Optics & Laser Technology 170 (2024) 110265

Contents lists available at ScienceDirect

Optics and Laser Technology

journal homepage: www.elsevier.com/locate/optlastec

Full length article

Joint optimization of splitter pattern and image reconstruction for metasurface-based color imaging systems

Haosen Liu, Edmund Y. Lam *

Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong Special Administrative Region of China

ARTICLE INFO

Dataset link: https://github.com/haosennn/Splitter

Keywords:
Color imaging system
Metasurface
Image sensor
Deep neural networks
Computational imaging

ABSTRACT

The metasurface-based splitters have demonstrated great potential in high-resolution, dark-scene, and ultra-fast imaging due to their excellent capabilities of preserving light intensity, attracting increasingly more research interests. This paper investigates the splitter pattern design, which determines the theoretical performance limit of metasurface-based color imaging systems, to guide their further development. We first establish a unified mathematical model for the splitter-based imaging process. This model reveals that the raw image detected by the splitters is degraded by noise, incomplete sampling, and blurring. Based on this, we develop a deep neural network capable of simultaneously tackling multiple image processing tasks to achieve color image reconstruction. We then conduct evaluation experiments for several existing patterns, through which we establish three design principles to guide the optimization of a splitter pattern. Furthermore, by jointly learning the parameters of our proposed imaging model and reconstruction network, we achieve an automatic optimization of a splitter pattern. The experiments demonstrate that the optimized pattern can achieve better quantitative and visual results than the original pattern, validating the effectiveness of this pattern optimization method.

1. Introduction

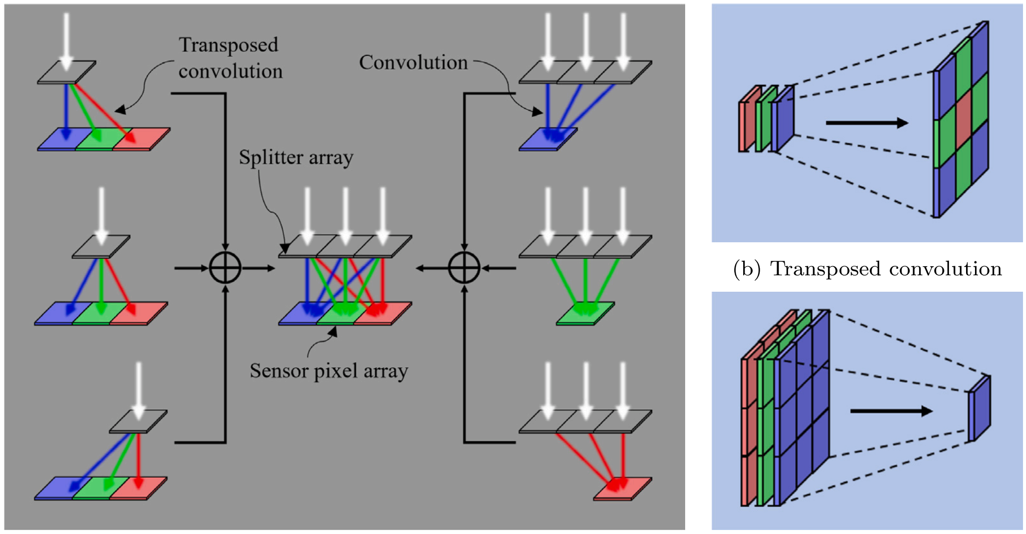
High-sensitivity image sensors are needed for many challenging applications, e.g., high-resolution, dark-scene, and/or ultra-fast imaging, where the incident light received at each sensor pixel is extremely limited. Conventional color image sensors are generally implemented with color filters [1]. They have an inherent drawback that the color filter will reduce the intensity of the incident light, and thus their light sensitivity cannot meet the requirement of those demanding applications. To tackle this problem, a recent development is to replace the color filter with a metasurface-based color splitter [2,3]. Metasurfaces are flat optical components composed of subwavelength structures [4]. By engineering and arranging these individual structures, metasurfaces are capable of tuning key properties of a light wave (e.g., amplitude and phase) at specific frequencies with subwavelength spatial resolution [5]. Consequently, metasurfaces can act as color splitters that can flexibly divert primary colors (i.e., red, green, and blue) at a single position to specified sensor pixels at different positions. An image sensor with such a metasurface-based splitter is expected to hold a light intensity sensitivity roughly three times higher than that with color filters [6], showing great potential in extremely challenging applications.

