

Hyperspectral Imaging with Random Printed Mask

Yuanyuan Zhao¹ Hui Guo¹ Zhan Ma¹ Xun Cao¹ Tao Yue^{1,2} Xuemei Hu¹

¹Nanjing University, Nanjing, China

²NJU institute of sensing and imaging engineering, Nanjing, China

{yuan_square, guo_hui}@smail.nju.edu.cn {mazhan, caoxun, yuetao, xuemeihu}@nju.edu.cn

Abstract

Hyperspectral images can provide rich clues for various computer vision tasks. However, the requirements of professional and expensive hardware for capturing hyperspectral images impede its wide applications. In this paper, based on a simple but not widely noticed phenomenon that the color printer can print color masks with a large number of independent spectral transmission responses, we propose a simple and low-budget scheme to capture the hyperspectral images with a random mask printed by the consumer-level color printer. Specifically, we notice that the printed dots with different colors are stacked together, forming multiplicative, instead of additive, spectral transmission responses. Therefore, new spectral transmission response uncorrelated with that of the original printer dyes are generated. With the random printed color mask, hyperspectral images could be captured in a snapshot way. A convolutional neural network (CNN) based method is developed to reconstruct the hyperspectral images from the captured image. The effectiveness and accuracy of the proposed system are verified on both synthetic and real captured images.

1. Introduction

Spectra can provide additional information of scenes beyond the ability of human eyes and commercial RGB cameras, having great potential to facilitate computer vision tasks [3, 9, 29]. However, the high complexity and cost of spectral imaging systems greatly raise the difficulty of acquiring spectral images, and thus limit the wide application of spectral imaging.

Traditionally, to capture hyperspectral images, the scanning based methods (either the spatial scanning [18] or the spectral scanning [15]) are required. These scanning based systems can capture images with several to hundreds spectral channels, while the ability to handle dynamic scenes are sacrificed. To take the spectral image in a single snapshot, the snapshot spectral imaging methods are proposed in the past few years [6, 10, 16, 25, 31, 32]. However, most of

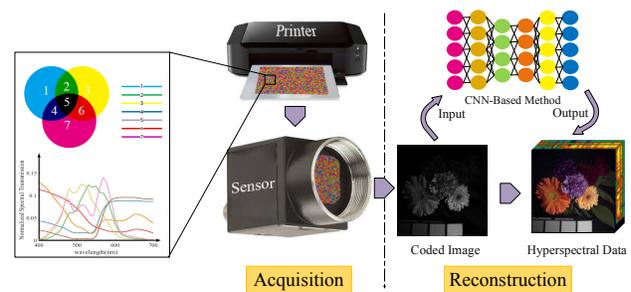


Figure 1. Overview of the proposed hyperspectral imaging system: with a consumer-level printer, a random color printed mask can be attached to the sensor of the camera to sample the hyperspectral images. Through randomly overlaying ink droplets, the spectral transmission responses of different points are rendered highly uncorrelated and facilitate the full spectrum recovery with details. A convolutional neural network is proposed to recover the hyperspectral images from the randomly coded images.

these systems suffer from the system complexity and calibration difficulty. Customized spectral filter based methods are proposed to realize compact hyperspectral imaging [5, 20, 23, 28], while these methods suffer from the fabrication cost of the filters. RGB cameras are also studied to be turned into spectrometers [2, 4, 27, 24, 33]. However, the uncorrelated spectral modulations (i.e. RGB Bayer filter) are limited and spectral details are hard to be retrieved.

In this paper, we propose a simple and low-cost spectral imaging scheme with a color mask, which can be printed by consumer-level printers. The idea of the proposed system is based on a simple but not widely noticed phenomenon, i.e. the spectral transmission responses of overlapping printed color dots are the multiplication of the spectral transmission response of each overlapped color dots. The multiplied spectral transmission response is linearly uncorrelated with that of the printer inks. Therefore, we can generate the color mask with a large number of uncorrelated spectral transmission responses. With the color mask, hyperspectral images can be encoded in high quality. We develop a CNN-based method to reconstruct the hyperspectral images from the captured image and built a prototype hyperspec-

tral imaging system. The proposed hyperspectral imaging method is validated on both the simulated and real captured data. The main contributions of the paper are:

- We *propose* to generate the color mask with a large number of uncorrelated spectral transmission responses by using consumer-level color printers and *propose* a simple hyperspectral imaging scheme based on the randomly printed color mask.
- We *develop* a CNN-based reconstruction network to recover the hyperspectral images from observations.
- We *build* a prototype imaging system to verify this approach and *demonstrate* the feasibility and effectiveness on both synthetic and experimental data.

2. Related Work

Snapshot hyperspectral imaging technologies have been evolving rapidly in the last few decades. To capture the 3D hyperspectral images with a 2D imaging sensor in a snapshot way, either spatial or spectral coding is introduced. In terms of the specific coding techniques, existing methods can be divided into three main categories: dispersion based spectral imaging methods [6, 16, 25, 31, 32], the scattering based spectral imaging methods [12, 30], and the spectral filter based methods [2, 4, 5, 13, 20, 23, 24, 27, 28, 33].

Dispersion based Spectral Imaging Methods. With dispersive elements, e.g. prisms or gratings, the spectrum of each point is spread spatially. Through introducing a spatial coding, the spectral information is coded indirectly and captured [6, 10, 16, 25, 31, 32]. While these spectral imaging methods could realize snapshot and high quality hyperspectral imaging, sophisticated calibration are always required and the system is bulky. Compared with these methods, our hyperspectral imaging method only requires to print a color mask and attach the color mask in front of the camera sensor, which is easy to implement, of low cost and promising for wider application in practice.

