Multi-Band NIR Colorization Using Structure-Aware Network

Min-Je Park, Ju-Han Lee, Sang-Ho Lee and Jong-Ok Kim†
* School of Electrical Engineering, Korea University, Seoul, Korea
E-mail: {minje234, jfl522, franky, jokim}@korea.ac.kr

***Abstract*—Near infrared (NIR) image can capture the scene in the low light condition without noise unlike RGB. Therefore, it has been widely used in low light vision problems and is often fused with its RGB counterpart for RGB image enhancement. However, there can be some situations that an RGB image can be hardly captured like extremely low-light condition. To cope with this problem, researches to convert NIR image to RGB have been recently conducted based on deep learning networks. These methods show a good performance relatively, but they have some limitations of performance. In this paper, we propose a deep network to colorize multi-band NIR images to RGB using our new dataset. The proposed method attempts to exploit the correlation between individual NIR band and RGB by using multi-band NIR images. It can successfully colorize the multi-band NIR images using two-branch structure and the constraint of the proportional gradient between NIR and RGB.**

I. INTRODUCTION

An NIR image usually captures the light of wavelengths between 700nm and 1000nm in the form of a single channel and it has been widely used for night-vision enhancement as it can capture the details and textures of the scene without noise even in the low-light condition unlike an RGB image which contains a large amount of noise. Also, based on its reflective characteristics, it can be also used for dehazing foggy scenes and agriculture and biology domains. Therefore, there have been conducted so many studies to fuse NIR and RGB images for enhancing details or textures of an RGB image and for noise removal [1, 2, 3, 4].

However, although NIR images contain textures and details that are not available to RGB, there are still some performance limitations to be applied to human visual perception or computer vision algorithms successfully as NIR images are invisible and are different from the RGB luminance channel. Therefore, some researches to colorize NIR images to RGB using deep learning have been conducted recently and show good performances to some extent [5-13].

However, these methods still have a fundamental limitation in that the conversion of NIR to RGB is a highly under-constrained problem. It requires some prior information for further improvement. In this paper, we consider the correlation between NIR and RGB signals that exists on structural domain such as gradient. NIR and RGB images are more closely correlated to each other on high frequency bands rather than low-frequency. Therefore, learning this correlation jointly with color information can boost up the performance of NIR-to-RGB conversion.

Also, estimating 3-channel RGB image from a single channel NIR is an ill-posed problem and the correlation between individual NIR bands and RGB spectrums is hardly considered for NIR-to-RGB conversion. We try to learn this correlation effectively with a deep network by choosing three NIR bands.

In this work, we address a novel NIR to RGB conversion method that uses the NIR-RGB correlation that exists on gradient domain and jointly learns the structural and color information of RGB images. Also, we propose a novel dataset that consists of multi-band NIR and RGB images pairs to learn the correlation on the spectrum domain. In summary, the main contributions of the paper are summarized as follows.

- We propose a novel two-branch network architecture that jointly learns structural and color relationship between NIR and RGB images.
- The proposed network architecture reflects the correlation between NIR and RGB gradient images to the loss as an additional constraint.
- We introduce a new dataset that contains multi-band NIR and RGB image pairs that enable us to learn the correlation between NIR and visible spectrum.

II. RELATED WORK

A. NIR image colorization

Colorizing NIR images to RGB ones can significantly be helpful for night vision applications, but researches were hardly conducted in a non-deep-learning way because of its severe ill-posed property. With the development of deep learning, the colorization of NIR images to RGB has been actively researched recently. There are some previous works that use CNN. [6] used a deep convolutional network and Laplacian pyramid, and [7] used a S-shaped

CNN to acquire clear RGB images around edges. Recently, GAN-based techniques have been proposed. [15] proposed a triple-GAN architecture to estimate R, G, B channels separately, and the works such as [5, 11, 13] proposed CycleGAN based methods to acquire visually rich-quality images. GAN-based methods can be effective to relieve the misalignment problem between NIR-RGB pairs when NIR and RGB image pairs are separately captured and mis-aligned. However, they are hard to be converged and can also produce unnatural artifacts on texture regions. [16] proposed a method to colorize the multi-band NIR image based on the transfer matrix estimation. It utilizes the color checker to estimate the conversion matrix and requires a gray image as well as multi-band NIR images as an input. However, the conversion performance is quite degraded without the gray input, which is a reasonable scenario.

