Supplementary Material for "Deep Gaussian Scale Mixture Prior for Spectral Compressive Imaging"

Tao Huang[1] Weisheng Dong[1]* Xin Yuan[2]* Jinjian Wu[1] Guangming Shi[1]
[1]School of Artificial Intelligence, Xidian University [2]Bell Labs
thuang_666@stu.xidian.edu.cn wsdong@mail.xidian.edu.cn xyuan@bell-labs.com
jinjian.wu@mail.xidian.edu.cn gmshi@xidian.edu.cn

In this supplementary material, we provide the RGB images of the testing scenes, more visualization results of the regularization parameters $\boldsymbol{w}$ and more visual comparison results on both synthetic data and real data.

The benchmark methods in our comparison include: three model-based hyperspectral image (HSI) reconstruction methods (i.e., TwIST [1], GAP-TV [8] and DeSCI [2]) and four deep learning based methods (i.e., λ-net [4], HSSP [5], DNU [6] and TSA-Net [3]). The peak-signal-to-noise (PSNR) and the structural similarity index (SSIM) [7] are employed to evaluate the performance of competing HSI reconstruction methods.

1. RGB images of the testing scenes and the regularization parameters $\boldsymbol{w}$

Fig. 1 shows the RGB images of the 10 scenes and its corresponding regularization parameters $\boldsymbol{w}$ which were estimated in the fourth stage. From Fig. 1, we can see that the values of $\boldsymbol{w}$ are consistent with the image edges and textures. Aided by this well-learned $\boldsymbol{w}$, the proposed method will pay more attentions to the edges and textures.

2. More visual comparison results on synthetic data

Fig. 2-11 show more visual comparison results of the best five competing methods with 28 spectral channels for 10 testing scenes. Ground truth, measurements, and RGB images are shown for reference. We compare the proposed methods with TSA-Net [3], DNU [6], HSSP [5] and λ-net [4]. From Fig. 2-11, it can be observed the proposed method can achieve high reconstruction quality and recover more details of the textures and edges than the other competing methods.

* Corresponding authors.

3. More visual comparison results on real data

Fig. 12-16 show more visual comparison results with 28 spectral channels for the 5 real scenes. We compare the proposed methods with TSA-Net [3], DeSCI [2], GAP-TV [8] and TwIST [1]. From Fig. 12-16, we can see that the proposed method can better suppress undesirable visual artifacts and recover more details of the textures and fine structures.

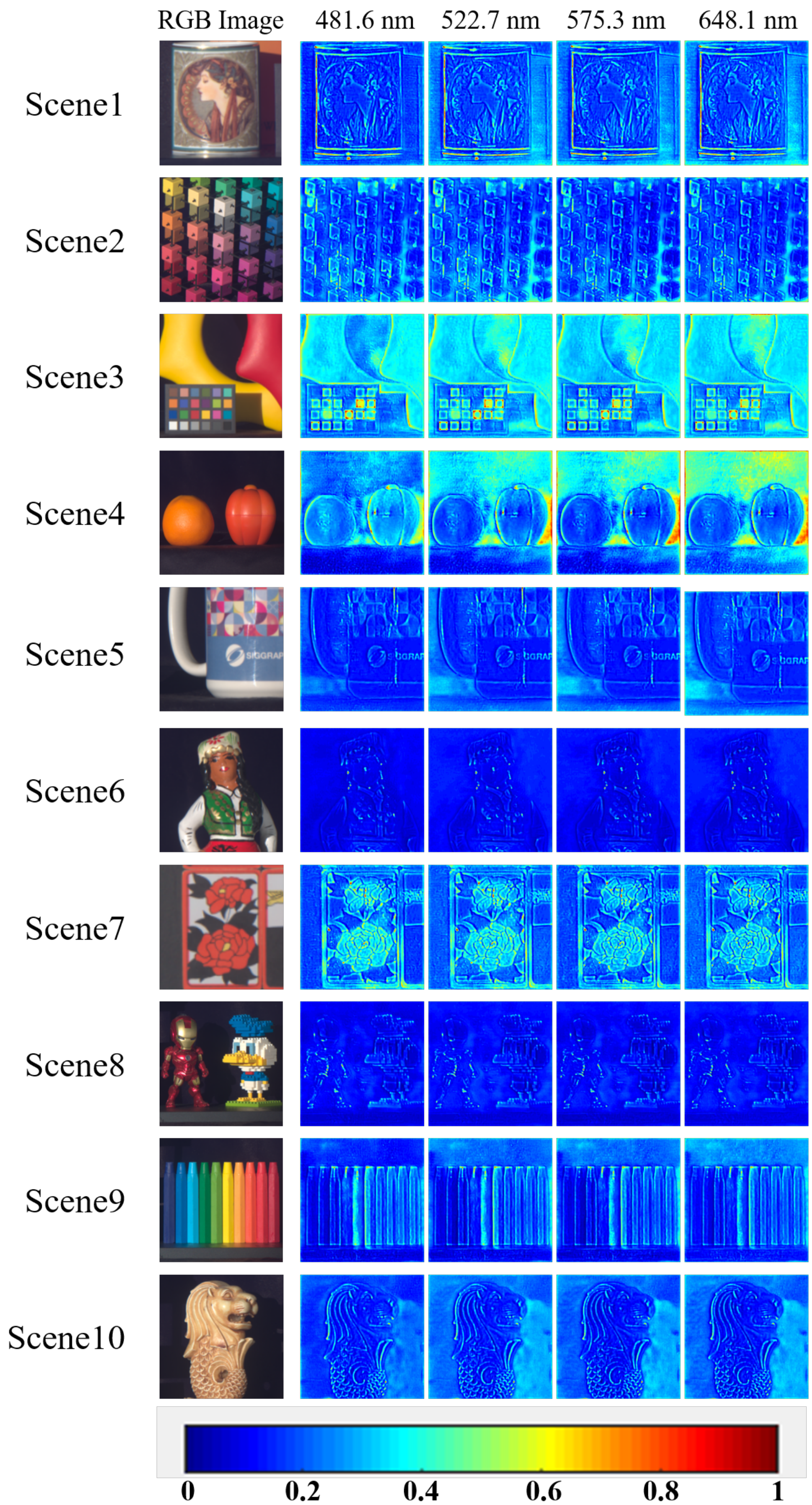

Figure 1. The RGB images of the 10 scenes and the visualization of the regularization parameters $\boldsymbol{w}$ estimated in the 4-th stage. Left: the corresponding RGB image; right: the $\boldsymbol{w}$ images associated with the four spectral bands (with normalization).

