DCT-based residual network for NIR image colorization

Hongcheng Jiang[1], Paras Maharjan[1], Zhu Li[1], Gorge York[2]
[1]University of Missouri-Kansas City, USA,
[2]US Air Force Academy

Objectives

- We propose a DCT-based residual network (DCT-RCAN) for NIR image colorization. Instead of stacking large-scale models with a large amount of computation and enormous parameters, we adopt lightweight but an effective module RIR to improve the efficiency of NIR image colorization without sacrificing much performance.
- Extensive experiments validate that our DCT-RCAN is computationally efficient and demonstrate competitive results against state-of-the-art NIR image colorization methods.

Proposed Methods

- We use the four subgroups generated by the 4×4 DCT of RGB images as the output, so that the coarse contents and sharp details are separated explicitly during training. The size of subgroups is 1/4th of the original image, which alleviates the learning difficulty of our network without information loss. The DCT and its inverse operation are both invertible, leading to no information loss. Thus, our network can easily generate the RGB images via the inverse DCT.
- We exploit the RCAN which includes shallow feature extraction and deep feature extraction. DCT can be seamlessly integrated to RCAN and add trivial computational cost. In addition, our DCT-RCAN can focus on more crucial underlying patterns in channel dimension in a lightweight manner.

Network Architecture

- The input of our model is the NIR image. Three convolutional layers (Conv) are used to extract the shallow feature. Then we stack N Residual in Residual (RIR) modules to form a deep network for the deep feature extraction, which allows abundant low-frequency information to be bypassed through multiple skip connections, making the main network focus on learning high-frequency information.