B. Correlation between NIR and RGB

Although RGB and NIR images have quite different characteristics and they are captured on different spectral bands, they share similar structural information. As shown in Fig. 1, NIR and RGB images are different from each other especially in low-frequency components, but high-frequency components like edges and structures are relatively more similar than low-frequency. However, as described in [14], there is inherent discrepancy of structures between NIR and RGB such as gradient magnitude variation, gradient direction divergence, gradient loss, shadow and highlight by flash. In [14], the authors tried to overcome this problem using the pixel-wise scale map obtained by solving the optimization problem. It originally targets to restore the noise-free RGB image using NIR as a guidance. This NIR-RGB correlation is reflected to the cost function of the proposed deep neural network.

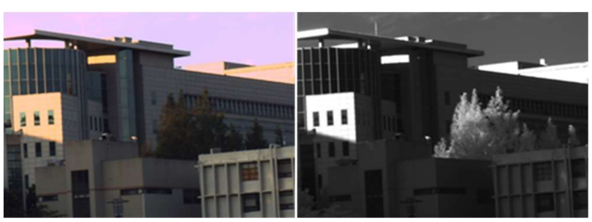


Figure 1. Comparison between NIR and RGB images

III. THE PROPOSED METHOD

A. Dataset

We captured hyperspectral images between 700nm and 1000nm's wavelengths with SPECIM FX-10E hyperspectral camera [17] and synthesized NIR-RGB image pairs corresponding to the NIR-RGB 4 channel camera's spectral response curve. Out dataset consists of 102 training NIR-RGB image pairs and 8 testing NIR-RGB image pairs. For multi-band NIR images, we selected three NIR bands whose wavelengths are 780nm, 850nm, and 940nm, similar to [16]. Some sample images of our dataset is shown in Figure 2.

Figure 2. Some samples of our dataset

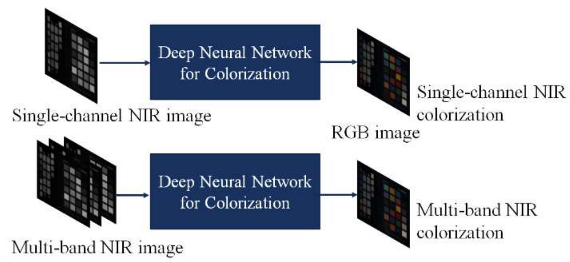


Figure 3. Overall diagram comparison of multi-band and single-channel NIR colorization

B. The Proposed Network Architecture

The proposed overall network architecture is illustrated in Fig. 4. It is based on the U-Net [18], which effectively learns the features of local and global regions using iterative up- and down-sampling. We construct each convolutional block with dense connections to prevent the information from disappearing. The encoder structure is the same as the baseline U-Net, but we construct multi-decoders (two branch) to generate both the gray and RGB images. One branch learns structural information by recovering a gray image from the feature extracted from the NIR input and the other learns color information by restoring an RGB image from the encoded feature. Two decoders are connected with the cross-branch loss which calculates SSIM (structural similarity) between the outputs of two branch networks. It is expressed by

$$L_{cross-branch} = \frac{1}{M}\sum_{i=1}^{M} 0.1 * [1 - SSIM(I^i_{gray}, rgb2gray(\hat{I}^i_{RGB}))] \quad (1)$$

where $\hat{I}_{RGB}$ is the reconstructed RGB image of the ground-truth, I_{RGB} and M is the number of the training images in the dataset.