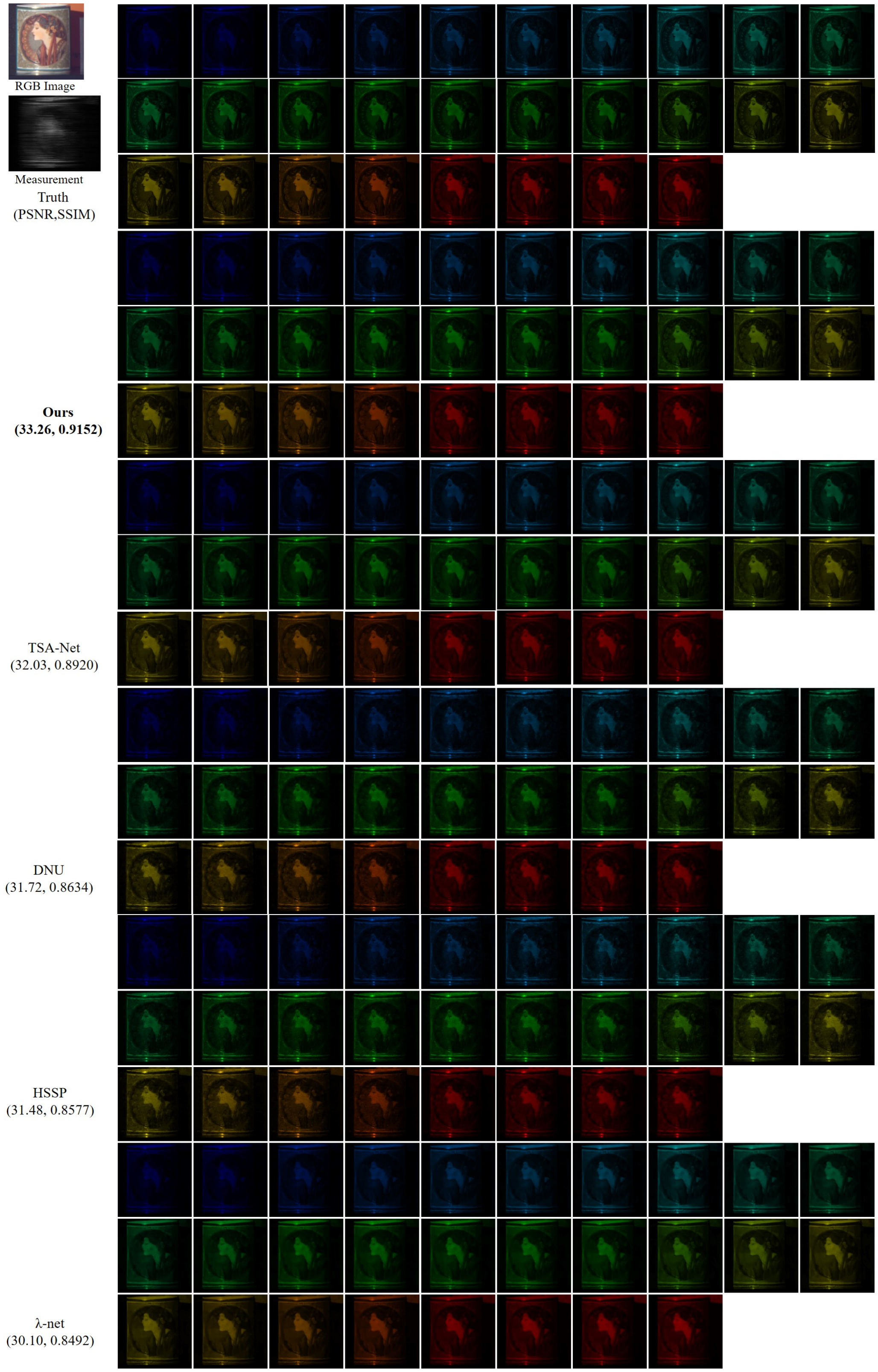

Figure 2. Simulation: RGB image, measurement, ground truth and reconstructed results by the proposed method (PSNR = **33.26**dB, SSIM = **0.9152**), TSA-Net [3] (PSNR = 32.03dB, SSIM = 0.8920), DNU [6] (PSNR = 31.72dB, SSIM = 0.8634), HSSP [5] (PSNR = 31.48dB, SSIM = 0.8577) and λ-net [4] (PSNR = 30.10dB, SSIM = 0.8492) for *Scene1*. Zoom in for better view.

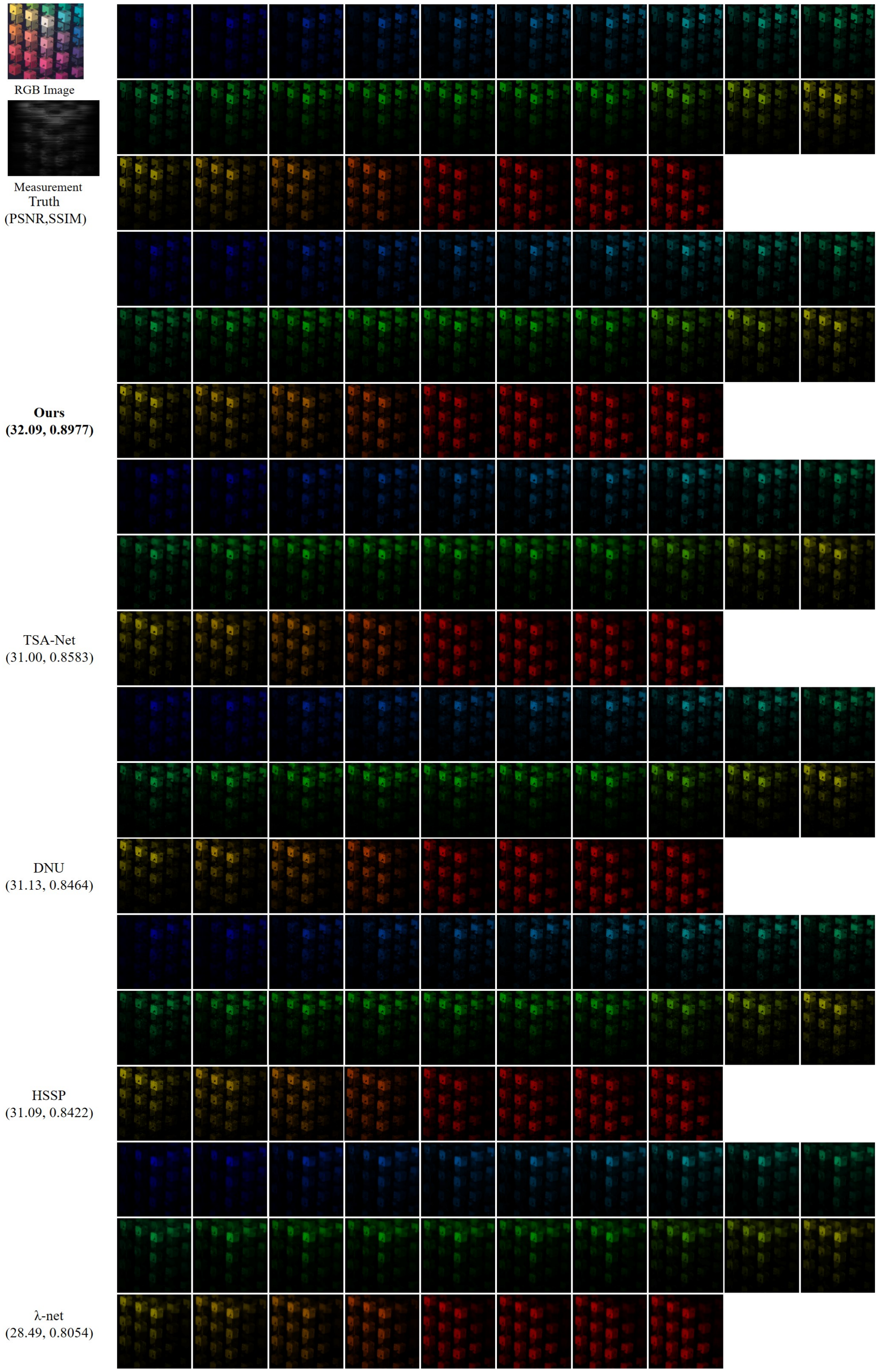

Figure 3. Simulation: RGB image, measurement, ground truth and reconstructed results by the proposed method (PSNR = **32.09**dB, SSIM = **0.8977**), TSA-Net [3] (PSNR = 31.00dB, SSIM = 0.8583), DNU [6] (PSNR = 31.13dB, SSIM = 0.8464), HSSP [5] (PSNR = 31.09dB, SSIM = 0.8422) and λ-net [4] (PSNR = 28.49dB, SSIM = 0.8054) for *Scene2*. Zoom in for better view.

