Progressive Deep Network with Channel Back-Projection for Hyperspectral Recovery from RGB

Sang-Ho Lee, Min-Je Park and Jong-Ok Kim
School of Electrical Engineering Korea University, Seoul, Korea
E-mail: {franky_, Minje234, jokim}@korea.ac.kr

Abstract—**Hyperspectral images are useful in a variety of fields such as remote sensing, medical diagnosis, and agriculture. But it requires very expensive professional equipment and a lot of time to obtain. In this paper, we propose a deep learning architecture that reconstructs hyperspectral images from RGB images that are easy to acquire in real time. Hyperspectral reconstruction is inherently difficult because much information is lost when hyperspectral bands are integrated into three RGB channels. To effectively overcome the problem of hyperspectral restoration, we design a neural network with the following three basic principles. First, it adopts a method in which channels are gradually increased through several steps to restore hyperspectral images. Second, it is learned on a group basis for efficient restoration. Hyperspectral bands are divided into three groups: R, G, and B. Finally, the concept of channel back projection is newly proposed. In the process of gradually performing hyperspectral reconstruction, the reconstructed image is refined by repeatedly projecting the reconstructed hyperspectral to RGB. In the experimental results, these three principles proved the performance that exceeds the state-of-the-art methods.**

I. INTRODUCTION

Hyperspectral imaging is to obtain multiple images from a scene on distinct spectral bands, and the resulting images commonly include scene characteristic information which cannot be perceived from RGB images. The spectral characteristic is often useful in a variety of fields such as remote sensing, medical diagnosis, and agriculture. However, hyperspectral imaging devices are very expensive, and take much more time for image acquisition rather than consumer RGB cameras with low cost. They have been primarily used as a measurement equipment for research and development. Despite the potential benefits of hyperspectral imaging, its usage is significantly limited to some specific areas. Thus, we need a cheaper and faster way to acquire hyperspectral images (HSI). This is the reason that we need a technology to convert RGB to hyperspectral images.

Hyperspectral signals are inherently three-dimensional, and it is very time-consuming to acquire those signals with a 2D imaging sensor. This considerably restricts either spatial or spectral resolution actually, and finer spectral information can be obtained at the expense of spatial resolution given time duration. Most common visual sensors can only obtain signals in a limited wavelength band that includes standard red, green, and blue in the visible spectrum to match the tricolor perception of the human visual system. Unlike hyperspectral imaging, it is easy to obtain RGB images in a real-time way, and researches for reconstructing HSI from RGB have been conducted in literature. Hyperspectral recovery is inherently difficult because much information has been already lost when hyperspectral bands are integrated into three RGB channels. In other words, it is a severely ill-posed problem due to the mapping from three RGB channels to so many spectral bands.

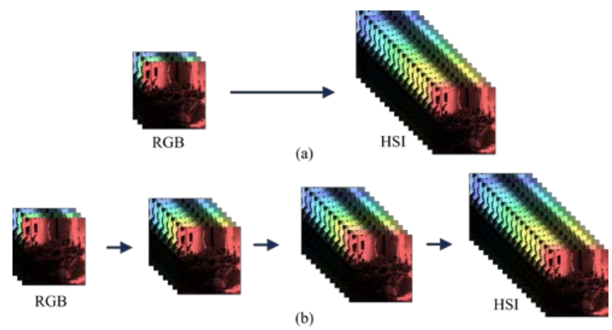

Fig. 1 The concept of hyperspectral recovery for (a) the existing approach and (b) the proposed method (progressive recovery).

In recent years, the deep learning based approach has been studied for hyperspectral recovery, and it accomplishes superior performances when compared with the traditional vision based approach. But, there are still a few works only to adopt a simple network architecture, and they can be further improved to solve the severe ill-posed problem.

In this paper, we propose a novel deep learning architecture for hyperspectral recovery from an RGB image. For effectively overcoming the ill-posed property of hyperspectral recovery, we design a neural network with three underlying principles, which are as follows. First, we adopt the concept of progressive conversion where three RGB bands are gradually increased to multiple hyperspectral ones (e.g., 31 bands in this work) by several phases, not at a time. This idea has been borrowed from image super-resolution where a low-resolution image is magnified gradually, and progressive super-resolution was proved to work effectively. Second, hyperspectral conversion is learned on a local group basis for efficient recovery. Hyperspectral bands are classified into three local groups, and the spectrum of each group dominantly belongs to one among R, G and B spectrums. This is conceptually similar to the factorization of a complex problem to several sub-problems. Finally, we propose a novel concept of channel back-projection. It is adopted to refine the converted signals matched to the RGB input. The gradually converted hyperspectral signals are repeatedly back-projected to the input RGB bands on spectral axis for consistent refinement.