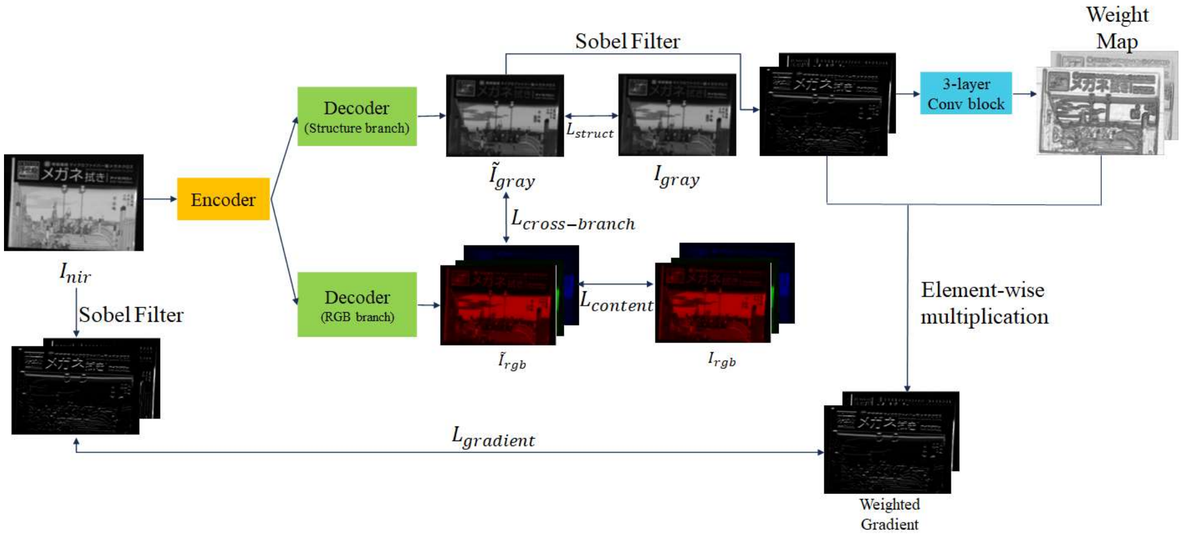


Figure 4. The proposed overall network architecture, which consists of one encoder and two decoder branches for generating gray and RGB images.

The correlation measure between NIR and RGB on the gradient domain is also integrated into the proposed network. After the gray image is recovered, the pixel-wise scale map is obtained using 3-layer convolution blocks from its gradient version, and it is pixel-wisely multiplied to the original gray image's gradient. Finally, the resulting weighted gradient of the gray image is compared to the input NIR gradient. The gradient loss is given by

$$L_{gradient} = \frac{1}{N}\sum_{i=1}^{N}\left|\nabla I_{NIR}^{i} - s * \nabla \tilde{I}_{gray}^{i}\right| \quad (2)$$

where N is the number of pixels for images, s is the pixel-wise scale map, and ∇ indicate the sobel gradient operator.

For the structure branch net, the loss compares its output with the gray version from the generated RGB, and is given by

$$L_{struct} = \frac{1}{M}\sum_{i=1}^{M} 0.1 * \left[1 - SSIM\left(I_{gray}^{i}, \tilde{I}_{gray}^{i}\right)\right] \quad (3)$$

For the RGB branch net, the loss measures the error of the RGB output, and is given by

$$L_{content} = \frac{1}{N}\sum_{i=1}^{N}\left|I_{RGB}^{i} - \hat{I}_{RGB}^{i}\right| \quad (4)$$

$$(for\ first\ 1{,}500\ epochs)$$

$$L_{content} = \frac{1}{M}\sum_{i=1}^{M} 0.1 * \left[1 - SSIM\left(I_{RGB}^{i}, \hat{I}_{RGB}^{i}\right)\right] \quad (5)$$

$$(for\ last\ 2{,}500\ epochs)$$

Finally, the total loss function is as follows.

$$L_{total} = L_{content} + L_{struct} + L_{cross-branc} + L_{gradient} \quad (6)$$

