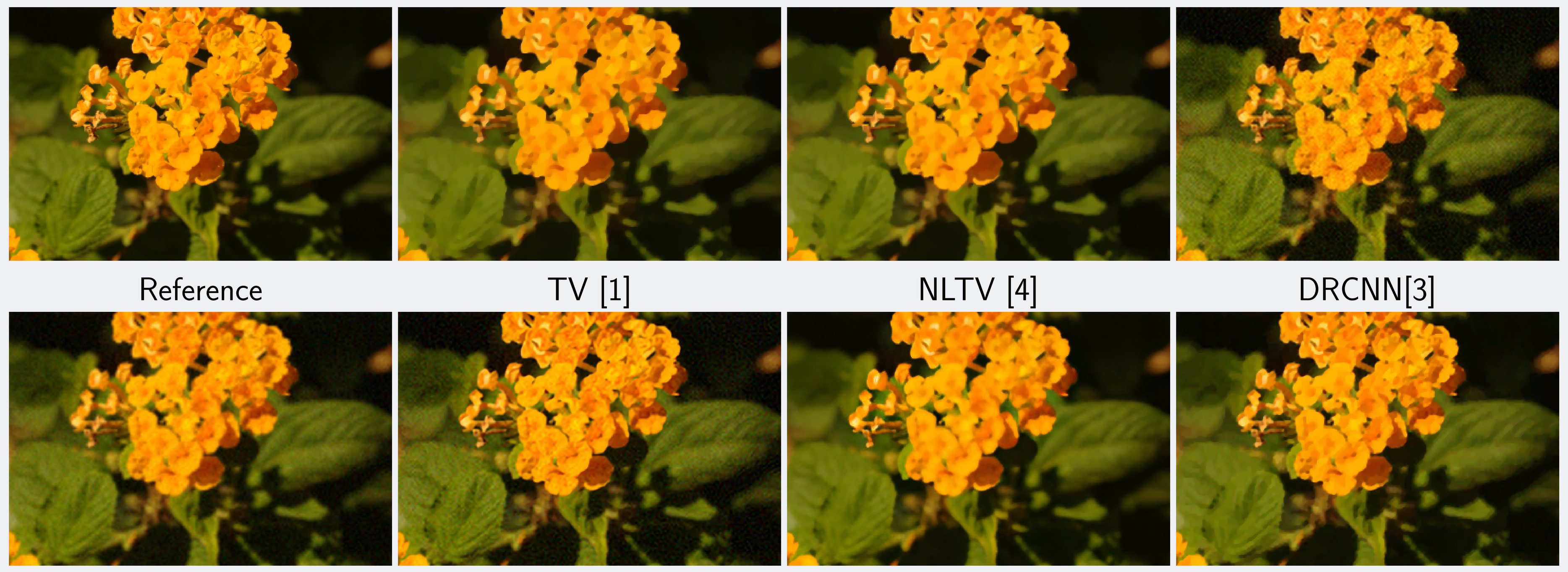

Figure: Close-ups of the results for single-image super-resolution on *Leaves*, with $s = 2$, $\sigma_{blur} = 0.75$ and $\sigma_{noise} = 5$. While TV, NLTV and BSRGAN provide over-smoothed results, DRCNN and IRCNN are very sensitive to noise. Instead, our method better recovers contours and texture while correctly removing the noise.

• Denoising example:

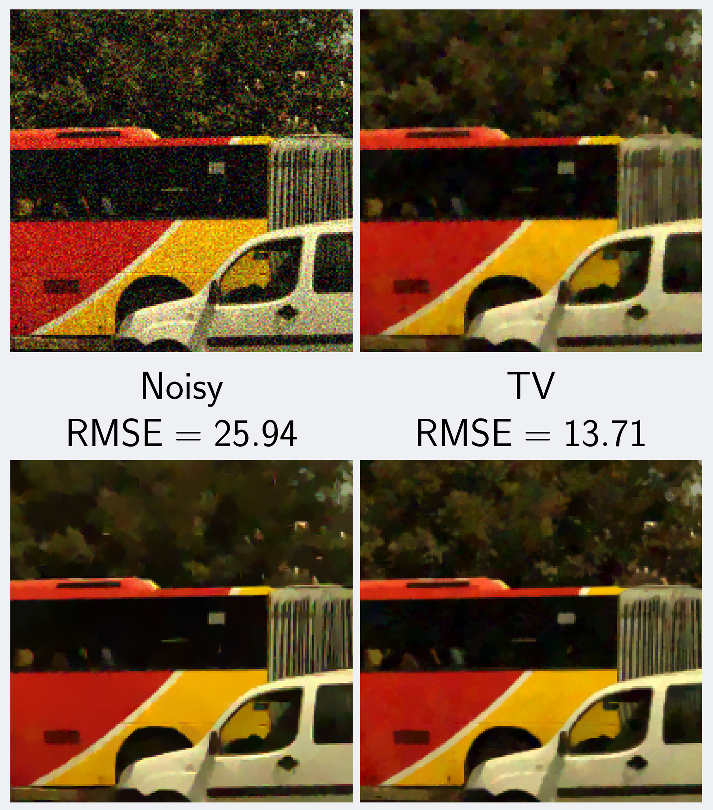

Figure: Close-ups of the results for image denoising on *Traffic* with $\sigma_{noise} = 30$. In contrast to TV and NLTV, our variational method is able to remove high-level noise while preserving texture and geometry.

