Supplementary Material: Semi-sparsity Priors for Image Structure Analysis and Extraction

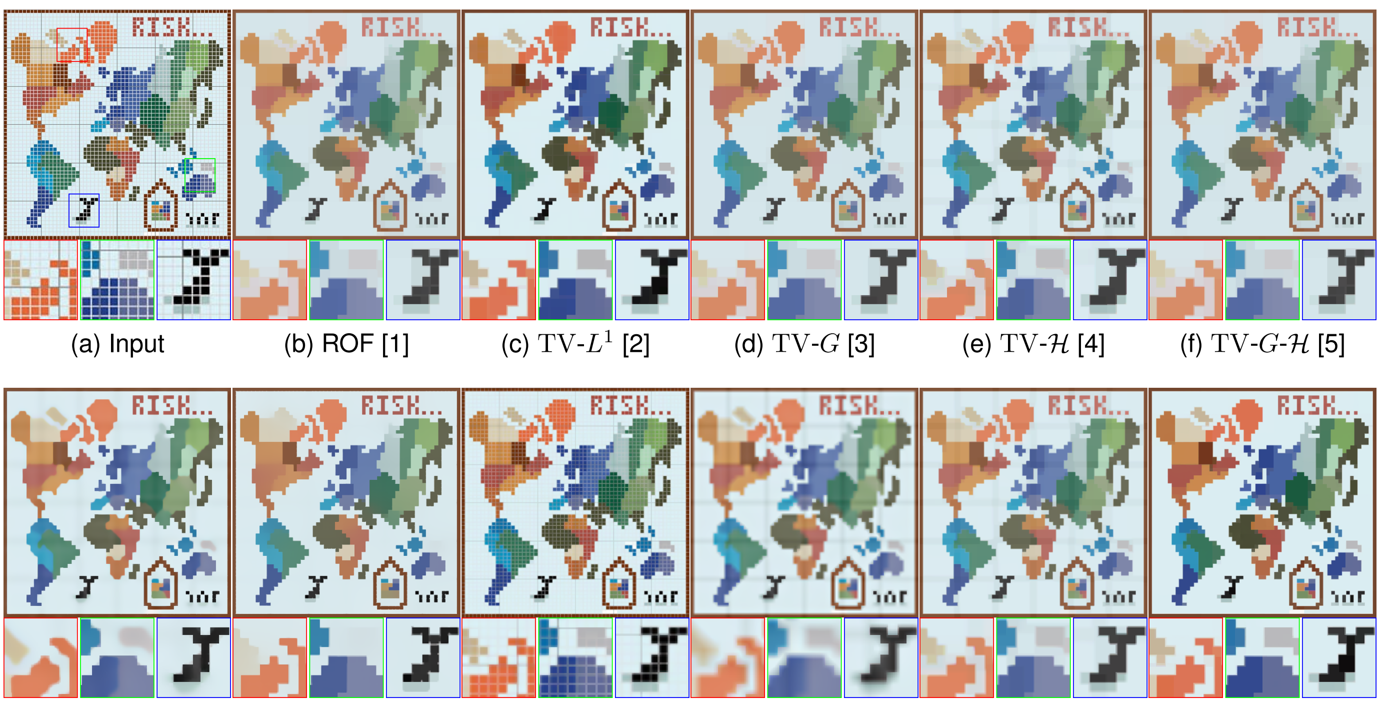
Junqing Huang, Haihui Wang, Michael Ruzhansky

1 Additional Experimental Results

We further provide some additional experimental results to show the benefits of our semi-sparse image decomposition. In Fig. 1, we compare the decomposed results in cases of piece-wise constant structures with large oscillating textures. In Fig. 2 and 3, we compare the decomposed results in cases of piece-wise constant and polynomial-smoothing structures with large oscillating textures. As shown in all cases, the proposed method produces similar comparable or superior results compared with the cutting-edge methods, which is consistent with the ones in the main paper. Additionally, we also show some results in Fig. 4, where the images are selected from our new dataset. Clearly, our semi-sparsity method is capable of decomposing image structures from a variety of textural backgrounds.

References

[1] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: nonlinear phenomena*, vol. 60, no. 1-4, pp. 259–268, 1992.
[2] V. Le Guen, "Cartoon+ texture image decomposition by the tv-l1 model," *Image Processing On Line*, vol. 4, pp. 204–219, 2014.
[3] Y. Meyer, *Oscillating patterns in image processing and nonlinear evolution equations: the fifteenth Dean Jacqueline B. Lewis memorial lectures*. American Mathematical Soc., 2001, vol. 22.