The design process of a metasurface-based splitter can be divided into three steps. The first step is to determine the splitter pattern, which involves the arrangement of a set of splitters and their focal points for different colors. Then, the second step is to calculate the phase profile based on the given focal points. Finally, the third step is to optimize the subwavelength structures of metasurfaces to approximate the required phase profile. Up to now, most of the existing works focus their attention on the latter two steps. The relation between focal points and the phase profile can be referred to [7]. The subwavelength structure arrangement can be optimized via various optimization methods, such as gradient descent [6,8], adjoint optimization [9–12] and deep learning methods [13]. However, not much prior work has carefully studied the first step, i.e., the splitter pattern design. As such, the design principles are still not clear, and existing patterns are not well analyzed and evaluated. This is problematic since it is the foundation for the subsequent two steps. In fact, the splitter pattern design determines the theoretical performance limit. As such, here we report for the first time a detailed study on the splitter pattern design in order to further promote the development of splitter-based sensors.

Based on the design philosophy of computational imaging [14], the front-end optics (e.g., the splitter pattern) and the post-processing image reconstruction method should be jointly designed. An efficient

* Corresponding author.
E-mail address: elam@eee.hku.hk (E.Y. Lam).

https://doi.org/10.1016/j.optlastec.2023.110265
Received 31 July 2023; Received in revised form 13 October 2023; Accepted 25 October 2023
Available online 6 November 2023



(a) Splitter-based imaging process (b) Transposed convolution (c) Convolution

Fig. 1. Illustration of our proposed mathematical model for splitter-based image sensors. (a) The splitter-based imaging process can be modeled in two equivalent forms. One focuses on how light information is diverted by color splitters, which can be modeled with (b) the transposed convolution operator. The other focuses on how light information is accumulated at image sensors, which can be modeled with (c) the convolution operator.

paradigm to achieve such a joint design is the differentiable imaging [15], which parameterizes both the optical system and reconstruction method with differentiable computational blocks (e.g., neural networks), concatenates these blocks, and jointly optimizes the parameters of all blocks with a specified objective function. This paradigm has been successfully applied to various fields such as engineering the point spread function [16–19], learning the phase-coded aperture for depth estimation [20–23], and so on [24–28]. Inspired by these works, we explore how to apply this paradigm to the splitter pattern design.

The key to applying differentiable imaging here is to parameterize the imaging process of splitter-based sensors and the reconstruction method. To achieve this, we first establish a differentiable mathematical model for the splitter pattern. This model has two equivalent forms, one of which indicates that the raw image detected by splitter-based sensors is degraded by noise, incomplete sampling, and blurring. Therefore, a reconstruction method capable of tackling all of these degradations is desired. Recent studies in the deep learning method have demonstrated that a single network can jointly process multiple image processing tasks [29–33] and generalize to different color filter array patterns [29–31]. Inspired by this, we develop a deep neural network to serve as the reconstruction method. With this network, we conduct an analysis and evaluation of several existing patterns, establishing three pattern design principles, i.e., not using the sub-pixel structure, splitting colors to a region as small as possible, and fully splitting incident light. By further concatenating and jointly optimizing the differentiable model and the reconstruction network, we can achieve an automatic optimization of a splitter pattern. To validate the efficiency of this method, we test it on two splitter patterns. The experimental results show that the two optimized patterns can achieve better quantitative and visual results than their original patterns.

2. Method

2.1. Imaging model for splitter-based sensors

The function of a splitter is to split colors at a position to specified sensor pixels at different positions. It is almost the same as the function of a transposed convolution illustrated in Fig. 1(b), which is to project features from a position to a set of positions [34]. Therefore, it is natural to model the effect of each splitter with the transposed convolution. Besides, as shown in Fig. 1(a), a splitter-based imaging system generally consists of a set of splitters that are not unique at different positions. To specify the positions where each type of splitter is placed, we adopt a sampling operator that preserves values at specified positions and replaces values at other positions with 0. As such, the imaging process of splitter-based system can be modeled as

$$\boldsymbol{g} = \sum_{s\in S} \left(\boldsymbol{f} \downarrow_s \star \boldsymbol{w}_s\right) + \boldsymbol{n}, \tag{1}$$