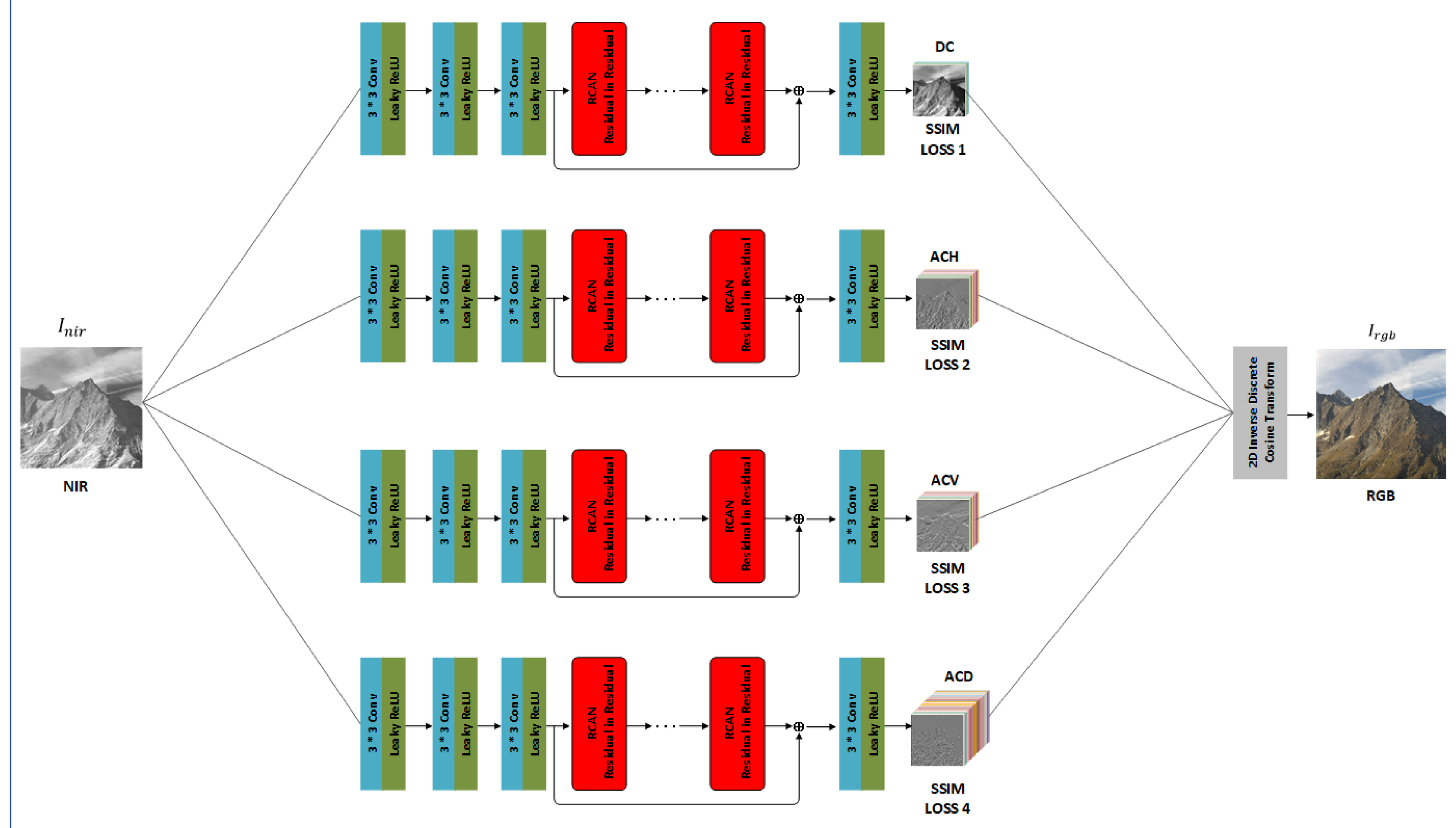

Figure 1. Proposed Network Structure

DCT Decomposition

- We used 4×4 DCT to decompose the compressed image and subsampled the resulting DCT image to create 16 sub-band images, each coefficient representing DC, AC1, AC2,..., AC15. DC coefficient represents the low-frequency component and has global information of the image. AC1 to AC15 represents high-frequency components of the image. Particularly, AC1, AC2, AC3 coefficients represent the horizontal subgroup (ACH), AC4, AC8, AC12 coefficient represent the vertical subgroup (ACV), and the rest of the AC coefficients represent the diagonal subgroup (ACD).