(a) Input (b) ROF [1] (c) TV-L^1 [2] (d) TV-G [3] (e) TV-$\mathcal{H}$ [4] (f) TV-G-$\mathcal{H}$ [5]

(g) BTF [6] (h) RTV [7] (i) TGV-L^1 [8] (j) TGV-$\mathcal{H}$ [9] (k) HTV-$\mathcal{H}$ [9] (l) Ours

Figure 1. Visual results of image structures in case of **piece-wise constant structures** and **large oscillating textures**. (a) Input, (b) ROF ($\lambda=0.0045, \alpha=0.001$) [1], (c) TV-$L^1$ ($\lambda=0.001, \alpha=0.002$) [2], (d) TV-$G$ ($\lambda=0.003, \alpha=0.0006, \gamma=0.0004$) [3], (e) TV-$\mathcal{H}$ ($\lambda=0.001$, $\alpha=0.01$) [4], (f) TV-G-$\mathcal{H}$ ($\lambda=0.004, \alpha=0.008, \gamma=0.005$) [5], (g) BTF ($\sigma=7.0, iter=3$) [6], (h) RTV ($\lambda=0.012, \sigma=3.0$) [7], (i) TGV-$L^1$ ($\lambda=0.001, \alpha=0.002, \beta=0.002$) [8], (j) TGV-$\mathcal{H}$ ($\lambda=0.002, \alpha=0.045, \beta=0.025$) [9], (k) HTV-$\mathcal{H}$ ($\lambda=0.004$, $\alpha=0.006, \beta=0.0015$) [9], and (l) Ours ($\lambda=0.002, \alpha=0.005, \beta=0.001$). Quantitative results with the $STR(C_0/C_1)$ metrics, (b)$\sim$(j): 22.84 (0.1908/0.0918), 22.83 (0.0550/0.1161), 22.89 (0.1771/0.0954), 22.78 (0.0392/0.0424), 22.81 (0.0416/0.0326), 22.88 (0.0271/0.0526), 22.85 (0.1371/0.0447), 22.84 (0.0077/0.1087), 22.78 (0.0370/0.1994), 22.85 (0.0507/0.0869), 23.01 (0.0774/0.0644). (Zoom in for better view.)

(a) Input (b) ROF [1] (c) TV-L^1 [2] (d) TV-G [3]
(e) TV-$\mathcal{H}$ [4] (f) TV-G-$\mathcal{H}$ [5] (g) BTF [6] (h) RTV [7]
(i) TGV-L^1 [8] (j) TGV-$\mathcal{H}$ [9] (k) HTV-$\mathcal{H}$ [9] (l) Ours

Figure 2. Visual results of image structures in case of **hybrid and complicated textural and structural components**. (a) Input, (b) ROF ($\lambda = 0.005, \alpha = 0.001$) [1], (c) TV-$L^1$ ($\lambda = 0.004, \alpha = 0.009$) [2], (d) TV-$G$ ($\lambda = 0.003, \alpha = 0.0006, \gamma = 0.0004$) [3], (e) TV-$\mathcal{H}$ ($\lambda = 0.001$, $\alpha = 0.01$) [4], (f) TV-G-$\mathcal{H}$ ($\lambda = 0.004, \alpha = 0.008, \gamma = 0.005$) [5], (g) BTF ($\sigma = 7.0, iter = 3$) [6], (h) RTV ($\lambda = 0.015, \sigma = 4.0$) [7], (i) TGV-$L^1$ ($\lambda = 0.001, \alpha = 0.002, \beta = 0.002$) [8], (j) TGV-$\mathcal{H}$ ($\lambda = 0.002, \alpha = 0.05, \beta = 0.02$) [9], (k) HTV-$\mathcal{H}$ ($\lambda = 0.004$, $\alpha = 0.002, \beta = 0.001$) [9], and (l) Ours ($\lambda = 0.006, \alpha = 0.004, \beta = 0.003$). Quantitative results with the $STR(C_0/C_1)$ metrics, (b)$\sim$(j): 18.54(0.2691/0.1004), 18.35(0.0409/0.0288), 18.60(0.2406/0.0222), 18.24(0.0897/0.0354), 18.32(0.0985/0.0235), 18.48(0.0126/0.0597), 18.72(0.0118/0.0406), 18.45(0.0908/0.1693), 18.26(0.0586/0.1220), 18.30(0.0935/0.1126), 18.51(0.0103/0.0217). (Zoom in for better view.)