II. RELATED WORK

A learning-based approach has been studied for hyperspectral recovery from RGB images due to insufficient input. It aims to learn the relationship between real hyperspectral images (HSI) and their RGB conversion by camera spectral sensitivity (CSS). In [1-2], an optimal CSS is determined by training a neural network, and the RGB-HSI relationship is learned. For the non-deep-learning approach, the problem of hyperspectral reconstruction was solved by sparse coding [3]. A dictionary is established using hyperspectral images collected in advance and is converted into its RGB version using the receptor spectral absorbance functions. An input RGB image is expressed as a sparse linear combination of RGB dictionary vectors, and its hyperspectral version can be obtained by replacing the RGB dictionary with the hyperspectral. In [4], HSCNN was proposed to reconstruct hyperspectral images with a residual based CNN. In [5], it is further extended by adding a local residual block (HSCNN-R) or a dense block (HSCNN-D). Although HSCNN-R has a simple structure in which a residual block is repeated, it achieves higher performance than state-of-the-art methods. [6] proposed to utilize an adversarial network and used the U-net architecture for the generator. [7] separately performed channel upsampling and reconstruction enhancement. A RGB input is first converted to HSI through an upsampling CNN, and then, reconstruction is enhanced through a residual CNN.

III. THE PROPOSED METHOD

Hyperspectral recovery is significantly challenging because 31 channels should be reconstructed from just 3 channel RGB. It is not efficient to recover hyperspectral images at a time. In this paper, we propose to progressively reconstruct hyperspectral channels in a coarse-to-fine way.

The proposed network basically extends HSCNN-R [5] by implementing two our contributions (i.e., the concept of local group based progressive recovery and channel back-projection). First, a 64-channel feature map is generated from an input RGB using a convolution layer that acts as a spectral interpolation. Hyperspectral channels are divided into three local groups, and they are separately recovered. This separate recovery process between groups probably makes a distinction at the border of groups, and thus, the network is designed so that three groups might be partially overlapped each other as shown in Fig. 3. The smooth reconstruction at the group border is ensured by adding a relevant loss. Also, the progressive recovery from 3 channels to 31 is done by three phases as shown in Fig. 1. We incorporate channel back-projection which is inspired by resolution back-projection used popularly in super-resolution [8].

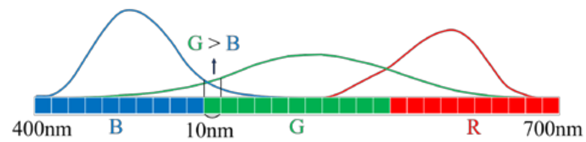

Fig. 2 Grouping of hyperspectral bands based on RGB spectral sensitivity curves.

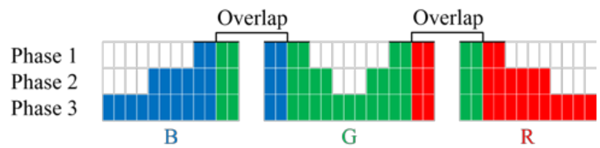

Fig. 3 Local group based progressive hyperspectral recovery.

A. *Progressive Reconstruction via Channel Grouping*

The visible spectral band is in the range between 400 nm and 700 nm as shown in Fig. 2. It is divided into 31 channels by 10 nm intervals. In other words, we should reconstruct 31 hyperspectral images from RGB. The 31 hyperspectral bands are grouped by three, which are named by B, G and R as shown in Fig. 2. We assign one among B, G and R to each hyperspectral band according to the portion of B, G and R spectral sensitivity curves on the spectral band. Namely, each hyperspectral band belongs to a group whose sensitivity curve has the largest portion.