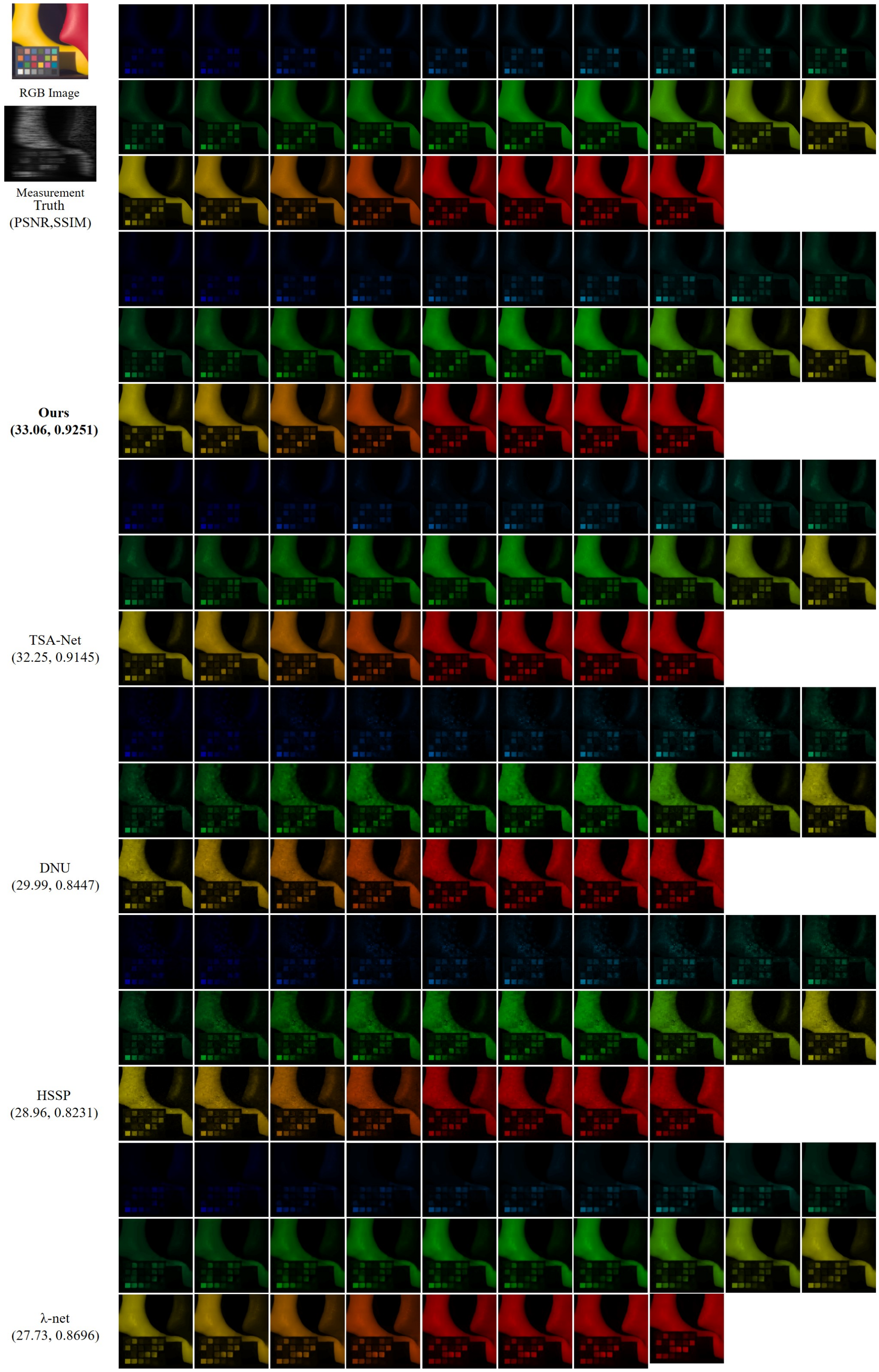

Figure 4. Simulation: RGB image, measurement, ground truth and reconstructed results by the proposed method (PSNR = **33.06**dB, SSIM = **0.9251**), TSA-Net [3] (PSNR = 32.25dB, SSIM = 0.9145), DNU [6] (PSNR = 29.99dB, SSIM = 0.8447), HSSP [5] (PSNR = 28.96dB, SSIM = 0.8231) and λ-net [4] (PSNR = 27.73dB, SSIM = 0.8696) for *Scene3*. Zoom in for better view.

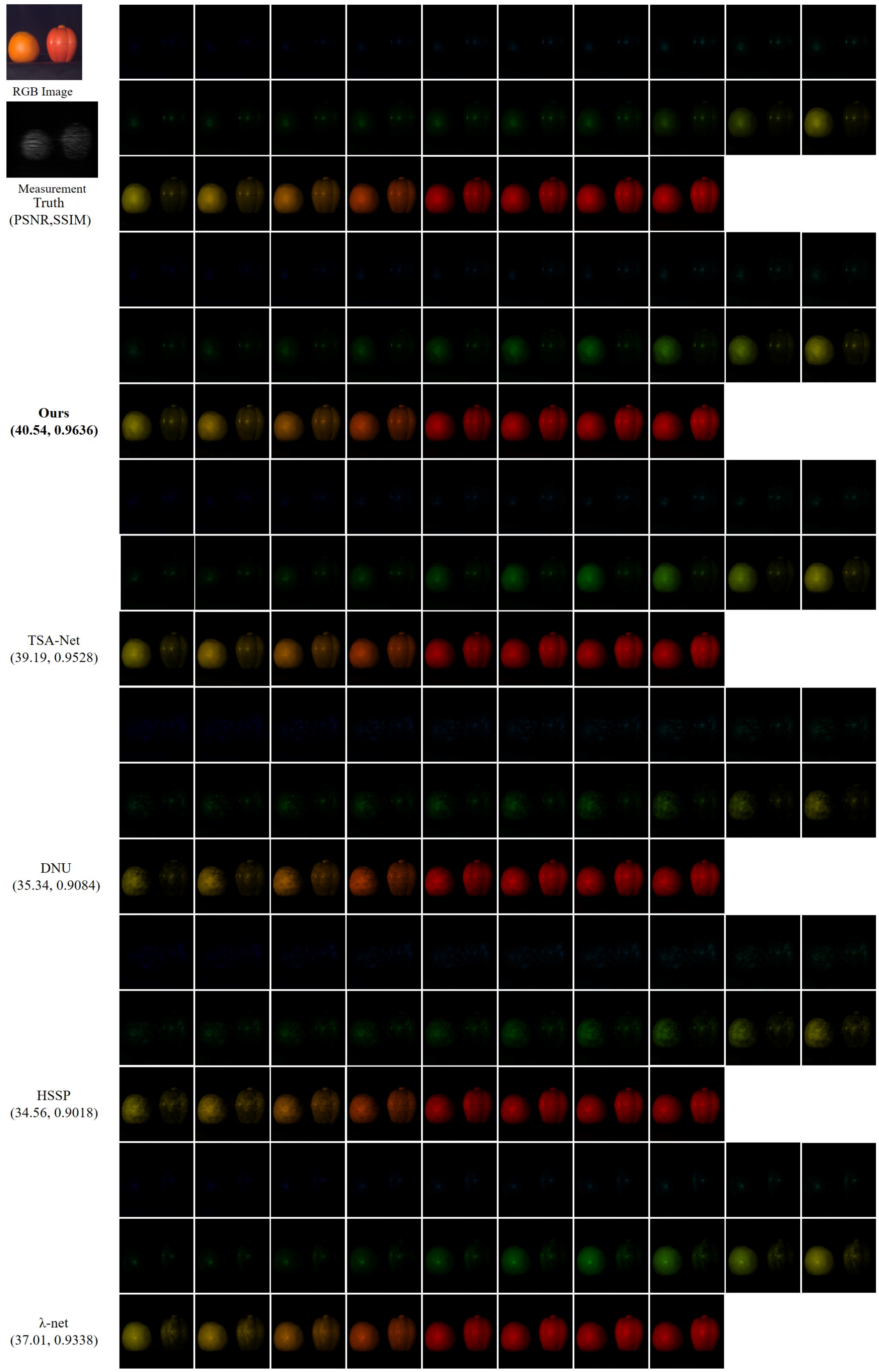

Figure 5. Simulation: RGB image, measurement, ground truth and reconstructed results by the proposed method (PSNR = **40.54**dB, SSIM = **0.9636**), TSA-Net [3] (PSNR = 39.19dB, SSIM = 0.9528), DNU [6] (PSNR = 35.34dB, SSIM = 0.9084), HSSP [5] (PSNR = 34.56dB, SSIM = 0.9018) and λ-net [4] (PSNR = 37.01dB, SSIM = 0.9338) for *Scene4*. Zoom in for better view.