Scattering based Spectral Imaging Methods. Besides spectral coding through introducing dispersers, scattering medium is also introduced to encode the spectral information with different speckle patterns [12, 30]. While these methods are promising for compact hyperspectral imaging, the spectral resolving ability is limited due to the speckle correlation among different wavelengths. In our method, we propose to use overlaying of ink drops to generate spectral modulation which is highly uncorrelated and could enable high quality encoding of hyperspectral imaging.

Spectral Filter based Spectral Imaging Methods. Other than those indirect spectral coding methods, spectral imaging could also be realized through direct spectral coding: designing and attaching the spectral filter in front of the camera sensor [5, 20, 23, 28]. Deep learning or dictionary learning based methods, which exploit the sophisticated spatial-spectral prior, are proposed for high quality

hyperspectral recovery [8, 13, 14]. These hyperspectral imaging methods may require high precision manufacturing of spectral filters, while our spectral imaging method only requires to print a color mask with consumer-level printers. Recently, turning commercial RGB cameras into hyperspectral imaging also has emerged as a hot research topic [2, 4, 24, 27, 33], which is promising for a low-cost spectrometer. However, the uncorrelated spectral modulations are limited and may not be enough to recover spectral details. Our method could provide various uncorrelated spectral modulations, making it possible to realize higher quality encoding and recovery of hyperspectral information.

In all, we propose a novel low-cost and easy-to-implement hyperspectral imaging technique. The calibration of our method is easy and most importantly, through exploiting the new generated spectral transmission responses of randomly overlaying ink droplets, detailed hyperspectral information could be encoded. We propose a CNN-based network model to extract the hyperspectral images and demonstrate the effectiveness of our method through both synthetic and physical experiments.

3. Transmission Model of Random Printed Mask

The basis of our work is that the spectral transmission responses of inks of different colors are various and the spectral information can be encoded with a randomly printed color mask. We investigate the printed color dot characteristics, formulate the mask printing model and analyze the color transmission responses with different printing parameters. Based on the analysis, we could choose the optimal physical parameters to print our spectral coding mask.

3.1. Characteristics of Random Printed Mask

Random Distribution Characteristics. We first observe the distribution of ink droplets at the micron level using a microscope with 20X magnification. The CMYK color mode, commonly used in color printing, contains four standard colors, i.e. cyan (C), magenta (M), yellow (Y), and black (K). Since black ink absorbs light of all wavelengths, we generated a background picture with CMYK mode in which case C, M and Y channel is set to 10 respectively and K channel is set to 0. We print a uniform mask with the CMYK value on a transparent film and observed the printed picture under the microscope, as shown in Fig. 2(a). It can be observed that the positions of droplets at the micron level are not regularly arranged, but relatively random and uncontrolled. Each individual drop of ink is approximately round. New colors such as purple, green and orange are produced due to the overlapping of the CMY ink droplets.

Multiplicative Stacking Characteristics. Observing the random distribution of ink drops, we further explore how

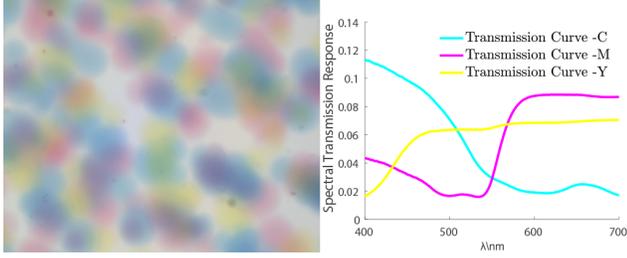


Figure 2. (a) Image of ink droplets distribution at micron level. (b) Spectral transmission response of CMY color printed film.

would the spectral transmission response change if different ink droplets are stacked. We assume that the transmission response would be the multiplication of the transmission responses of overlapped ink droplets as,

$$\mathbf{c}_p = \prod_i \mathbf{c}_i, \quad (1)$$

where \mathbf{c}_p means the transmission response of overlapped layers and \mathbf{c}_i is the transmission response of the i th layer.

We conduct an experiment to verify the assumed stacking model. A single point spectrometer (ASD TerraSpec 4 Standard-Res Mineral Analyzer) is used to measure the transmission responses. After removing the influences of light source and camera spectral sensitivity response, the transmission responses of CMY inks are shown in Fig. 2(b). We calculate the transmission response of the overlapped ink droplets of magenta-magenta and cyan-magenta and measure that with the single spectrometer. As shown in Fig. 3, the measured transmission responses match well with the estimated one by the multiplication model (Eq. 1).

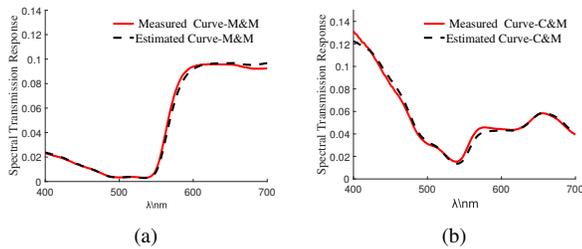


Figure 3. (a) Transmission response with the same ink stacking. (b) Transmission response with different ink stacking. The estimated spectral transmission responses are calculated by Eq. (1).

3.2. Modeling the Transmission of Random Printed Mask

With the multiplicative stacking model, we proceed to model the multilayer monochromatic mask to simulate the transmission response matrix of printed mask. The main factors to be considered in modeling are ink density p , number of layers l , diameter of single droplet d and number of

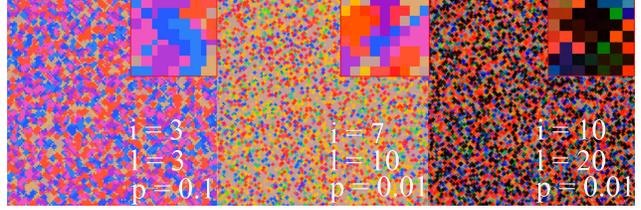


Figure 4. The RGB images of the color masks with different printing parameters.

ink colors. For one monochromatic layer, the printed mask can be formulated as

$$\mathbf{M}_i = \mathbf{I}(p) * \mathbf{K}(d, \mathbf{c}_i), \quad (2)$$

where $\mathbf{I}(p)$ denotes the random 0-1 printing pattern of dimension $H \times W$. \mathbf{M}_i means the mask transmission response matrix, \mathbf{K} denotes the circle kernel whose diameter is d (the shape of printed drop is approximately circular in practice). \mathbf{c}_i is the spectral transmission response of the i th layer. $*$ means the spatial convolutional operation.