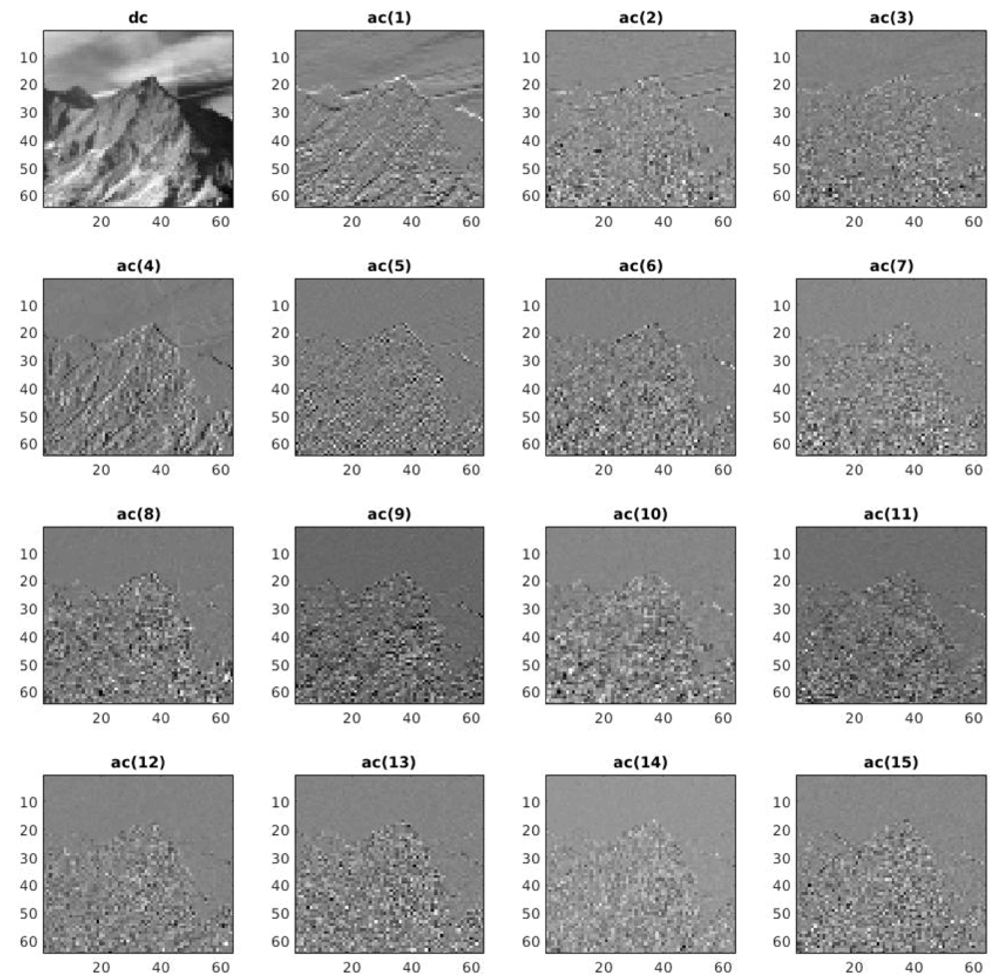

Figure 2. DCT Decomposition

4x4 DCT basis

The Discrete Cosine Transform (DCT) is one of many transforms that takes its input and transforms it into a linear combination of weighted basis functions. DCT basis functions for N = 4 can be seen in figure 3.

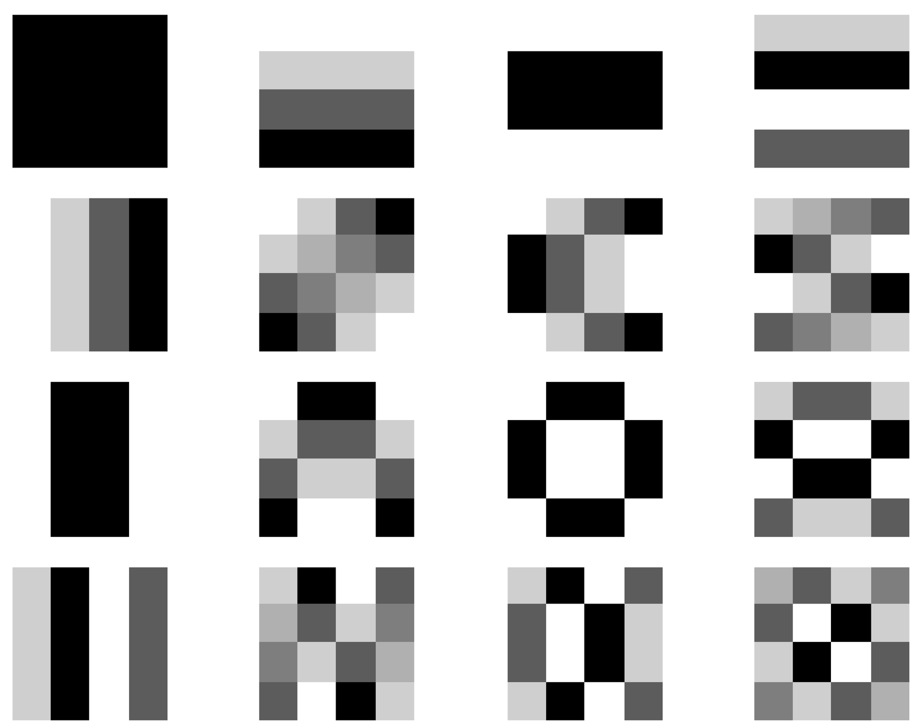

Figure 3. 4x4 DCT basis

Results

	Evaluation Metrics		
	PSNR (dB)	SSIM	AE (degrees)
MFF [4]	17.39	0.61	4.69
ATcycleGAN [5]	20.67	0.68	3.97
SST [6]	14.26	0.57	5.61
SPADE [7]	19.24	0.59	4.59
NIR-GNN [8]	17.50	0.60	5.22
Proposed Method	**22.15**	**0.77**	**3.40**

Table 1. Average PSNR, SSIM, AE for the Validation Dataset

	Evaluation Metrics		
	PSNR (dB)	SSIM	AE (degrees)
Without DCT	13.40	0.37	5.95
With DCT	**22.15**	**0.77**	**3.40**