where S denotes the set of splitters, $\boldsymbol{g} \in \boldsymbol{R}^{H\times W}$ denotes the detected raw image, $\boldsymbol{f} \in \boldsymbol{R}^{H\times W\times 3}$ denotes the color image to be detected, $\boldsymbol{n} \in \boldsymbol{R}^{H\times W}$ denotes the noise, $\star$ denotes the transposed convolution, $\downarrow_s$ and $\boldsymbol{w}_s \in \boldsymbol{R}^{M\times N\times 3}$ denote the sampling operator and parameters of the transposed convolution for the s-th splitter. Furthermore, it is worth noting that this model has an equivalent form as

$$\boldsymbol{g} = \sum_{c\in C} \left(\left(\boldsymbol{f} * \boldsymbol{k}_c\right) \downarrow_c\right) + \boldsymbol{n}, \tag{2}$$

where C denotes the set of colors in the raw image, $*$ denotes the convolution operator illustrated in Fig. 1(c), $\downarrow_c$ and $\boldsymbol{k}_c \in \boldsymbol{R}^{M\times N\times 3}$ denote the sampling operator and the convolution kernel for the c-th color in the raw image. Note that this sampling operator specifies the positions where each type of convolution kernel is applied.

Both of these two forms are illustrated in Fig. 1(a). They are mathematically equivalent, but they hold different meanings. The first one focuses on how light information is diverted by each splitter. Parameters $\boldsymbol{w}_s$ involved in Eq. (1) are directly related to the splitter pattern. Therefore, we implement this form as a splitter network to achieve the joint optimization of the splitter pattern and reconstruction method. By contrast, the second form focuses on how light information is accumulated at each sensor pixel. It can be used to analyze the quality of the raw image captured by detector sensors. More specifically, since Eq. (2) consists of the convolution operator, the sampling operator, and noise variables, it indicates that the detected raw image is blurred among spatial positions and/or color channels, incompletely sampled, and corrupted by noise.

2.2. Reconstruction method

To tackle all of the degradations involved in the splitter-based imaging process and achieve a high-quality recovery of the ground truth color image $\boldsymbol{f}$, we develop a deep convolutional neural network

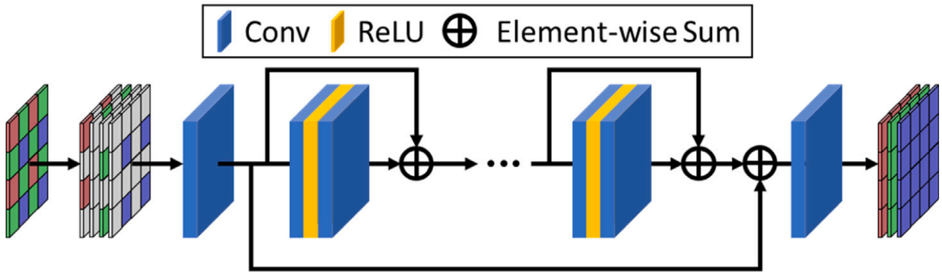


Fig. 2. Reconstruction network. This network first samples and concatenates different colors in the input raw image to restore the spatial invariance. Then, it processes the sampled tensor with 1 convolutional layer, 10 residual blocks, and 1 convolutional layer.

here. As illustrated in Fig. 2, the network first samples the input raw image based on colors, where different colors are put into different channels and the missing color information is padded with 0. This is to convert the original raw image, of which colors are spatially variant, to a tensor holding the spatially invariant color vector. After the sampling layer, the sampled tensor is processed by 1 convolutional layer, 10 residual blocks [35], and 1 convolutional layer. There is a long skip connection between the output of the first convolutional layer and the input of the final convolutional layer. The optimizer and the loss function used for training this network are respectively the Adam algorithm [36] and the mean squared error (MSE) function as

$$\mathcal{L}(\Theta) = \frac{1}{T}\sum_{t=1}^{T} \|\boldsymbol{f}_t - \mathcal{D}(\boldsymbol{g}_t;\Theta)\|^2, \tag{3}$$

where T denotes the number of training images, $\|\cdot\|$ denotes the ℓ_2 norm, and $\mathcal{D}(\cdot;\Theta)$ denotes the reconstruction network with parameters Θ.