For a mask of mixed colors, it is equivalent to print multiple layers of various colors. According to the multiplicative stacking characteristics, multilayer monochromatic mask \mathbf{M} can be modeled as

$$\mathbf{M} = \prod_{i=1}^L \mathbf{M}_i, \quad (3)$$

where \prod means element-wise product and L denotes the layer number.

According to the model, we simulate the mask printed with different parameters and discuss the effects of these parameters on spectral retrieval fidelity to select the best print settings in the next section.

3.3. Characteristic Analysis

In this chapter, we analyze the spectral reconstruction performance of different physical parameters for printing the color mask and choose the optimal one for our imaging technique. Fig. 4 shows pseudo-color images of the simulated masks under different parameters according to the calibrated spectral response of a commercial camera (point grey Grasshopper3).

We build the imaging model first and gives a very simple reconstruction method, so that the quality of random printed mask can be assessed by using the accuracy of the reconstruction results. The imaging model of encoding the spectral information from the scene with the printed mask can be expressed as

$$\mathbf{y} = \mathbf{C}\mathbf{s}, \quad (4)$$

where $\mathbf{y} = [y_1, y_2, \dots, y_n]^T$ denotes the signal encoded by printed mask, $\mathbf{C} = [\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n]^T$ refers to transmission

matrix of printed mask. \mathbf{s} means the spectra of scene which could be estimated through the least square problem with Tikhonov regularization [17]

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s}} (\|\mathbf{y} - \mathbf{C}\mathbf{s}\|_2^2 + \alpha \|\mathbf{s}\|_2^2), \quad (5)$$

where α is the regularization coefficient and setted as 0.01.

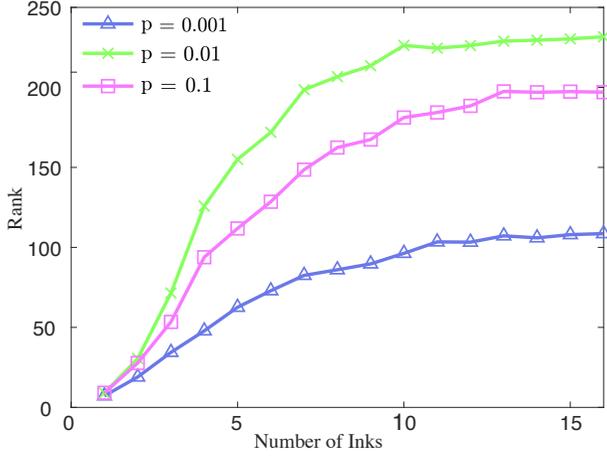


Figure 5. Statistical analysis of the relation between the maximum rank of transmission matrix and printing parameters p, i .

As for printing the color mask, there are four physical parameters: the diameter of each ink droplet d , the print density p , the number of inks i and the overlapping layer number l . In terms of the diameter, it is physically determined by the printer nozzle and fixed to 4 pixels in simulation, which is consistent with the size in real experiments. It is well known that the higher the rank of \mathbf{C} in Eq.(4), the more linearly uncorrelated spectral measurements we can get, facilitating higher fidelity spectral recovery. Thus, we are committed to seeking parameters to make the rank of printed mask as high as possible. Considering the implementation practicality, we limit the number of layers to 150 layers. To choose the best parameter values, we first randomly simulate masks 100 times under different parameter settings, and statistically analyze the relation between the maximum rank of transmission matrix and parameters p and i , as shown in Fig. 5. It is obvious that the larger the ink number, the higher rank can be achieved when the density p is fixed. Nevertheless, the rank of transmission matrix decreases whether the print is too thick or too sparse. Thus, $p = 0.01$ is chosen as the printing density in our experiments. It is worth noting that when i reaches 10, the rank growth slowly for $p = 0.01$. Hence $i = 10$ is chosen as the final printing ink number in our experiments.

In order to more intuitively establish the relationship between rank of mask under different parameters and reconstruction performance, we conduct three sets of experiments, as shown in Fig. 6. We firstly synthesize the color

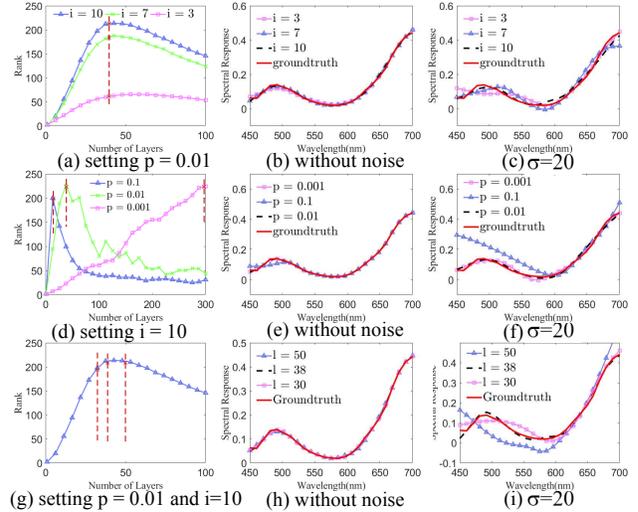


Figure 6. Setting print density p to 0.01: (a) rank analysis for different number of inks used when setting print density as 0.01, (b)-(c) reconstructed spectrums with parameters marked in (a), white Gaussian noise ($\sigma = 20$) is added in (c). Setting the number of inks i to 10: (d) rank analysis for different layers and print density. (e)-(f) reconstructed spectrums with parameters marked in (d), white Gaussian noise ($\sigma = 20$) is added in (f). Setting i to 10 and p to 0.01: (g) rank analysis for different layers when p is 0.01 and i is 10, (h)-(i) reconstructed spectrums at parameters marked in (g), white Gaussian noise ($\sigma = 20$) is added in (i).