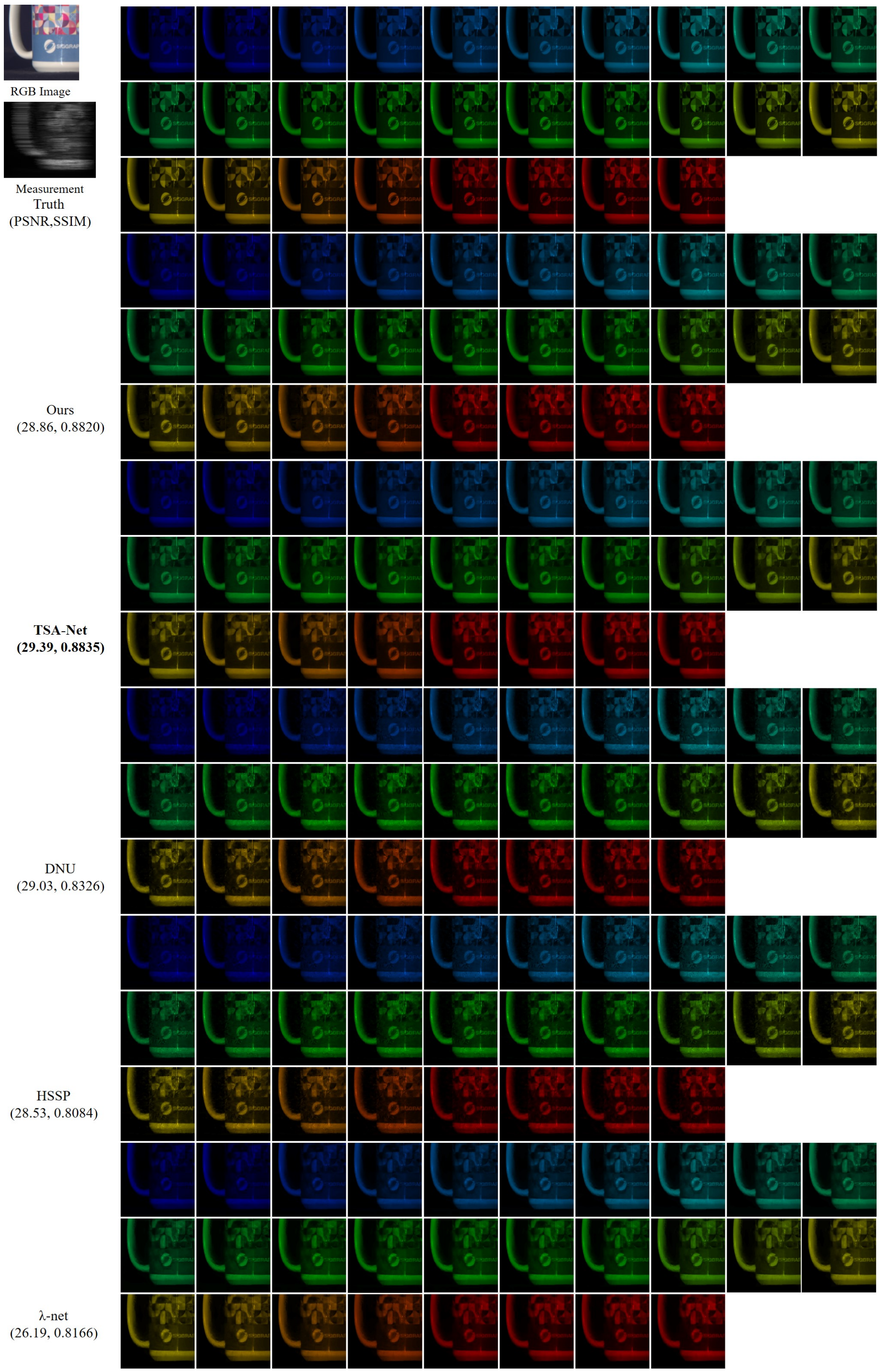

Figure 6. Simulation: RGB image, measurement, ground truth and reconstructed results by the proposed method (PSNR = 28.86dB, SSIM = 0.8820), TSA-Net [3] (PSNR = **29.39**dB, SSIM = **0.8835**), DNU [6] (PSNR = 29.03dB, SSIM = 0.8326), HSSP [5] (PSNR = 28.53dB, SSIM = 0.8084) and λ-net [4] (PSNR = 26.19dB, SSIM = 0.8166) for *Scene5*. Zoom in for better view.

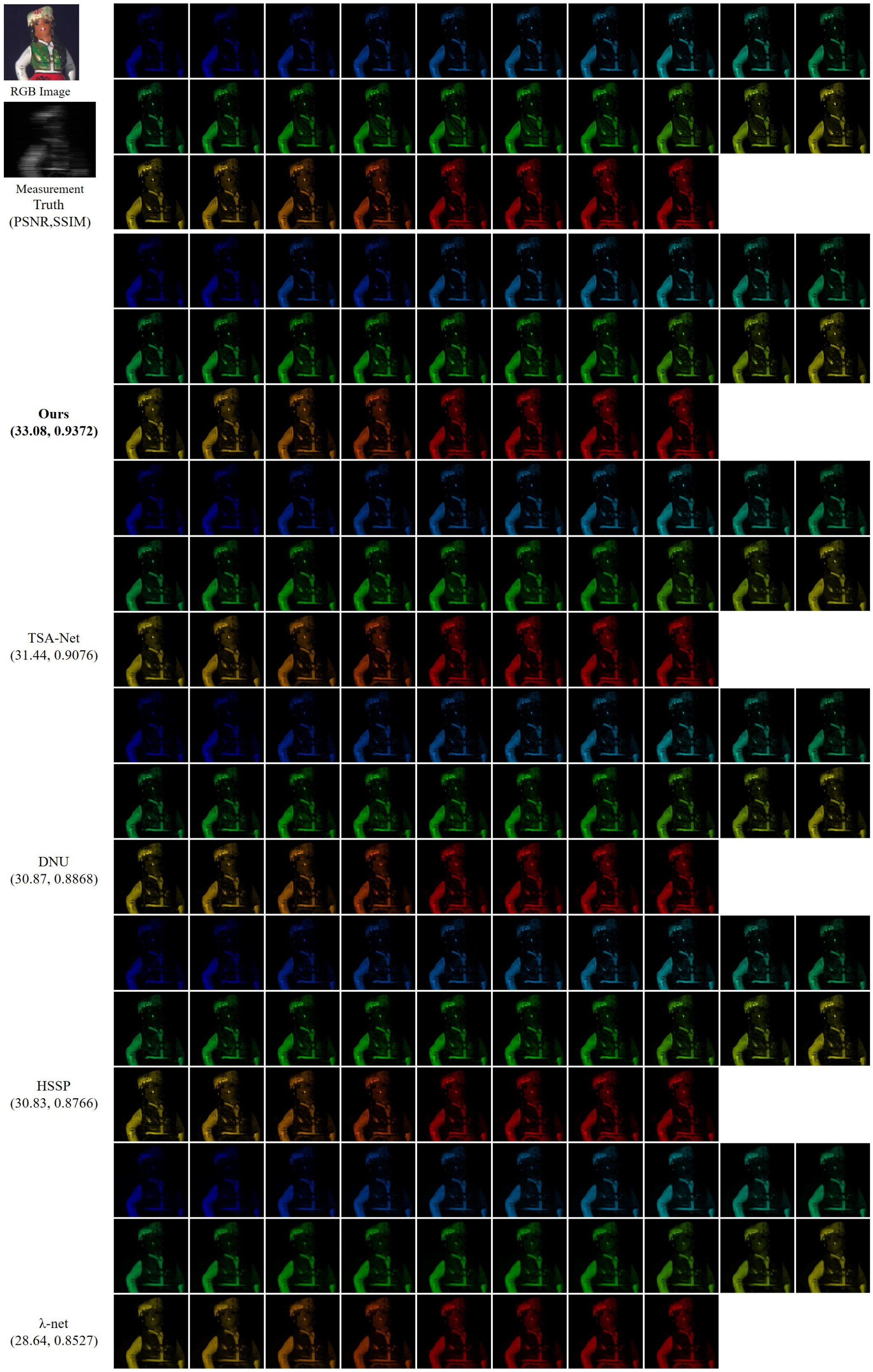

Figure 7. Simulation: RGB image, measurement, ground truth and reconstructed results by the proposed method (PSNR = **33.08**dB, SSIM = **0.9372**), TSA-Net [3] (PSNR = 31.44dB, SSIM = 0.9076), DNU [6] (PSNR = 30.87dB, SSIM = 0.8868), HSSP [5] (PSNR = 30.83dB, SSIM = 0.8766) and λ-net [4] (PSNR = 28.64dB, SSIM = 0.8527) for *Scene6*. Zoom in for better view.

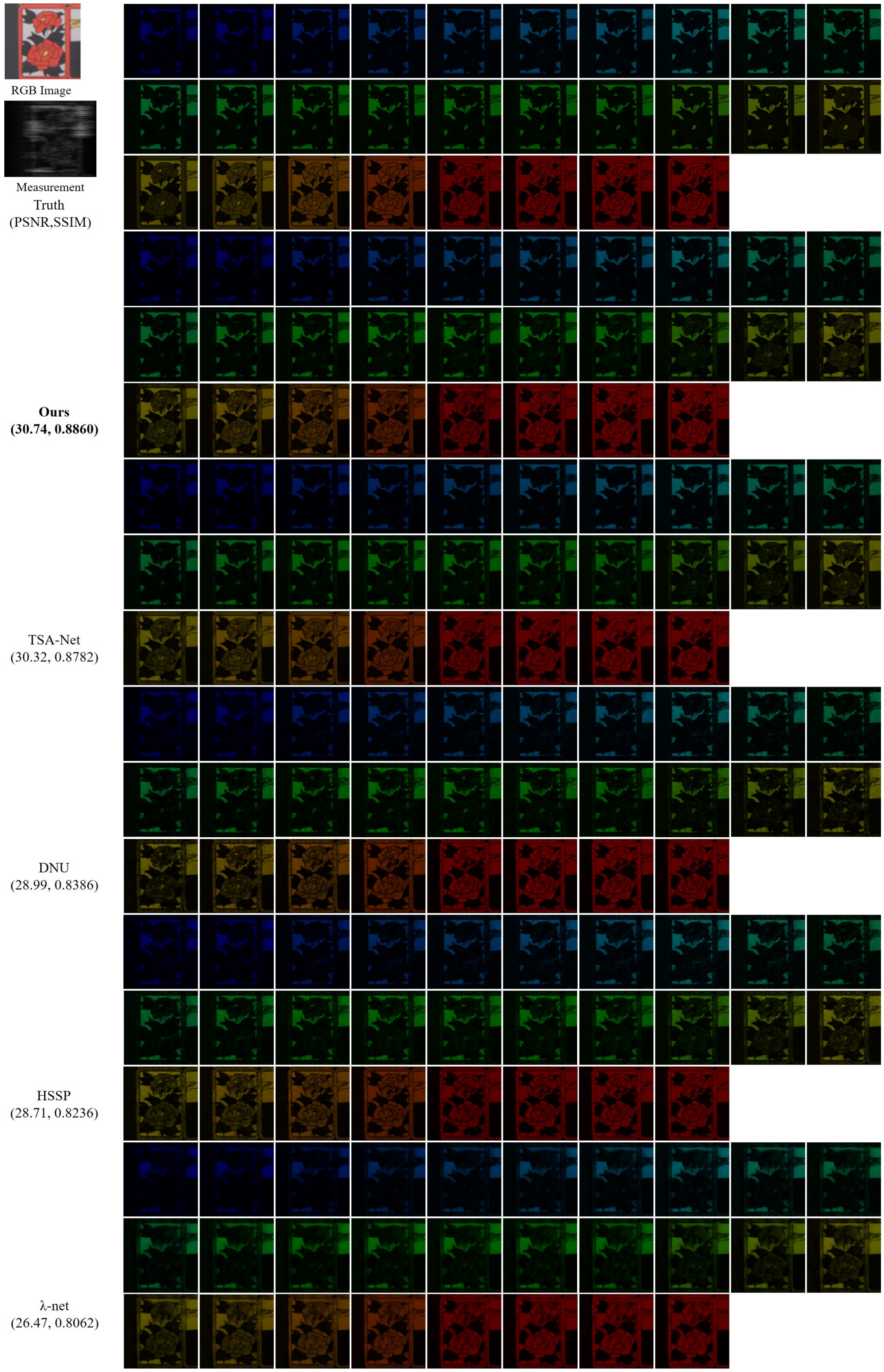

Figure 8. Simulation: RGB image, measurement, ground truth and reconstructed results by the proposed method (PSNR = **30.74**dB, SSIM = **0.8860**), TSA-Net [3] (PSNR = 30.32dB, SSIM = 0.8782), DNU [6] (PSNR = 28.99dB, SSIM = 0.8386), HSSP [5] (PSNR = 28.71dB, SSIM = 0.8236) and λ-net [4] (PSNR = 26.47dB, SSIM = 0.8062) for *Scene7*. Zoom in for better view.