After grouping hyperspectral bands into B, G, and R, hyperspectral images are progressively reconstructed by three phases as shown in Fig. 3. In the first phase, 8 channels are reconstructed from 3 RGB, and then, 20 and 31 channels are reconstructed from the output of the previous phase in a sequential way. The 8 channels in the first phase is chosen as ones between group borders because they can be reconstructed from at least two channels among RGB. It is thought that it is better to restore a hyperspectral channel whose information is more available to RGB. Four channels are produced from B and G, while the other four channels are from G and R. In the next phases, we gradually reconstruct more hyperspectral channels in both left and right directions from 8 ones already obtained in the first phase as shown in Fig. 3. Note that in order to keep the reconstruction consistent among three local groups, 4 channels are overlapped on both RG and BG groups. The overall proposed architecture is illustrated in Fig. 4.

B. *Channel Back-Projection*

In channel back-projection of Fig. 5, the concept of back-projection in super-resolution is applied to the spectral domain by converting input hyperspectral images to RGB through spectral downsampling. And upsampling the residual between the projected version and the ground truth RGB is added to the module's input. Spectral downsampling is replaced by 1x1 convolution, and when hyperspectral images are projected to RGB domain, B, G, and R groups are directly

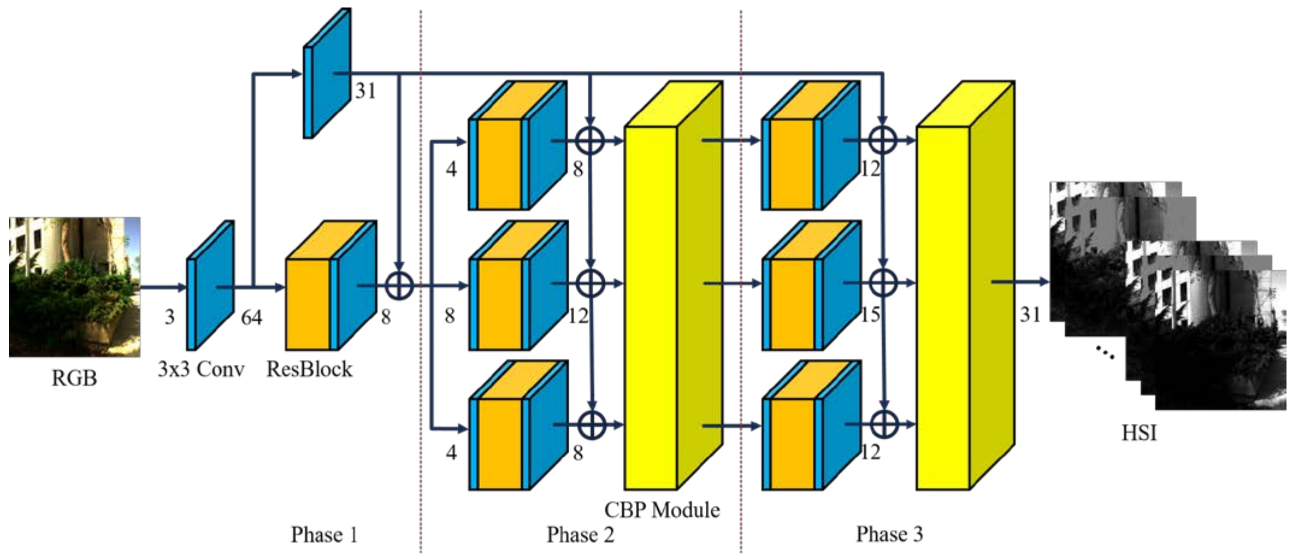

Fig. 4 The proposed progressive deep learning architecture.

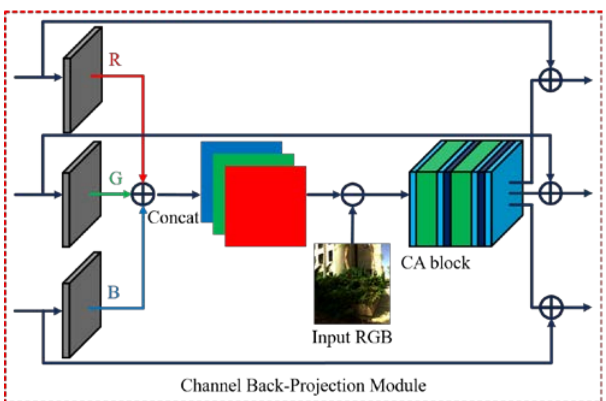

Fig. 5 The architecture of channel back-projection.

converted to B, G, and R channels, respectively. After being converted to an RGB image, its residual associated with the ground truth is obtained and upsampled through convolution and channel attention [9] as shown in Fig. 5.