mask with different i and l when $p = 0.01$, as shown in Fig. 6(a). It can be observed that with fixed i , the rank increases first and then decreases as l increases. Printing too many layers on the same sheet would decrease the light throughput and deteriorate the spectral coding ability of the color mask. We visualize the corresponding reconstruction results under different noise conditions with different parameters in Fig. 6(a) marked with red cross. As we can see from Figs. 6(b)-(c), with $p = 0.01$ and $l = 38$, the more inks, the rank of the transmission response is higher and the spectrum is reconstructed with higher fidelity.

We then synthesize the color mask with different printed density and layer number, fixing the other parameters. The changing trend of rank under different print density and layers is shown in Fig. 6(d). We further compare the corresponding reconstruction results under different noise conditions with different parameters in Fig. 6(d) marked with red cross to choose the best parameters to print, as shown in Figs. 6(e)-(f). The standard deviation σ of the added noise is 20 in Fig. 6(f). It can be observed that printing 38 layers when $p = 0.01$ or printing 300 layers when $p = 0.001$ can reconstruct the spectrum with high-fidelity and high tolerance to noise. However, print too many layers is impractical for implementation, so we choose the $p = 0.01$. To further determine the optimal number of print layers, we analyze different layers to print when setting p and i fixed, as

shown in Fig. 6(g) marked in red cross. The number of layers respectively is 30, 38 and 50. The reconstruction results are shown in Figs. 6(h)-(i). It can be seen that printing too many layers may result in a decrease in randomness, and on the contrary, printing too few causes insufficient information collection capability. In summary, printing 35-45 layers corresponds to the reasonably good performance.

Based on the above analysis, we choose the print density to be 0.01, diameter 4, color number 10 and the layers number 35-41 as our final printing parameters of the color mask (see Sec. 5.2 for implementation details). In the following, we will develop a CNN-based method for reconstructing hyperspectral images with our imaging technique.

4. Image Formation Model and Inverse Method

By placing the color mask in front of the sensor, the spectral images are encoded, integrated along wavelength and captured by a 2D sensor. The image formation model is

$$\mathbf{I}(x, y) = \sum_{\lambda} \Phi(x, y, \lambda) \mathbf{S}(x, y, \lambda), \quad (6)$$

where Φ denotes the spectral transmission matrix of the color mask, \mathbf{S} denotes the 3D hyperspectral images, and \mathbf{I} is the captured image. Reconstructing the 3D hyperspectral images \mathbf{S} from the captured 2D coded image is highly underdetermined, where sophisticated sparsity prior of hyperspectral images is required. Dictionary based methods have been pioneeringly applied to introduce the statistical prior of hyperspectral images into the inverse problem and successfully recover high quality hyperspectral images [1, 14, 25]. Beyond that, CNN which is expert in learning statistical priors from data, has also been applied and achieve remarkable performance in hyperspectral image reconstruction [8, 13, 24].

In this paper, we propose a CNN-based method for spectral reconstruction. In order to increase the receptive field (RF) [21] of the model and enable it to integrate the coded information of mask of different size level, a multiscale network model is employed. As shown in Fig. 7, the proposed method learns an end-to-end mapping from a large number of coded images and ground truth hyperspectral image pairs. The input of network is a two-dimensional image encoded by a printing mask, and the output is the reconstructed spectral image. The CNN-based model we used is represented as \mathcal{F} . The input and output pairs fed to \mathcal{F} is represented as $\{\mathbf{I}_i, \mathbf{S}_i\}_1^N$. \mathbf{I} is obtained by Eq.(6) in synthetic experiment with ground truth hyperspectral images \mathbf{S} , where Φ means the transmission response matrix of the color mask, $\Phi \in R^{H \times W \times \Lambda}$ and $\mathbf{S} \in R^{H \times W \times \Lambda}$, Λ denotes the spectral resolution. The output of the network is

Network Structure. Our model (Fig.7) is based on the multiscale structure, which is downsampled four times with

maximum poolings. The size of feature maps is shrunk to half of previous layer after each downsampling. Bilinear upsampling instead of deconvolution operation is used to prevent the checkboard effect. Bottleneck in Resnet [19] is added following upsampling to smooth the feature map. Skipping connections are introduced in our network structure to improve the disappearance of gradients during back-propagation and merge shallow and deep information. The multiscale scheme is proposed basically on the purpose to extract the correlation information from pixels in different scales of receptive fields for better spectral reconstruction.

$$\hat{\mathbf{S}} = \mathcal{F}(\mathbf{I}). \quad (7)$$

Loss Function. Parameters of each layer in \mathcal{F} are defined as $\theta = \{\mathbf{W}_l, \mathbf{b}_l\}_1^{d+1}$, and d is the number of hidden layers. These parameters are trained with the loss function in Eq. 8. The first term is the mean squared error (MSE) of the groundtruth hyperspectral image \mathbf{S} and the predicted hyperspectral image $\hat{\mathbf{S}}$ from the network. In order to recover the detailed characteristic information in spectral dimension, we add the spectrum constancy loss as the second term in Eq. 8, which constraints the first-order derivative of \mathbf{S} and $\hat{\mathbf{S}}$ in the wavelength dimension to be similar. In addition, decay term of weights is included to avoid overfitting.