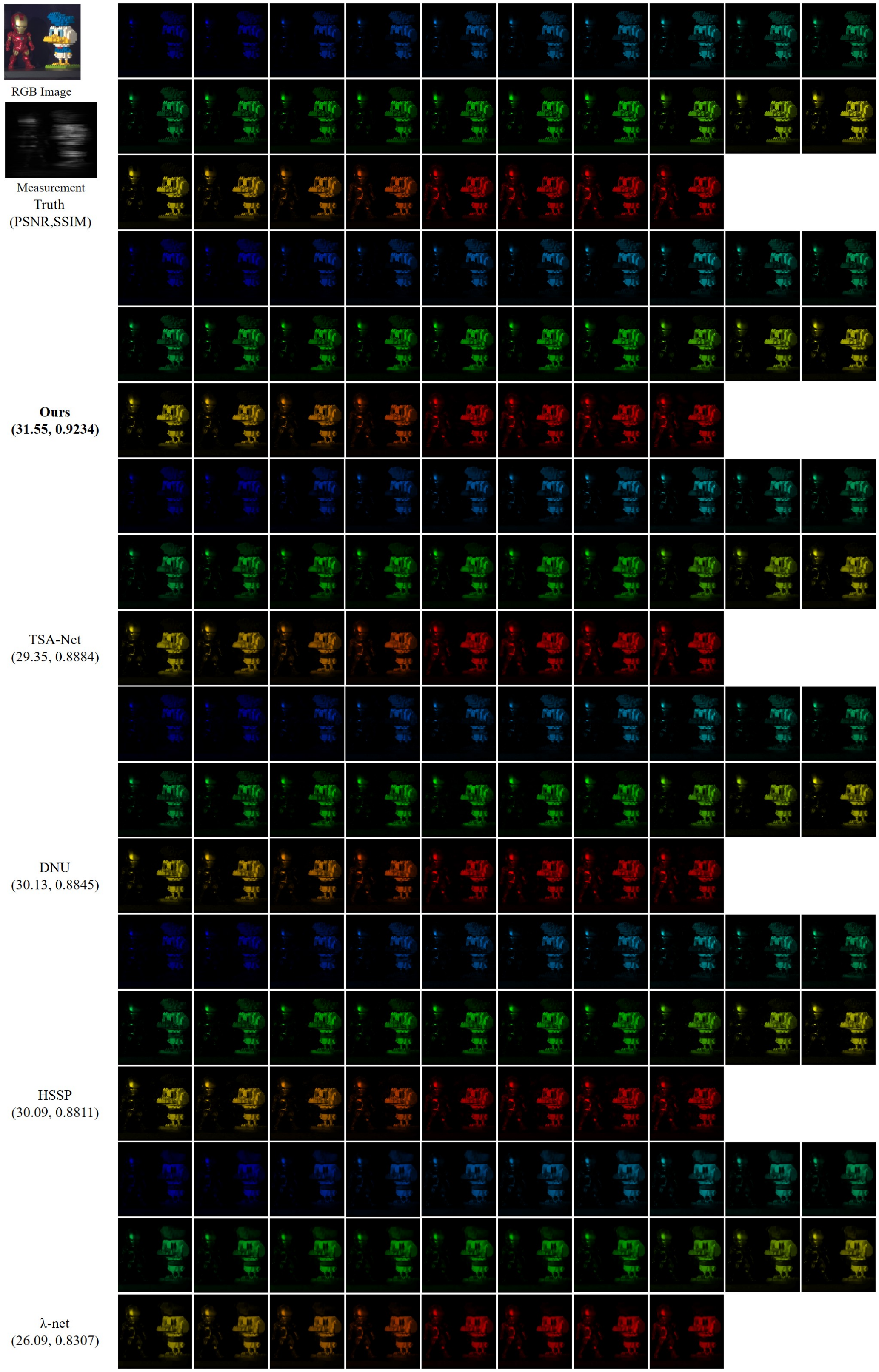

Figure 9. Simulation: RGB image, measurement, ground truth and reconstructed results by the proposed method (PSNR = **31.55**dB, SSIM = **0.9234**), TSA-Net [3] (PSNR = 29.35dB, SSIM = 0.8884), DNU [6] (PSNR = 30.13dB, SSIM = 0.8845), HSSP [5] (PSNR = 30.09dB, SSIM = 0.8811) and λ-net [4] (PSNR = 26.09dB, SSIM = 0.8307) for *Scene8*. Zoom in for better view.

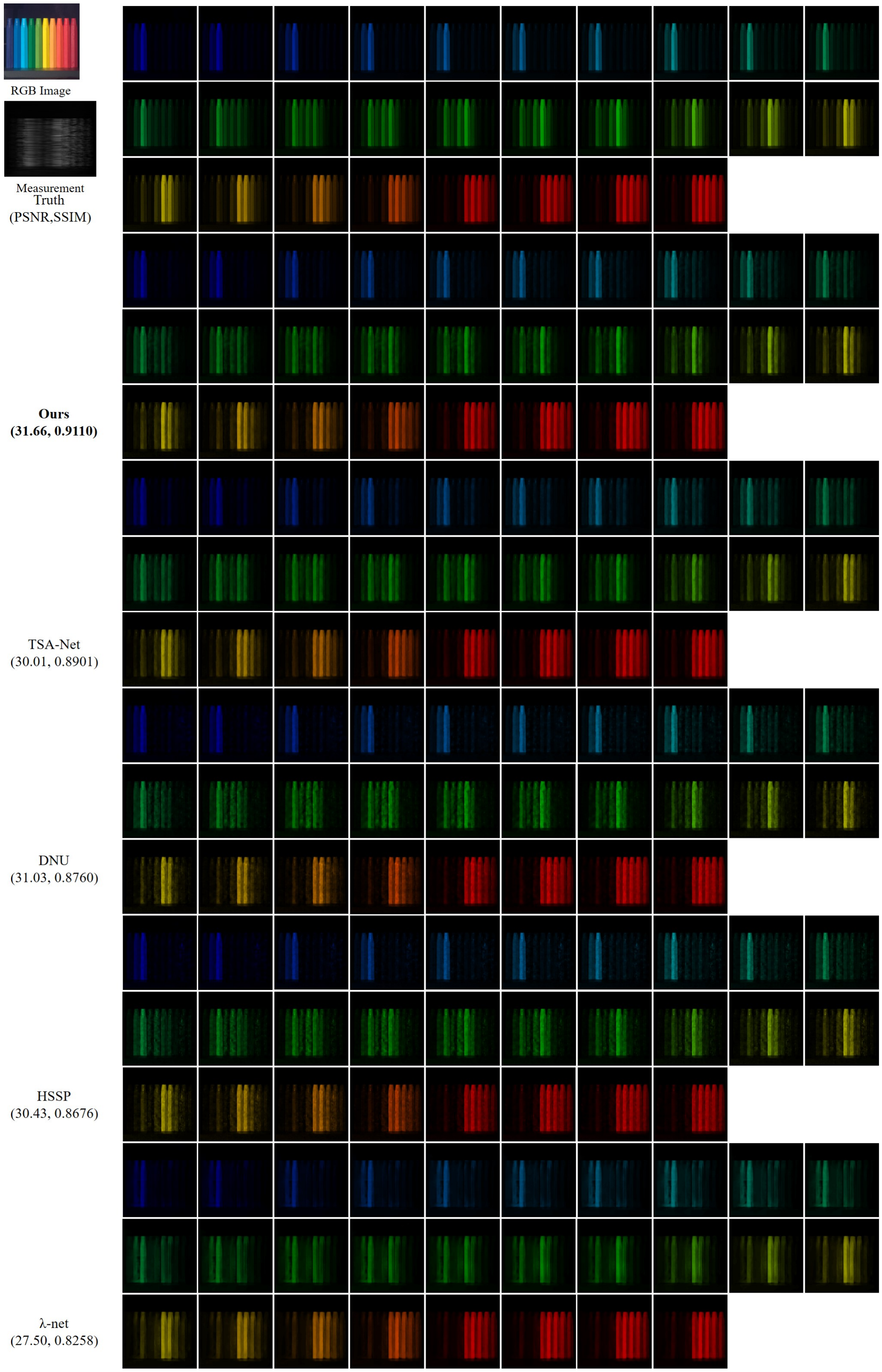

Figure 10. Simulation: RGB image, measurement, ground truth and reconstructed results by the proposed method (PSNR = **31.66**dB, SSIM = **0.9110**), TSA-Net [3] (PSNR = 30.01dB, SSIM = 0.8901), DNU [6] (PSNR = 31.03dB, SSIM = 0.8760), HSSP [5] (PSNR = 30.43dB, SSIM = 0.8676) and λ-net [4] (PSNR = 27.50dB, SSIM = 0.8258) for *Scene9*. Zoom in for better view.