C. Loss Function

Mean relative absolute error (MRAE) was adopted as a loss function for training. It is generally considered to be suitable for measuring errors in hyperspectral images composed of multiple channels with different dynamic ranges.

$$MRAE = \frac{1}{p}\sum_{j=1}^{p}\left(\frac{\left|I_{HS}^{j} - I_{GT}^{j}\right|}{I_{GT}^{j}}\right)$$

where I_{HS}^{j} and I_{GT}^{j} are the j th pixel value of the reconstructed hyperspectral image and ground truth, respectively, and p is the total number of pixels. The proposed network progressively reconstructs hyperspectral channels through the three phases, and the loss for each phase is added to the total loss function as follows.

$$L = L_1 + L_2 + L_3$$

where L_n is the MRAE value of the image reconstructed in the n th phase. Note that L_2 is calculated before channel back-projection, and L_3 is calculated after that.

IV. Experimental Results

The proposed network was trained with ICVL Dataset [3]. A 1392x1300 hyperspectral image is converted to its RGB version using the CIE 1960 color matching function (CMF), and 50x50 patches are extracted with 80 stride to construct a pair of RGB and HSI patches. The proposed network is trained for 36 hours using two NVIDIA 2080 Ti GPUs.

Fig. 6 subjectively compares the reconstruction error for each method. The conventional methods have a tendency for errors to appear outstandingly in the object at the center of the image, but the proposed method can restore the object better in particular. In Fig. 7, we plot a hyperspectral curve at a small local region. Namely, we plot the average value of a 11x11 patch for every hyperspectral channel on a spectral axis. We can see that the curve shape of the proposed method is very close to the ground truth.

In Table 1, the performance of the proposed method is compared with state-of-the-art methods in terms of MRAE and root mean square error (RMSE). Even if channel back-projection is not applied to the proposed network, both MRAE and RMSE show a superior performance gain over the existing methods including HSCNN-R, so it can be confirmed that channel grouping and progressive reconstruction are effective for hyperspectral image reconstruction. In addition, channel back-projection can contribute to further improve the