$$L = \underbrace{\beta_1 \|\mathcal{F}(\mathbf{I}) - \mathbf{S}\|_2^2}_{\text{data term}} + \underbrace{\beta_2 \|\nabla_{\lambda} \hat{\mathbf{S}} - \nabla_{\lambda} \mathbf{S}\|_2^2}_{\text{spectrum constancy loss}} + \underbrace{\frac{\tau_w}{2} \sum_1^{d+1} \|\mathbf{W}_l\|_2^2}_{\text{decay term}} \quad (8)$$

Implementation Details. The databases used for training are publicly available, including Harvard [7], Columbia [34], KAIST [8], and Manchester [26, 11] spectral image database. Data augmentation method is used to preprocess datasets by means of cutting, scaling, rotation, etc. 40000 enhanced data pairs of size $(256 \times 256 \times 31, 256 \times 256)$ are sampled. The augmented dataset is divided into training and validation set by 4 : 1. Pytorch framework is employed to train our model. ADAM [22] is adopted for optimization. The leaning rate is set to 10^{-4} initially, and scaled to 1/3 of previous one starting from the 6th epoch. β_2 is gradually increased from the 6th epoch, scaled by 1.1 each epoch. The weight τ_w for the decay term is set to 10^{-8} . The network training process lasted about 24 hours. The hardware platform we used is configured with an Intel(R) CPU E5-2609 with 64GB memory and NVIDIA Tesla P100 with 16GB of memory. Zero mean simulated Gaussian noise with a standard deviation of 5 is added during training and testing.

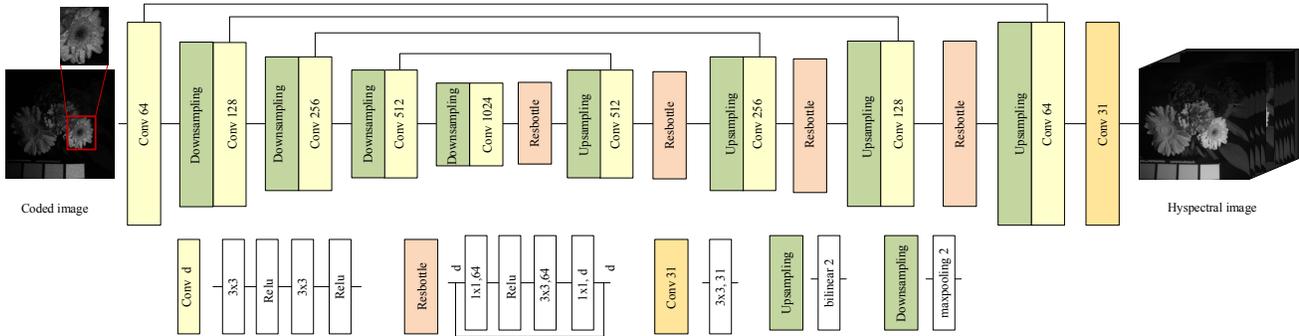


Figure 7. Neural network structure. The input is the spectrally encoded images, and the output is the recovered hyperspectral images of 31 channels. The multiscale network is composed of convolution layers, bilinear upsampling, maxpooling and Resbottle modules, exploiting the correlation among pixels over receptive fields of different scales.

Table 1. Quantitative comparisons with state-of-the-art methods

Methods	PSNR(dB)	SSIM	MSE
Sparse Coding [25]	25.97	0.86	0.0008
Deep CASSI [8]	34.47	0.91	0.01
Our CNN+RGB	21.5	0.57	0.03
Our Method	34.74	0.93	0.0004

5. Experiment

In this section, we demonstrate the effectiveness of our method through both simulation and physical experiments.

5.1. Experiments on Synthetic Spectral Data

We compare our method against the other three state-of-the-art methods: sparse coding based method (Sparse Coding) [25], CNN-based reconstruction with coded aperture snapshot spectral imaging system (Deep CASSI) [8] (network model of [8] provided by the authors), and our CNN network with RGB image as input (Our CNN+RGB). For fair comparison, parameters of the three methods are traversed and set to the ones with the best performance.

Quantitative Comparisons of Reconstruction. As shown in Tab. 5.1, we verify our method on four validation datasets (Harvard [7], Columbia [34], KAIST [8], Manchester [11, 26] datasets). The average peak signal to noise ratio (PSNR), structure similarity index metrics (SSIM), mean squared error (MSE). As shown, the performance of our method is comparable to the other state-of-the-art methods in PSNR, SSIM and MSE. As a whole, we demonstrate the effectiveness of our method through extensive comparisons.

Qualitative Comparisons of Reconstruction. To visualize the comparison, we show the reconstruction result of our method and the other three methods on four images, each from a different database. We first calculate the corresponding RGB images with the reconstructed hyperspectral images and RGB spectral response of a commercial RGB camera (point grey Grasshopper3), as shown in 1-4th

rows in Fig. 8. Through comparing with the other methods, the color images calculated from the recovered hyperspectral images with our method are the most similar with the ground truth. This could be demonstrated through comparing the error maps of RGB images of different methods, as shown in the 5-8th rows in Fig. 8. Furthermore, we randomly pick four points from these four scene respectively and compare the recovered spectrum directly in the first column, 5-8th rows. As shown, our methods could recover most spectral details, demonstrating the effectiveness of our method.

5.2. Experiments on Real Captured Data

To prove the validity of our method with real captured data, we build a prototype system with the printed mask. To make the mask more qualified and random, we investigate the transmission curves of 13 inks (C13T761280, C13T761380, C13T761480, C13T761580, C13T761680, LAMY T52-Cyan, T52-Magenta, T52-Yellow, and five Pilot INK-30 series) in the market, and compare the correlation coefficients among them. Since the RGB Bayer pattern of the camera already introduces 3 color filter with different transmission curves, we further choose 7 inks which are least correlated with each other. We randomly printed the inks on the transparent film by ink-jet printers (two XP-245 printer, one loaded 3 inks and the other 4 inks.). For each printer, the number of colorful inks on one layer each time depends on the loaded ink number, i.e. 3 or 4. In our experiment, we print 6 layers with the 3-inks printer and 5 layers with the 4-inks printer.