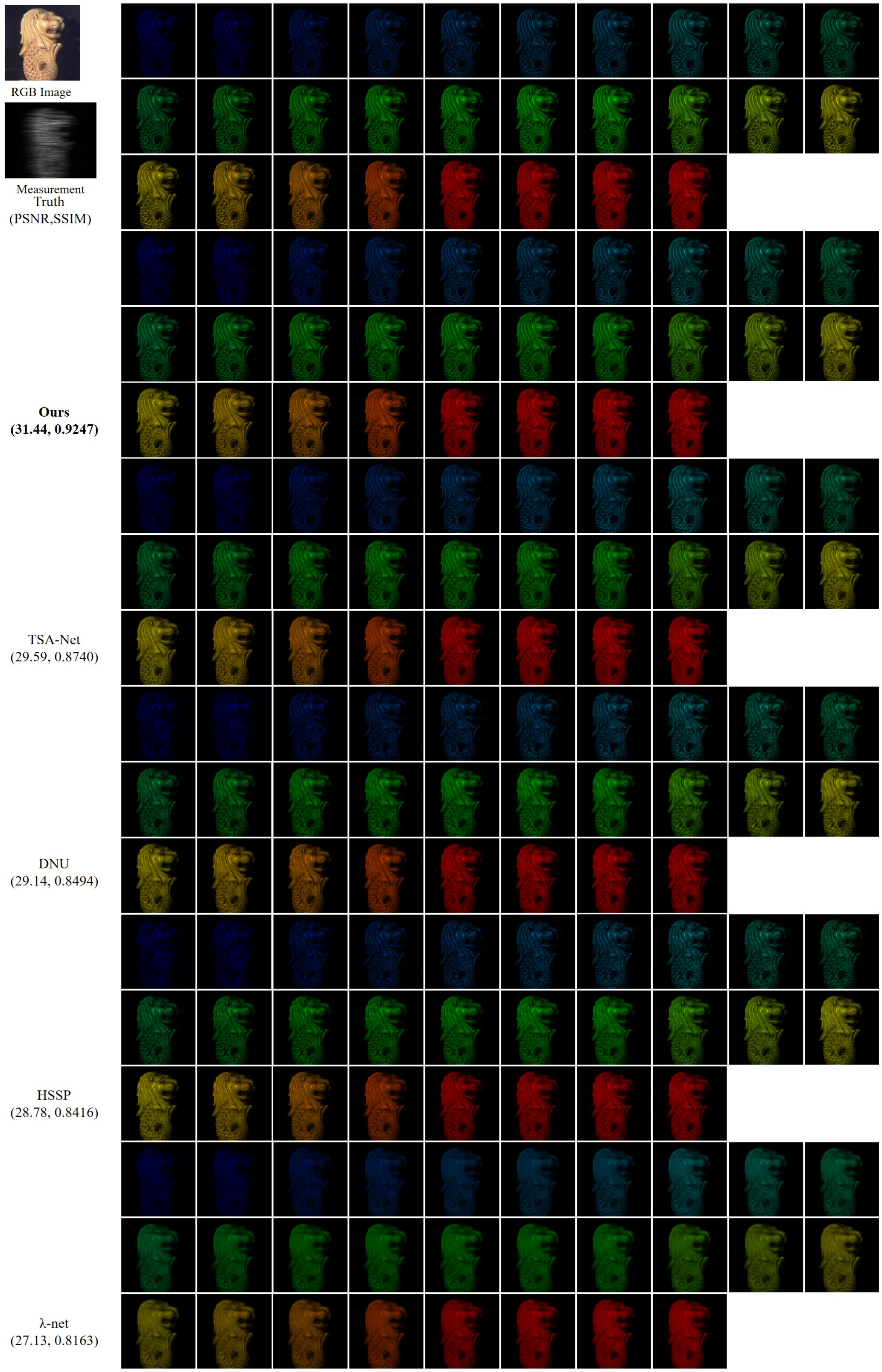

Figure 11. Simulation: RGB image, measurement, ground truth and reconstructed results by the proposed method (PSNR = **31.44**dB, SSIM = **0.9247**), TSA-Net [3] (PSNR = 29.59dB, SSIM = 0.8740), DNU [6] (PSNR = 29.14dB, SSIM = 0.8494), HSSP [5] (PSNR = 28.78dB, SSIM = 0.8416) and λ-net [4] (PSNR = 27.13dB, SSIM = 0.8163) for *Scene10*. Zoom in for better view.

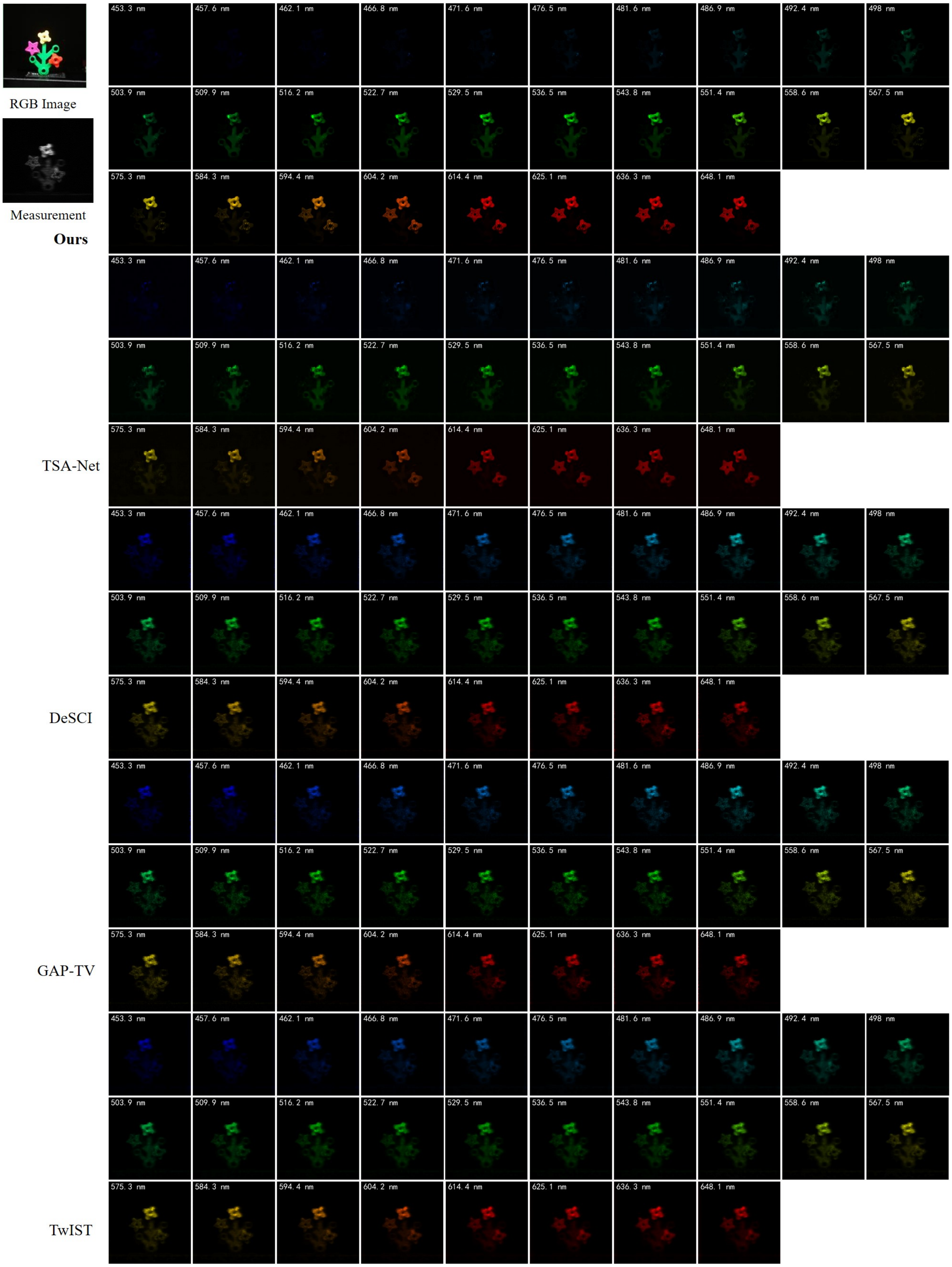

Figure 12. Real data: RGB image, measurement and reconstructed results by the proposed method, TSA-Net [3], DeSCI [2], GAP-TV [8] and TwIST [1] for *Scene1*. Zoom in for better view.