IV. Experimental Result

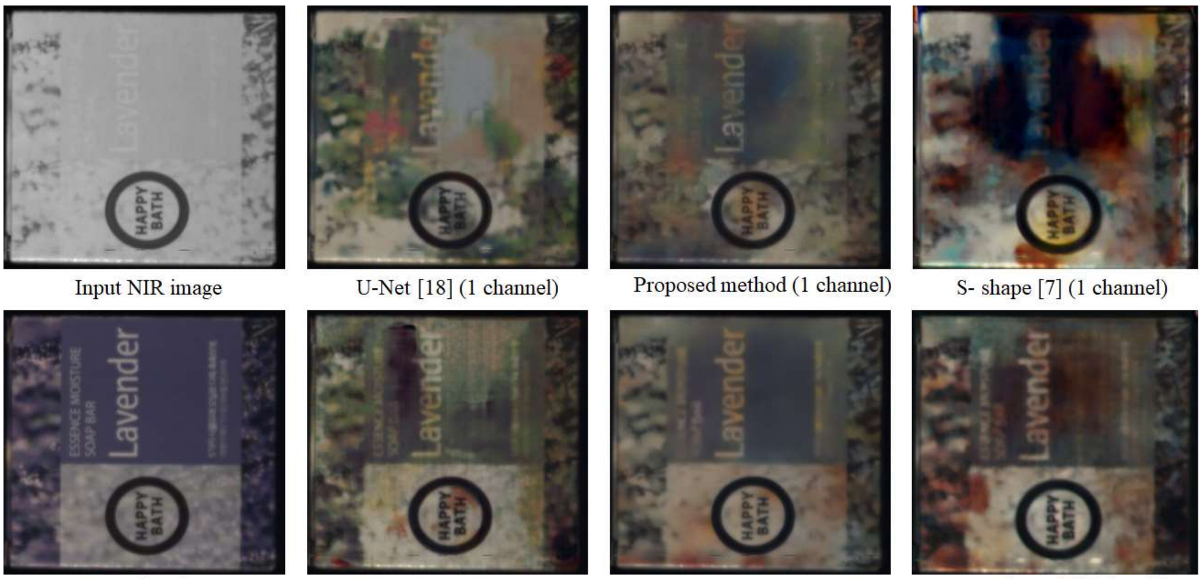
The proposed network was trained with our proprietary datset, which consists of 110 NIR/RGB pairs. Among them 102 NIR/RGB pairs were utilized for training, and 8 pairs were for evaluation. We trained both our model and comparison methods for 4,000 epochs. Fig. 5 shows the visual comparison of the colorized results. The proposed method achieves the best color reconstruction quality among the methods while restraining noises. In Table 1, the performance of the proposed method is quantitatively compared with the existing methods. Feeding three NIR band images makes a high-quality RGB conversion rather than a single NIR channel for the two existing networks as well as the proposed. This means that it is effective to leverage the NIR-RGB correlation on spectral domain. Also, the proposed method accomplishes a higher color restoration quality than the existing methods for three visual quality measures as listed in Table 1. Note that the angular error is a measure for color reproduction quality.

V. Conclusion

In this paper, we proposed a deep network with two-branch decoder and NIR-RGB correlation constraint on the gradient domain to effectively colorize multi-band NIR images. Also, we built a new dataset for NIR-RGB conversion. Utilizing multiple NIR band images enables the network to learn the correlation between NIR and RGB effectively, leading to the further improvement of the colorization performance. Also, the proposed network generates a gray image which provides an additional

	U-Net [18] (1 channel)	S-shape [7] (1 channel)	Proposed method (1 channel)	U-Net (3 bands)	S-shape (3 bands)	Proposed method (3 bands)
PSNR	20.5861	17.7096	21.7937	24.1180	22.2086	24.6006
SSIM	0.6839	0.4549	0.7412	0.7382	0.6660	0.7770
Angular Error	7.8454	16.0870	6.7087	6.9998	9.5480	6.6662

Table 1. Comparison of the colorized results in terms of PSNR, SSIM, and Angular error metrics



Input NIR image | U-Net [18] (1 channel) | Proposed method (1 channel) | S- shape [7] (1 channel)

Ground truth | U-Net [18] (3 bands) | Proposed method (3 bands) | S- shape [7] (3 bands)

Figure 6. Visual comparison of the colorized results

regularization on the consistency between NIR and RGB gradient domains. Experimental results show the superiority of the proposed network quantitatively and qualitatively.