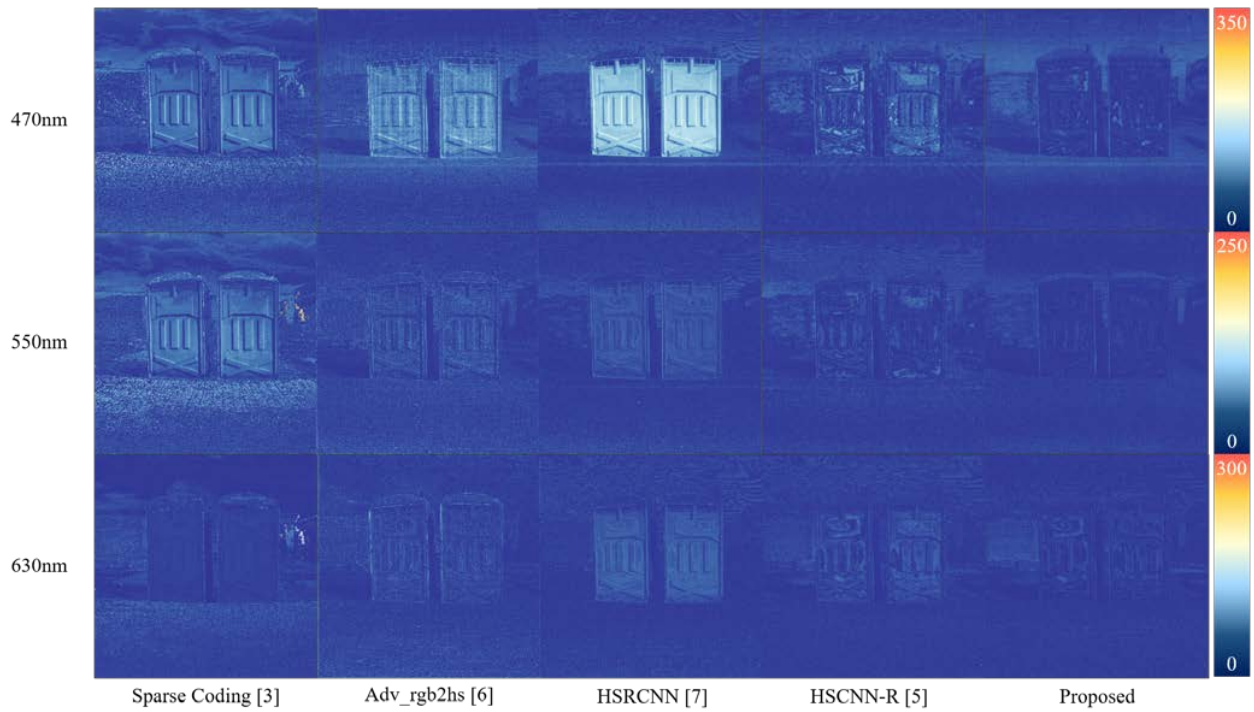

Fig. 6 Comparison of the reconstructed hyperspectral images.

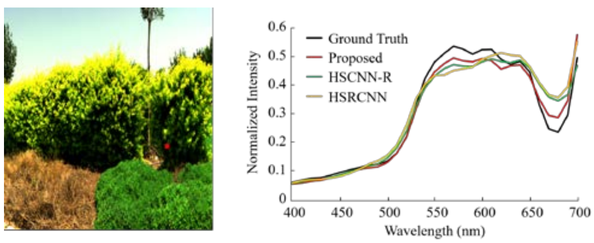

Fig. 7 Comparison of hyperspectral curves.

Method	MRAE	RMSE
Sparse Coding [3]	0.06518	3.9153
Adv_rgb2hs [6]	0.02540	2.3419
HSRCNN [7]	0.02105	1.9350
HSCNN-R [5]	0.01612	1.6539
Proposed (No CBP)	0.01560	1.5534
Proposed	0.01497	1.3958

Table. 1 Comparison of quantitative results.

performance of MRAE by about 4% and RMSE by about 11%.

V. Conclusion

In this paper, we propose a network that can effectively reconstruct hyperspectral images from RGB using channel grouping and progressive reconstruction. In addition, channel back-projection is proposed to further improve the performance. Unlike existing methods that perform hyperspectral recovery at a time, the proposed method gradually reconstructs 8, 20, and 31 channels by three phases. This is motivated by the severe ill-posed property of hyperspectral recovery. Also, hyperspectral channels are partitioned into three groups according to their correlation with R, G, and B, and they are separately reconstructed progressively. From experimental results, it is confirmed that the proposed method can achieve better reconstruction performance quantitatively and qualitatively compared to the existing methods.

Acknowledgment

This work is supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2020R1A4A4079705)

References

[1] Nie, Shijie, et al. "Deeply learned filter response functions for hyperspectral reconstruction." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018.J. Clerk Maxwell, *A Treatise on Electricity and Magnetism, 3rd* ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.

[2] Fu, Ying, et al. "Joint camera spectral sensitivity selection and hyperspectral image recovery." *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018.

[3] Arad, Boaz, and Ohad Ben-Shahar. "Sparse recovery of hyperspectral signal from natural RGB images." *European Conference on Computer Vision*. Springer, Cham, 2016.
[4] Xiong, Zhiwei, et al. "Hscnn: Cnn-based hyperspectral image recovery from spectrally undersampled projections." *Proceedings of the IEEE International Conference on Computer Vision Workshops*. 2017.
[5] Shi, Zhan, et al. "Hscnn+: Advanced cnn-based hyperspectral recovery from rgb images." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2018.
[6] Arad, Boaz, et al. "Ntire 2020 challenge on spectral reconstruction from an rgb image." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. 2020.
[7] Han, Xian-Hua, Boxin Shi, and Yinqiang Zheng. "Residual hsrcnn: Residual hyper-spectral reconstruction cnn from an rgb image." *2018 24th International Conference on Pattern Recognition (ICPR)*. IEEE, 2018.
[8] Haris, Muhammad, Gregory Shakhnarovich, and Norimichi Ukita. "Deep back-projection networks for super-resolution." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.
[9] Hu, Jie, Li Shen, and Gang Sun. "Squeeze-and-excitation networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.