Table 2. Average PSNR, SSIM, AE for the Validation Dataset

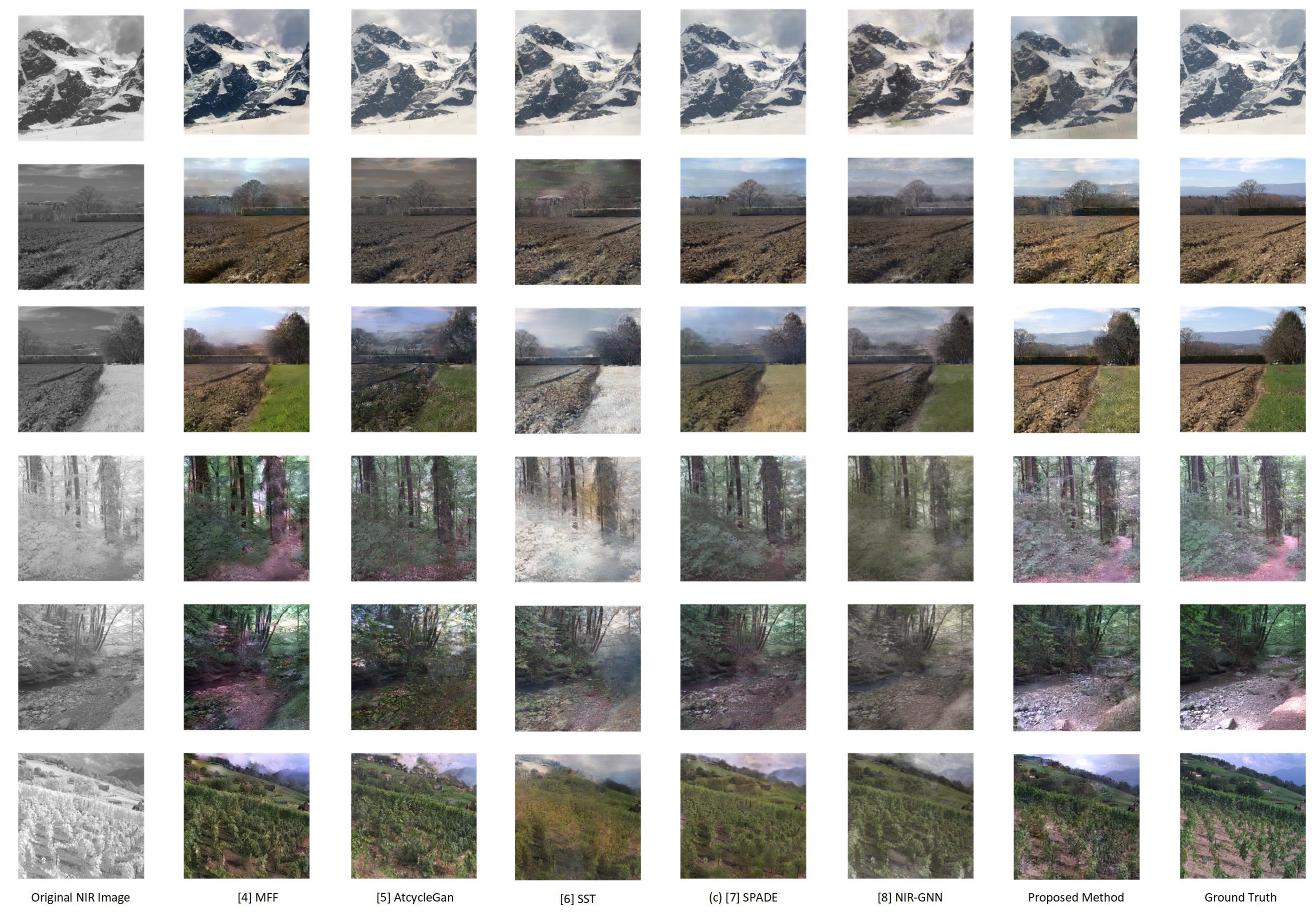

Figure 4. Comparison of the colorized NIR images

References

1. Matthew Brown and Sabine S¨usstrunk. Multi-spectral SIFT for scene category recognition. In CVPR 2011, pages 177–184, 2011. doi: 10.1109/CVPR.2011.5995637.
2. Ziyue Dong, Sei-ichiro Kamata, and Toby P. Breckon. Infrared Image Colorization Using a S-Shape Network. In 2018 25th IEEE International Conference on Image Processing (ICIP), pages 2242–2246, 2018. doi: 10.1109/ICIP.2018.8451230.
3. Patricia L. Su´arez, Angel D. Sappa, and Boris X. Vintimilla. Infrared Image Colorization Based on a Triplet DCGAN Architecture. In 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pages 212–217, 2017. doi: 10.1109/CVPRW.2017.32.
4. Longbin Yan, Xiuheng Wang, Min Zhao, Shumin Liu, and Jie Chen. A Multi-Model Fusion Framework for NIR-to-RGB Translation. In 2020 IEEE International Conference on Visual Communications and Image Processing (VCIP), pages 459– 462, 2020. doi: 10.1109/VCIP49819.2020.9301787.