Acknowledgment

This work was supported in part by the National Research Foundation of Korea (NRF) Grant funded by the Korean Government (MSIT) under Grant 2020R1A4A4079705 and in part by the Ministry of Science and ICT (MSIT), South Korea, through the Information Technology Research Center (ITRC) Support Program supervised by the Institute of Information & Communications Technology Planning & Evaluation (IITP) under Grant IITP-2021-2020-0-01749.

References

[1] Zhang, Xiaopeng, Terence Sim, and Xiaoping Miao. "Enhancing photographs with near infra-red images." *2008 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2008.

[2] Awad, Mohamed, Ahmed Elliethy, and Hussein A. Aly. "Adaptive near-infrared and visible fusion for fast image enhancement." *IEEE Transactions on Computational Imaging* 6 (2019): 408-418.

[3] Jung, Cheolkon, Kailong Zhou, and Jiawei Feng. "FusionNet: Multispectral fusion of RGB and NIR images using two stage convolutional neural networks." *IEEE Access* 8 (2020): 23912-23919.

[4] Shibata, Takashi, Masayuki Tanaka, and Masatoshi Okutomi. "Versatile visible and near-infrared image fusion based on high visibility area selection." *Journal of Electronic Imaging* 25.1 (2016): 013016.

[5] Sun, Tian, et al. "Nir to rgb domain translation using asymmetric cycle generative adversarial networks." *IEEE Access* 7 (2019): 112459-112469.

[6] Limmer, Matthias, and Hendrik PA Lensch. "Infrared colorization using deep convolutional neural networks." *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE, 2016.

[7] Dong, Ziyue, Sei-ichiro Kamata, and Toby P. Breckon. "Infrared image colorization using a s-shape network." *2018 25th IEEE International Conference on Image Processing (ICIP)*. IEEE, 2018.

[8] Sun, Tian, and Cheolkon Jung. " Sun, Tian, and Cheolkon Jung. "NIR Image Colorization Using SPADE Generator and Grayscale Approximated Self-Reconstruction." *2020 IEEE*

International Conference on Visual Communications and Image Processing (VCIP). IEEE, 2020.

[9] Valsesia, Diego, Giulia Fracastoro, and Enrico Magli. "NIR image colorization with graph-convolutional neural networks." *2020 IEEE International Conference on Visual Communications and Image Processing (VCIP)*. IEEE, 2020.

[10] Wang, Fengqiao, Lu Liu, and Cheolkon Jung. "Deep Near Infrared Colorization with Semantic Segmentation and Transfer Learning." *2020 IEEE International Conference on Visual Communications and Image Processing (VCIP)*. IEEE, 2020.

[11] Yang, Zaifeng, and Zhenghua Chen. "Learning From Paired and Unpaired Data: Alternately Trained CycleGAN for Near Infrared Image Colorization." *2020 IEEE International Conference on Visual Communications and Image Processing (VCIP)*. IEEE, 2020.

[12] Sekiguchi, Shoichiro, and Masao Yamamoto. "Near-Infrared Image Colorization by Convolutional Neural Network with Perceptual Loss." *2020 IEEE 9th Global Conference on Consumer Electronics (GCCE)*. IEEE, 2020.

[13] Yan, Longbin, et al. "A Multi-Model Fusion Framework for NIR-to-RGB Translation." *2020 IEEE International Conference on Visual Communications and Image Processing (VCIP)*. IEEE, 2020.

[14] Yan, Qiong, et al. "Cross-field joint image restoration via scale map." *Proceedings of the IEEE International Conference on Computer Vision*. 2013.

[15] Suárez, Patricia L., Angel D. Sappa, and Boris X. Vintimilla. "Infrared image colorization based on a triplet dcgan architecture." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2017.

[16] Fu, Qingtao, Cheolkon Jung, and Chen Su. "Color Recovery from Multi-Spectral NIR Images Using Gray Information." *Proceedings of the ACM Multimedia Asia*. 2019. 1-6.

[17] https://www.specim.fi/products/specim-fx10/

[18] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." *International Conference on Medical image computing and computer-assisted intervention*. Springer, Cham, 2015.