• Depth filtering example:

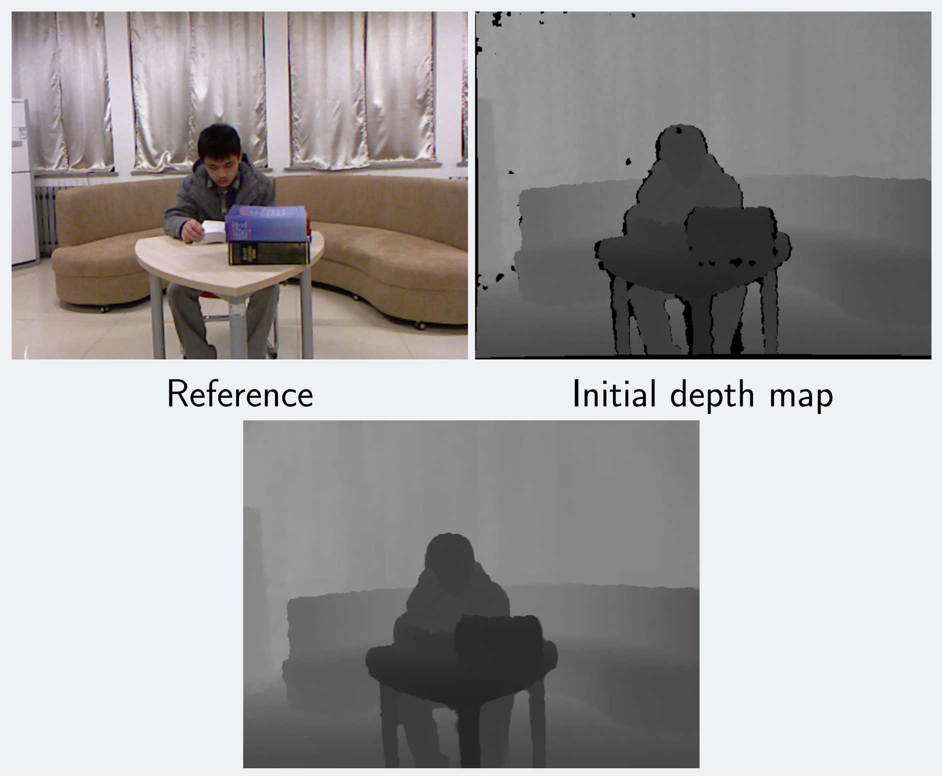

Figure: Result for depth filtering and interpolation on data from [5]. The proposed method is able to filter and interpolate the initial map preserving the geometrical structure of the objects, avoiding aliasing effects and providing sharp edges.

4. Conclusions

We have presented a new variational model for image restoration which takes advantage of image self-similarities in the data-fidelity term instead of the regularization term. We have experimentally checked that our approach outperforms classical variational methods with nonlocal total variation. Despite being a simple model, it has compared favourably with deep learning techniques for image super-resolution and it has shown robustness to high-level noise.

5. References

[1] A. Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20(1), 2004.

[2] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1), 2011.

[3] C. Dong, C.C. Loy, K. He, and X. Tang. Learning a deep convolutional network for image super-resolution. In *Proceedings of ECCV*. Springer, 2014.

[4] G. Gilboa, J. Darbon, S. Osher, and T. Chan. Nonlocal convex functionals for image regularization. *UCLA CAM-report*, 2006.

[5] J. Yang, X. Ye, K. Li, C. Hou, and Y. Wang. Color-guided depth recovery from rgb-d data using an adaptive autoregressive model. *IEEE Trans. Image Process.*, 23(8), 2014.

[6] K. Zhang, J. Liang, L. Van Gool, and R. Timofte. Designing a practical degradation model for deep blind image super-resolution. In *Proceedings of the IEEE/CVF on ICCV*, 2021.

[7] K. Zhang, W. Zuo, S. Gu, and L. Zhang. Learning deep cnn denoiser prior for image restoration. In *Proceedings of IEEE on CVPR*, 2017.

Acknowledgements

The authors were supported by Ministerio de Ciencia e Innovación under grant PID2021-125711OB-I00. (MINECO/AEI/FEDER, UE)