Since removing the protective glass in front of the sensor and stick the printed color mask to the sensor may require sophisticated manufacturing technique to avoid potential artifacts, here we leave this part for future work and build a relay system to simply demonstrate our method.

As shown in Fig. 10, the acquisition system we propose includes objective lens, printed mask, relay lens and

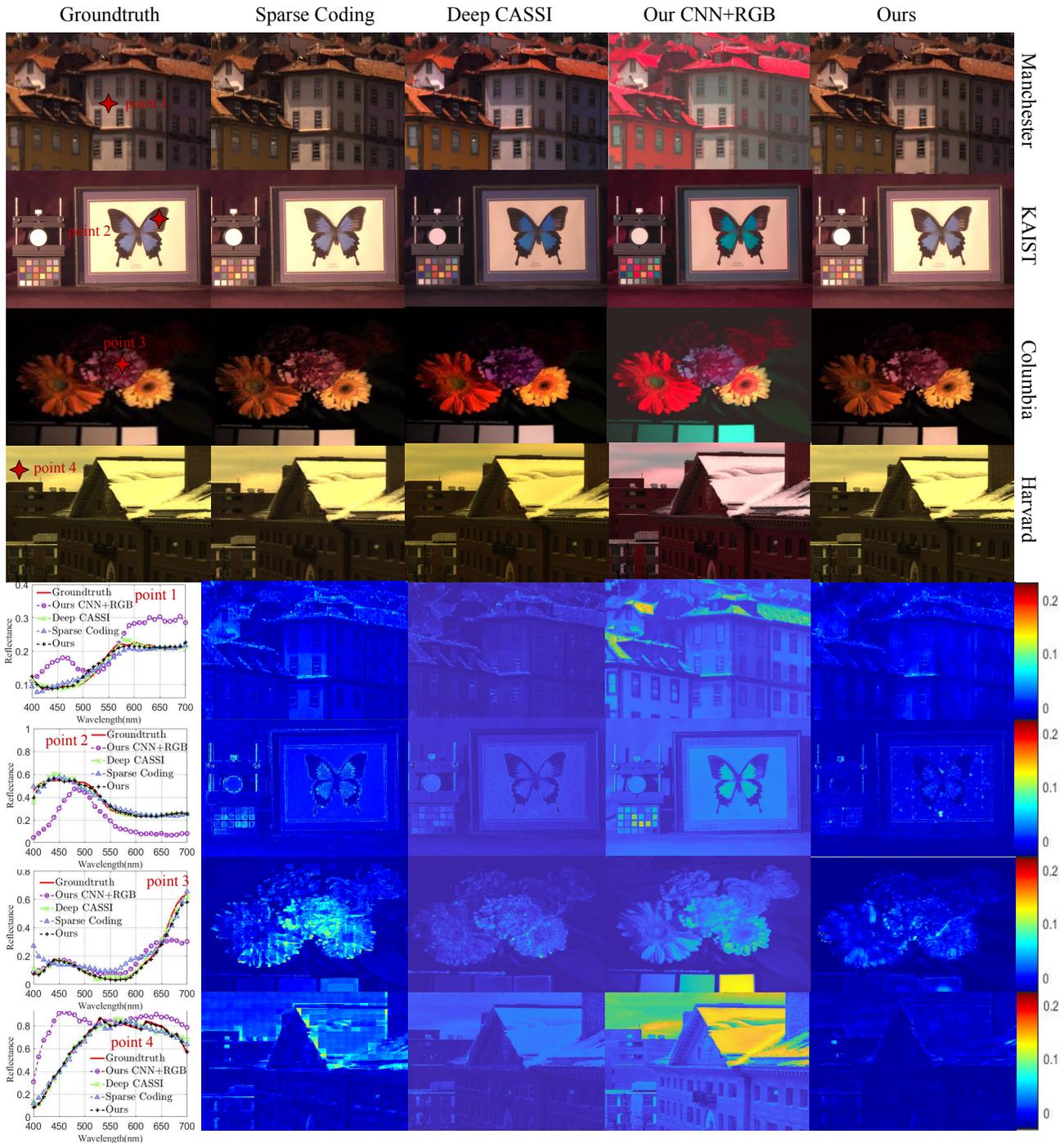


Figure 8. Simulation results on four images, each from a hyperspectral database (Manchester, Columbia, Harvard, KAIST). We compare the synthetic RGB images, the recovered spectrums and error maps with the other three methods.

the imaging sensor. Objective lens is used to focus the light from the scene on the printed mask, spectral information is modulated by the printed mask. The modulated light is finally imaged by the sensor through the relay lens. The spectral transmission response of the printed mask is

calibrated using a high resolution spectrophotometer (with spectral resolution: 0.1nm). We change the emission wavelength of the spectrophotometer and direct the monochromatic light into the integrating sphere to produce a spatially uniform light. The transmitted images are captured every