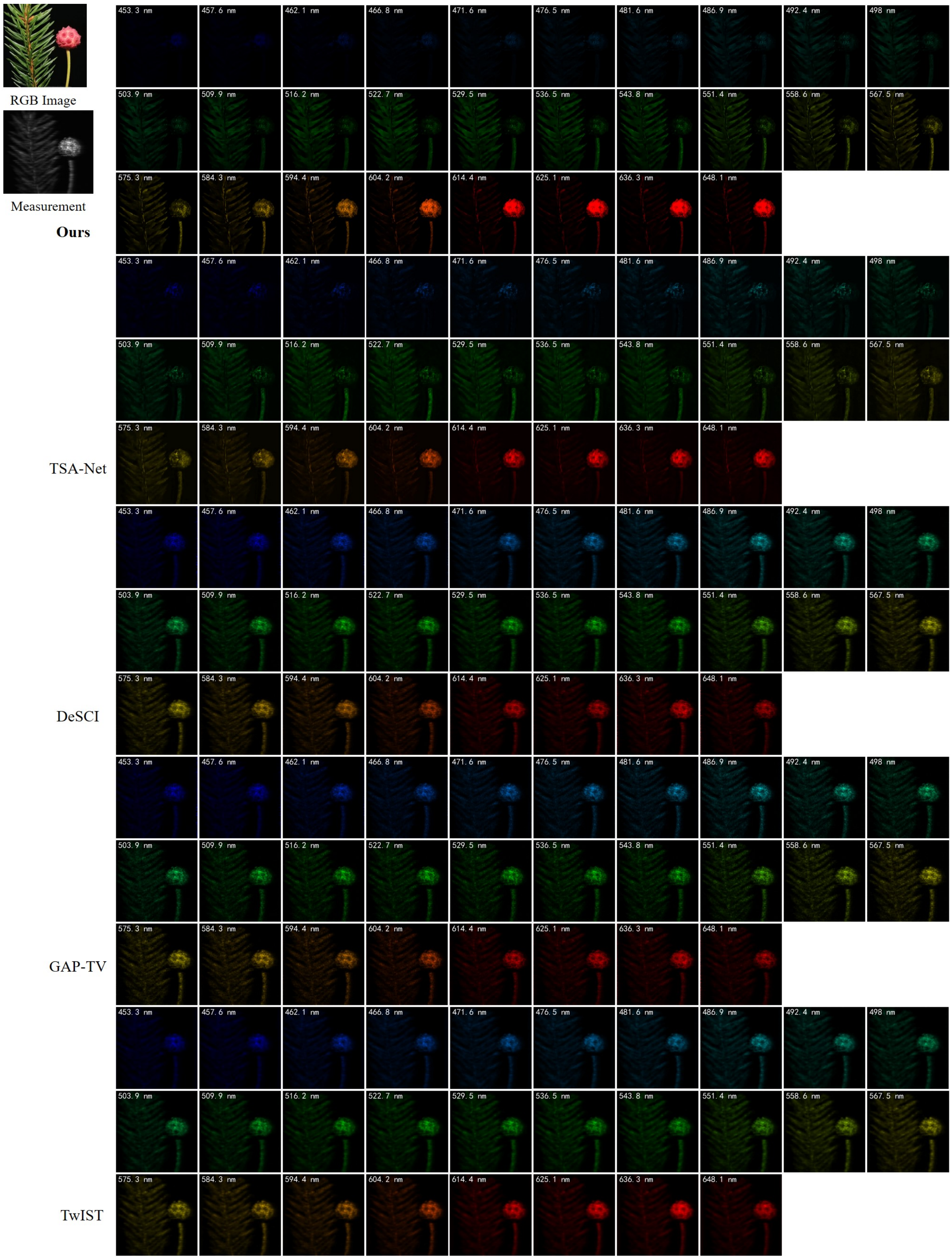

Figure 13. Real data: RGB image, measurement and reconstructed results by the proposed method, TSA-Net [3], DeSCI [2], GAP-TV [8] and TwIST [1] for *Scene2*. Zoom in for better view.

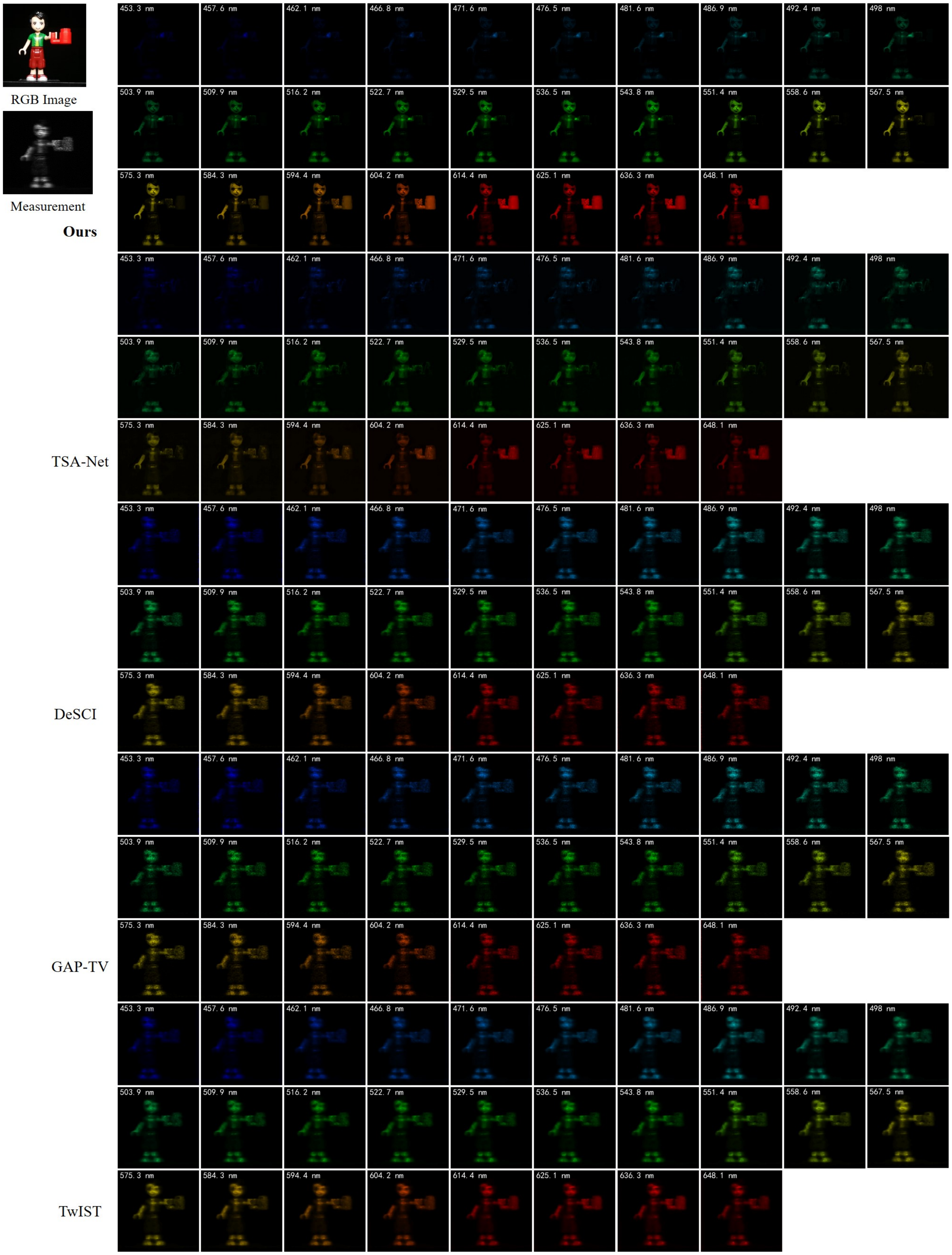

Figure 14. Real data: RGB image, measurement and reconstructed results by the proposed method, TSA-Net [3], DeSCI [2], GAP-TV [8] and TwIST [1] for *Scene3*. Zoom in for better view.