[4] S. Osher, A. Solé, and L. Vese, “Image decomposition and restoration using total variation minimization and the h,” *Multiscale Modeling & Simulation*, vol. 1, no. 3, pp. 349–370, 2003.

[5] J. Xu, W. Shang, and Y. Hao, “A new cartoon+ texture image decomposition model based on the sobolev space,” *Signal, Image and Video Processing*, vol. 16, no. 6, pp. 1569–1576, 2022.

[6] H. Cho, H. Lee, H. Kang, and S. Lee, “Bilateral texture filtering,” *ACM Transactions on Graphics (TOG)*, vol. 33, no. 4, pp. 1–8, 2014.

[7] L. Xu, Q. Yan, Y. Xia, and J. Jia, “Structure extraction from texture via relative total variation,” *ACM transactions on graphics (TOG)*, vol. 31, no. 6, pp. 1–10, 2012.

[8] K. Bredies, K. Kunisch, and T. Valkonen, “Properties of l1-tgv2: The one-dimensional case,” *Journal of Mathematical Analysis and Applications*, vol. 398, no. 1, pp. 438–454, 2013.

[9] M. Jung and M. Kang, “Simultaneous cartoon and texture image restoration with higher-order regularization,” *SIAM Journal on Imaging Sciences*, vol. 8, no. 1, pp. 721–756, 2015.

(a) Input (b) ROF [1] (c) $\mathrm{TV}\text{-}L^1$ [2] (d) $\mathrm{TV}\text{-}G$ [3]

(e) $\mathrm{TV}\text{-}\mathcal{H}$ [4] (f) $\mathrm{TV}\text{-}G\text{-}\mathcal{H}$ [5] (g) BTF [6] (h) RTV [7]

(i) $\mathrm{TGV}\text{-}L^1$ [8] (j) $\mathrm{TGV}\text{-}\mathcal{H}$ [9] (k) $\mathrm{HTV}\text{-}\mathcal{H}$ [9] (l) Ours

Figure 3. Visual results of image structures in case of **hybrid and complicated textural and structural components**. (a) Input, (b) ROF ($\lambda = 0.004, \alpha = 0.001$) [1], (c) $\mathrm{TV}\text{-}L^1$ ($\lambda = 0.002, \alpha = 0.003$) [2], (d) $\mathrm{TV}\text{-}G$ ($\lambda = 0.003, \alpha = 0.0006, \gamma = 0.0003$) [3], (e) $\mathrm{TV}\text{-}\mathcal{H}$ ($\lambda = 0.004$, $\alpha = 0.03$) [4], (f) $\mathrm{TV}\text{-}G\text{-}\mathcal{H}$ ($\lambda = 0.005, \alpha = 0.004, \gamma = 0.005$) [5], (g) BTF ($\sigma = 5.0, iter = 6$) [6], (h) RTV ($\lambda = 0.015, \sigma = 3.0$) [7], (i) $\mathrm{TGV}\text{-}L^1$ ($\lambda = 0.004, \alpha = 0.007, \beta = 0.003$) [8], (j) $\mathrm{TGV}\text{-}\mathcal{H}$ ($\lambda = 0.006, \alpha = 0.05, \beta = 0.04$) [9], (k) $\mathrm{HTV}\text{-}\mathcal{H}$ ($\lambda = 0.005$, $\alpha = 0.015, \beta = 0.01$) [9], and (l) Ours ($\lambda = 0.0007, \alpha = 0.0008, \beta = 0.0005$). Quantitative results with the $STR(C_0/C_1)$ metrics, (b)$\sim$(j): 18.22(0.1650/0.0251), 18.15(0.0295/0.0446), 18.19(0.1473/0.0120), 18.14(0.0696/0.0432), 18.21(0.0690/0.0453), 18.19(0.0350/0.0835), 18.24(0.0959/0.0207), 18.23(0.0106/0.1038), 18.18(0.0493/0.2023), 18.23(0.0659/0.1370), 18.17(0.0321/0.0065). (Zoom in for better view.)

Figure 4. Visual results of image structures in case of an image consisting of complicated textural and structural components. (Zoom in for better view).