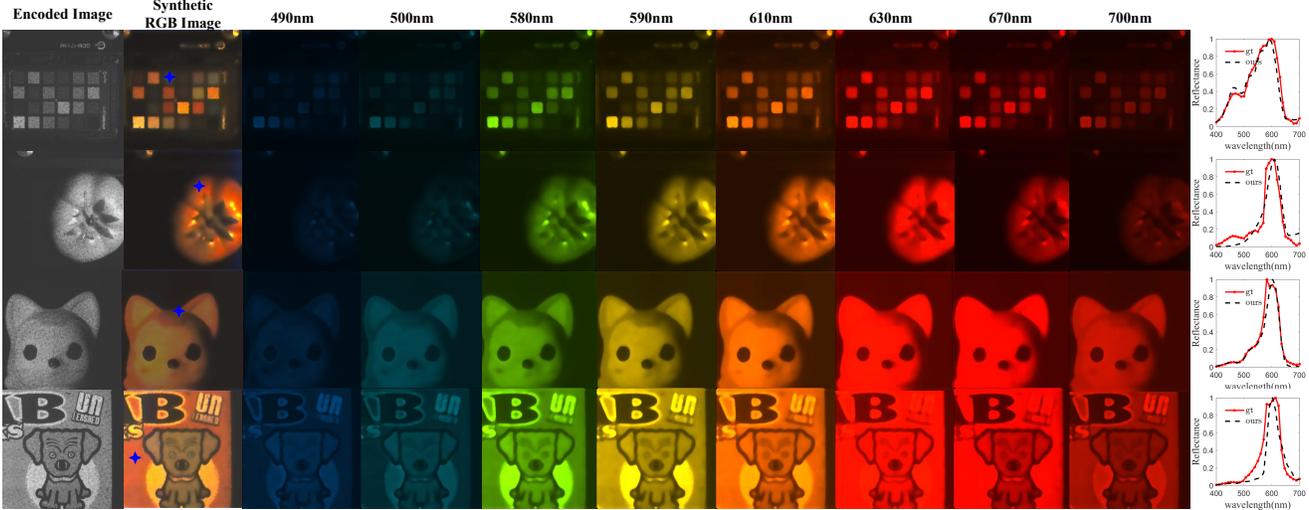


Figure 9. Experimental results of different scenes, the first column are the captured spectral-encoded images, the second column are the synthetic RGB images with the recovered spectrums, the 3-10th columns are the recovered hyperspectral channels, and the last column show the comparisons between the recovered spectrums with the ground truth spectrums captured by the ASD spectrometer.

10 nm from 400 nm to 700 nm. Here we show 8 of the 31 different spectral transmission response image in Fig. 10.

With the calibrated transmission spectral response and the hyperspectral image database, we synthesize 40000 pairs of training data. It took approximately 24 hours to train the network. With the trained network, in physical experiments, we capture the spectrally coded image with the prototype system and recover the hyperspectral image with the trained network. The calibration of the acquisition system is required only once to get the spectral transmission response Φ . Under the irradiation condition of iodine tungsten lamp light source, we collect several coded images as shown in the first column in Fig. 9. The results are shown in Fig. 9. The first column are the captured spectral-encoded images of different scenes. The second column are the synthetic RGB images with the recovered hyperspectral images. The 3-10th columns are the single band spectral images of 490 nm, 500 nm, 580 nm, 590 nm, 610 nm, 630 nm, 670 nm and 700 nm.

To verify the effectiveness of our method, we further

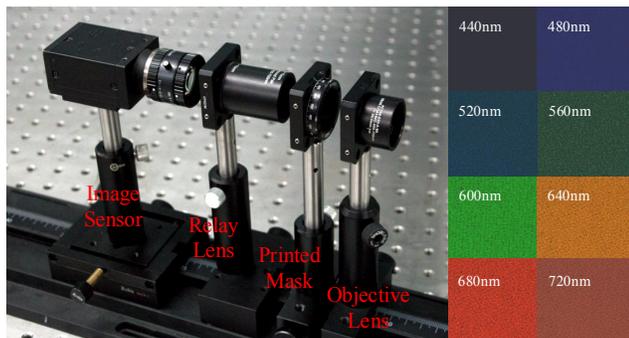


Figure 10. Prototype hyperspectral imaging system and the captured transmission images at different wavelengths for calibration.

capture the groundtruth spectrum with an ASD spectrometer and compare the reconstructed spectrum with the spectrum captured with the ASD. As shown in the last column (marked with blue points in the 2nd column), our method could recover most of the spectral details. In all, we demonstrate that through combing the random color spectral encoding and CNN-based hyperspectral reconstruction, our method could realize high quality hyperspectral imaging.

6. Conclusion

In this paper, we propose a simple and low-budget hyperspectral imaging method. We observe that the transmission spectrum of overlapping printed color mask is the multiplication of the spectral transmission response of each layer, which could introduce a large number of uncorrelated spectral transmission responses. Through printing multilayer random color pattern, we could get an efficient spectral-coding color mask. Combined with the novel spectral coding color mask, we develop a CNN-based network model to recover the hyperspectral information from the coded image with our color mask. As demonstrated in simulated and real captured data, our hyperspectral imaging is of the state-of-the-art spectral retrieving quality.

Future work would be to develop a compact spectrometer based on our method. Although our prototype system is not compact in its current relay-system implementation, our method is indeed promising for a compact hyperspectral imager by attaching the spectral mask on the sensor.

Acknowledgement. This work was supported by Grant No. 61671236 from National Science Foundation of China, Grant No. BK20160634 from National Science Foundation for Young Scholar of Jiangsu Province.