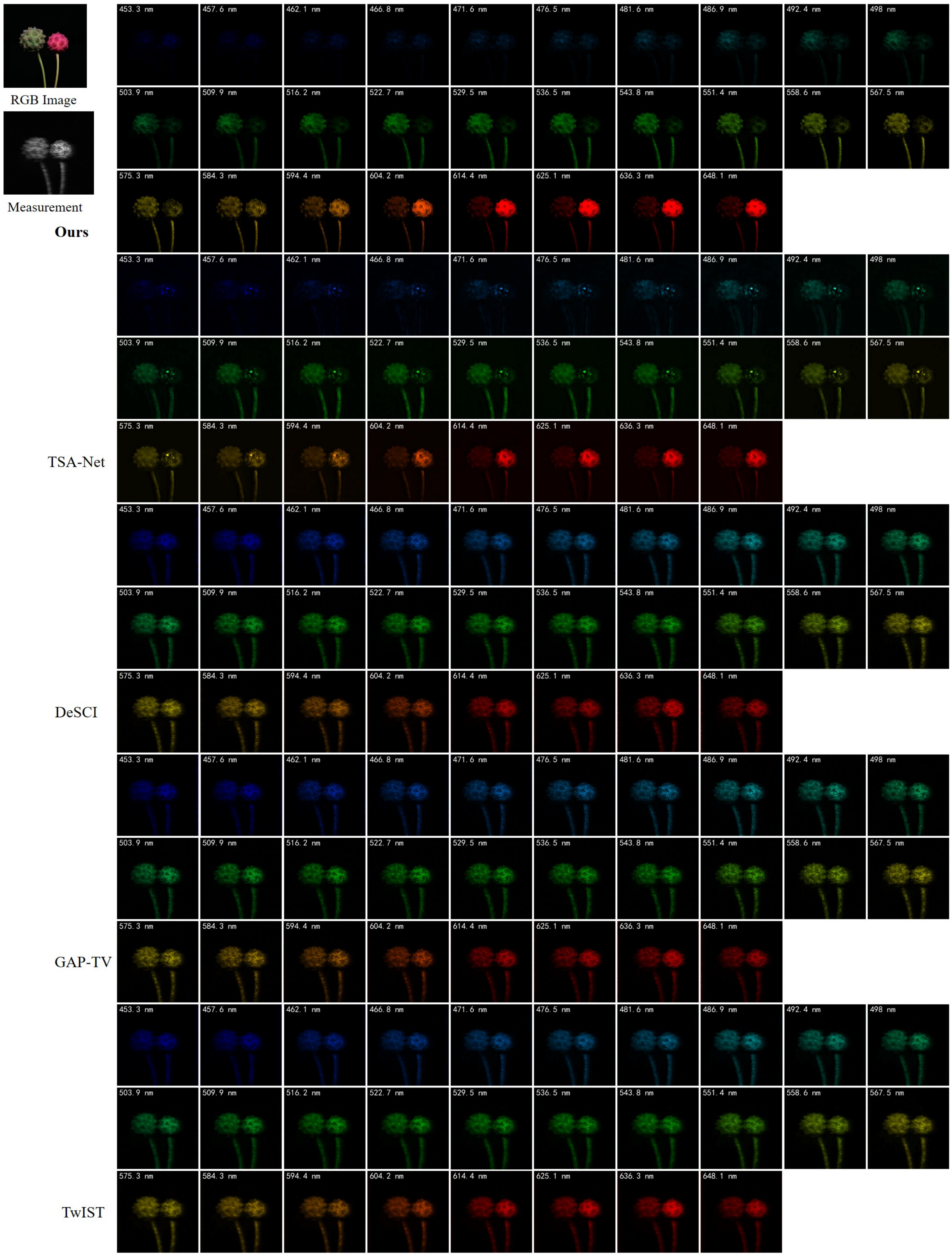

Figure 15. Real data: RGB image, measurement and reconstructed results by the proposed method, TSA-Net [3], DeSCI [2], GAP-TV [8] and TwIST [1] for *Scene4*. Zoom in for better view.

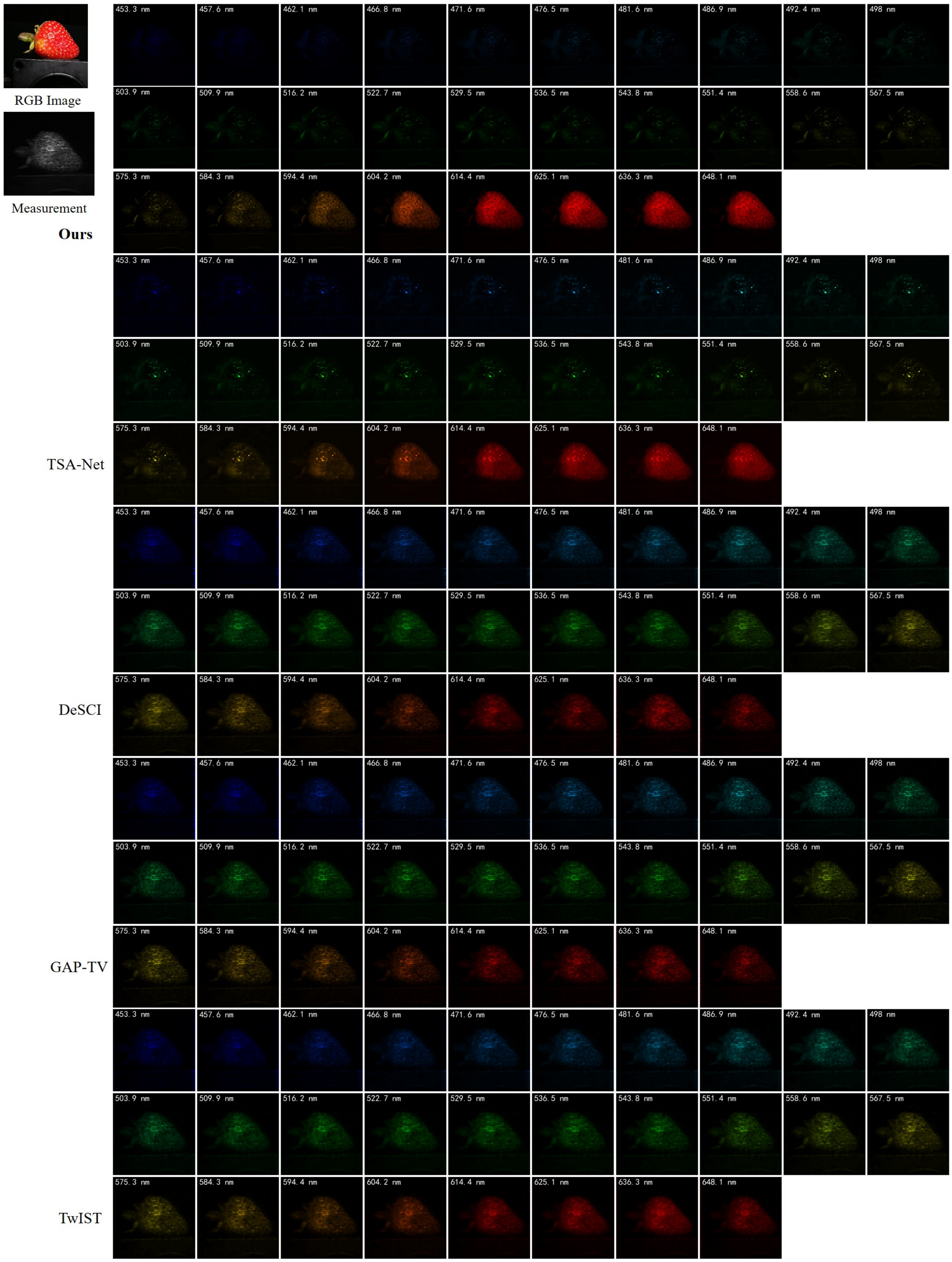

Figure 16. Real data: RGB image, measurement and reconstructed results by the proposed method, TSA-Net [3], DeSCI [2], GAP-TV [8] and TwIST [1] for *Scene5*. Zoom in for better view.

References

[1] José M Bioucas-Dias and Mário AT Figueiredo. A new twist: Two-step iterative shrinkage/thresholding algorithms for image restoration. *IEEE Transactions on Image processing*, 16(12):2992–3004, 2007. 1, 13, 14, 15, 16, 17

[2] Yang Liu, Xin Yuan, Jinli Suo, David J Brady, and Qionghai Dai. Rank minimization for snapshot compressive imaging. *IEEE transactions on pattern analysis and machine intelligence*, 41(12):2990–3006, 2018. 1, 13, 14, 15, 16, 17

[3] Ziyi Meng, Jiawei Ma, and Xin Yuan. End-to-end low cost compressive spectral imaging with spatial-spectral self-attention. In *European Conference on Computer Vision*, pages 187–204. Springer, 2020. 1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17

[4] Xin Miao, Xin Yuan, Yunchen Pu, and Vassilis Athitsos. lambda-net: Reconstruct hyperspectral images from a snapshot measurement. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 4058–4068. IEEE, 2019. 1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

[5] Lizhi Wang, Chen Sun, Ying Fu, Min H Kim, and Hua Huang. Hyperspectral image reconstruction using a deep spatial-spectral prior. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8032–8041, 2019. 1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

[6] Lizhi Wang, Chen Sun, Maoqing Zhang, Ying Fu, and Hua Huang. Dnu: Deep non-local unrolling for computational spectral imaging. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1661–1671, 2020. 1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

[7] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004. 1

[8] Xin Yuan. Generalized alternating projection based total variation minimization for compressive sensing. In *2016 IEEE International Conference on Image Processing (ICIP)*, pages 2539–2543. IEEE, 2016. 1, 13, 14, 15, 16, 17