References

- [1] B. Arad and O. Ben-Shahar. Sparse recovery of hyperspectral signal from natural rgb images. In *Proc. ECCV*, pages 19–34. Springer, 2016. 5
- [2] B. Arad and O. Ben-Shahar. Filter selection for hyperspectral estimation. In *Proc. CVPR*, pages 21–26, 2017. 1, 2
- [3] V. Backman, M. B. Wallace, L. Perelman, J. Arendt, R. Gurjar, M. Müller, Q. Zhang, G. Zonios, E. Kline, T. McGilligan, et al. Detection of preinvasive cancer cells. *Nature*, 406(6791):35, 2000. 1
- [4] S.-H. Baek, I. Kim, D. Gutierrez, and M. H. Kim. Compact single-shot hyperspectral imaging using a prism. *ACM Trans. on Graph. (TOG)*, 36(6):217, 2017. 1, 2
- [5] J. Bao and M. G. Bawendi. A colloidal quantum dot spectrometer. *Nature*, 523(7558):67, 2015. 1, 2
- [6] X. Cao, H. Du, X. Tong, Q. Dai, and S. Lin. A prism-mask system for multispectral video acquisition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(12):2423–2435, 2011. 1, 2
- [7] A. Chakrabarti and T. Zickler. Statistics of Real-World Hyperspectral Images. In *Proc. CVPR*, pages 193–200, 2011. 5, 6
- [8] I. Choi, D. S. Jeon, G. Nam, D. Gutierrez, and M. H. Kim. High-quality hyperspectral reconstruction using a spectral prior. *ACM Trans. on Graph. (TOG)*, 36(6):218, 2017. 2, 5, 6
- [9] W. Debskia, P. Walczykowska, A. Klewska, and M. Zyznowskib. Analysis of usage of multispectral video technique for distinguishing objects in real time. In *20th IS-PRS*, 2004. 1
- [10] M. Descour and E. Dereniak. Computed-tomography imaging spectrometer: experimental calibration and reconstruction results. *App. Opt.*, 34(22):4817–4826, 1995. 1, 2
- [11] D. H. Foster, K. Amano, S. M. Nascimento, and M. J. Foster. Frequency of metamerism in natural scenes. *JOSA A*, 23(10):2359–2372, 2006. 5, 6
- [12] R. French, S. Gigan, and O. L. Muskens. Speckle-based hyperspectral imaging combining multiple scattering and compressive sensing in nanowire mats. *Opt. Lett.*, 42(9):1820–1823, 2017. 2
- [13] Y. Fu, T. Zhang, Y. Zheng, D. Zhang, and H. Huang. Joint camera spectral sensitivity selection and hyperspectral image recovery. In *Proc. ECCV*, pages 812–828. Springer, 2018. 2, 5
- [14] Y. Fu, Y. Zheng, I. Sato, and Y. Sato. Exploiting spectral-spatial correlation for coded hyperspectral image restoration. In *Proc. CVPR*, pages 3727–3736, 2016. 2, 5
- [15] N. Gat. Imaging spectroscopy using tunable filters: a review. In *Wave. Appl. VII*, volume 4056, pages 50–65. International Society for Optics and Photonics, 2000. 1
- [16] M. Gehm, R. John, D. Brady, R. Willett, and T. Schulz. Single-shot compressive spectral imaging with a dual-disperser architecture. *Opt. Express*, 15(21):14013–14027, 2007. 1, 2
- [17] G. H. Golub, P. C. Hansen, and D. P. O’Leary. Tikhonov regularization and total least squares. *SIAM J. on Matr. Anal. and App.*, 21(1):185–194, 1999. 4
- [18] R. O. Green, M. L. Eastwood, C. M. Sarture, T. G. Chrien, M. Aronsson, B. J. Chippendale, J. A. Faust, B. E. Pavri, C. J. Chovit, M. Solis, et al. Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (aviris). *Remote Sens.S of Env.*, 65(3):227–248, 1998. 1
- [19] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proc. CVPR*, pages 770–778, 2016. 5
- [20] J. Jia, K. J. Barnard, and K. Hirakawa. Fourier spectral filter array for optimal multispectral imaging. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(4):1530–1543, 2016. 1, 2
- [21] E. R. Kandel, J. H. Schwartz, T. M. Jessell, D. of Biochemistry, M. B. T. Jessell, S. Siegelbaum, and A. Hudspeth. *Principles of neural science*, volume 4. McGraw-hill New York, 2000. 5
- [22] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 5
- [23] P.-J. Lapray, X. Wang, J.-B. Thomas, and P. Gouton. Multispectral filter arrays: Recent advances and practical implementation. *Sensors*, 14(11):21626–21659, 2014. 1, 2
- [24] H. Li, Z. Xiong, Z. Shi, L. Wang, D. Liu, and F. Wu. Hsvcn: Cnn-based hyperspectral reconstruction from rgb videos. In *Proc. ICIP*, pages 3323–3327. IEEE, 2018. 1, 2, 5
- [25] X. Lin, Y. Liu, J. Wu, and Q. Dai. Spatial-spectral encoded compressive hyperspectral imaging. *ACM Trans. on Graph. (TOG)*, 33(6):233, 2014. 1, 2, 5, 6
- [26] S. M. Nascimento, F. P. Ferreira, and D. H. Foster. Statistics of spatial cone-excitation ratios in natural scenes. *JOSA A*, 19(8):1484–1490, 2002. 5, 6
- [27] R. M. Nguyen, D. K. Prasad, and M. S. Brown. Training-based spectral reconstruction from a single rgb image. In *Proc. ECCV*, pages 186–201. Springer, 2014. 1, 2
- [28] S. Nie, L. Gu, Y. Zheng, A. Lam, N. Ono, and I. Sato. Deeply learned filter response functions for hyperspectral reconstruction. In *Proc. CVPR*, pages 4767–4776, 2018. 1, 2
- [29] Z. Pan, G. Healey, M. Prasad, and B. Tromberg. Face recognition in hyperspectral images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(12):1552–1560, 2003. 1
- [30] S. K. Sahoo, D. Tang, and C. Dang. Single-shot multispectral imaging with a monochromatic camera. *Optica*, 4(10):1209–1213, 2017. 2
- [31] A. Wagadarikar, R. John, R. Willett, and D. Brady. Single disperser design for coded aperture snapshot spectral imaging. *App. Opt.*, 47(10):B44–B51, 2008. 1, 2
- [32] L. Wang, Z. Xiong, D. Gao, G. Shi, W. Zeng, and F. Wu. High-speed hyperspectral video acquisition with a dual-camera architecture. In *CVPR*, pages 4942–4950, 2015. 1, 2
- [33] S. Wug Oh, M. S. Brown, M. Pollefeys, and S. Joo Kim. Do it yourself hyperspectral imaging with everyday digital cameras. In *Proc. CVPR*, pages 2461–2469, 2016. 1, 2
- [34] F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar. Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum. *IEEE Trans. on Imag. Proc.*, 19(9):2241–2253, 2010. 5, 6