2.3. Joint optimization

To further enable the joint optimization of the splitter pattern and the reconstruction network, we implement the imaging model Eq. (1) as a splitter network $\boldsymbol{g} = \mathcal{E}(\cdot;\boldsymbol{w})$, substitute $\boldsymbol{g}$ in the loss function Eq. (3) with this network, and set $\boldsymbol{w}$ as optimization variables. In this way, the loss function is modified as

$$\mathcal{L}(\Theta,\boldsymbol{w}) = \frac{1}{T}\sum_{t=1}^{T} \|\boldsymbol{f}_t - \mathcal{D}(\mathcal{E}(\boldsymbol{f}_t;\boldsymbol{w});\Theta)\|^2. \tag{4}$$

To meet physical requirements, reasonable constraints have to be imposed on parameters $\boldsymbol{w}$ during the joint optimization. Here, we consider two constraints, i.e., (a) elements of $\boldsymbol{w}$ should be non-negative and (b) elements of $\boldsymbol{w}_s$ for each splitter should sum to 1. To impose these two constraints while avoiding the challenging constrained optimization for networks, we obtain $\boldsymbol{w}_s$ by imposing the softmax function $\mathcal{T}(\cdot)$ on auxiliary variables $\boldsymbol{z}_s$, i.e., $\boldsymbol{w}_s = \mathcal{T}(\boldsymbol{z}_s)$, since the outputs of a softmax function can naturally meet the above two constraints. Therefore, the final imaging model and the loss function used for joint optimization are

$$\boldsymbol{g} = \sum_{s\in S} \left(\boldsymbol{f}\downarrow_s \star \mathcal{T}(\boldsymbol{z}_s)\right) + \boldsymbol{n}, \tag{5}$$

and

$$\mathcal{L}(\Theta,\boldsymbol{z}) = \frac{1}{T}\sum_{t=1}^{T} \|\boldsymbol{f}_t - \mathcal{D}(\mathcal{E}(\boldsymbol{f}_t;\mathcal{T}(\boldsymbol{z}));\Theta)\|^2. \tag{6}$$

By initializing $\boldsymbol{z}$ with a splitter pattern and optimizing this loss function with the Adam optimizer, we can achieve an automatic optimization of this splitter pattern.

3. Results and analysis

3.1. Splitter pattern evaluation

Splitter patterns are evaluated using the pipeline illustrated in Fig. 3(a). We first generate degraded raw images for each splitter pattern, then estimate the reference image from each raw image using the reconstruction network shown in Fig. 2, and finally evaluate each recovered image. The degraded raw image $\boldsymbol{g}$ is generated as Eq. (1), where noise $\boldsymbol{n}$ is assumed to be white Gaussian noise and parameters of transposed convolution $\boldsymbol{w}$ are determined by the splitter pattern to be evaluated. The five patterns [2,6,8,37,38] involved in evaluation are illustrated in Figs. 3(b)–3(f). For each pattern, a reconstruction network is trained. The reference images used for training are from the dataset DIV2K [39], and those for testing are from the datasets Kodak24[1] and McMaster [40]. During training, Gaussian noise of standard deviations σ ranging from 0 to 30 is added so that the trained model can be applicable to different noise levels. In each training batch, 80 image patches of size 64 × 64 are extracted and augmented by flipping horizontally or vertically and rotating 90 degrees. The learning rate is set as 10^{-4}. To quantitatively evaluate the quality of recovered images, we adopt peak signal-to-noise ratio (PSNR) and structure similarity index measure (SSIM) as evaluation indices. All of the average index values of each pattern at each test noise level are provided in Table 1. For visual assessments, several representative results are shown in Figs. 4 and 5. By comparing and analyzing the results of all splitter patterns, we establish three principles as follows:

Avoid using the sub-pixel structure. The sub-pixel structure refers to the structure adopted by Pattern B [8] and Pattern C [37]. As illustrated in Figs. 3(c), 3(d), such a structure treats incident light at a set of positions as a whole and splits different colors to its color sub-pixels. This kind of structure is most widely adopted in existing splitter patterns [3,8–11,37]. However, as Table 1 shows, evaluation indices achieved by pattern B and pattern C are the worst among the competing patterns in all cases. Also, the visual results shown in Figs. 4 and 5 demonstrate that these two patterns of the sub-pixel structure tend to result in false textures and/or severe blurring effects. The reason behind this phenomenon is that the sub-pixel structure will lead to a severe loss of position information. Since all of the sub-pixels in a single unit share the same position information, it is impossible to recover the lost position information by using the relations among these sub-pixels. Besides patterns using the sub-pixel structure, other patterns are also likely to result in the blurring effect. One example is the result of pattern A [6] in Fig. 4. This is because colors detected by each sensor pixel are collected from different positions. It is inevitable to result in the loss of position to some extent for all splitter patterns. However, different from the case of the sub-pixel structure, the lost position information for other patterns is likely to be recovered by using the

[1] This dataset can be downloaded from the website: https://r0k.us/graphics/kodak/.

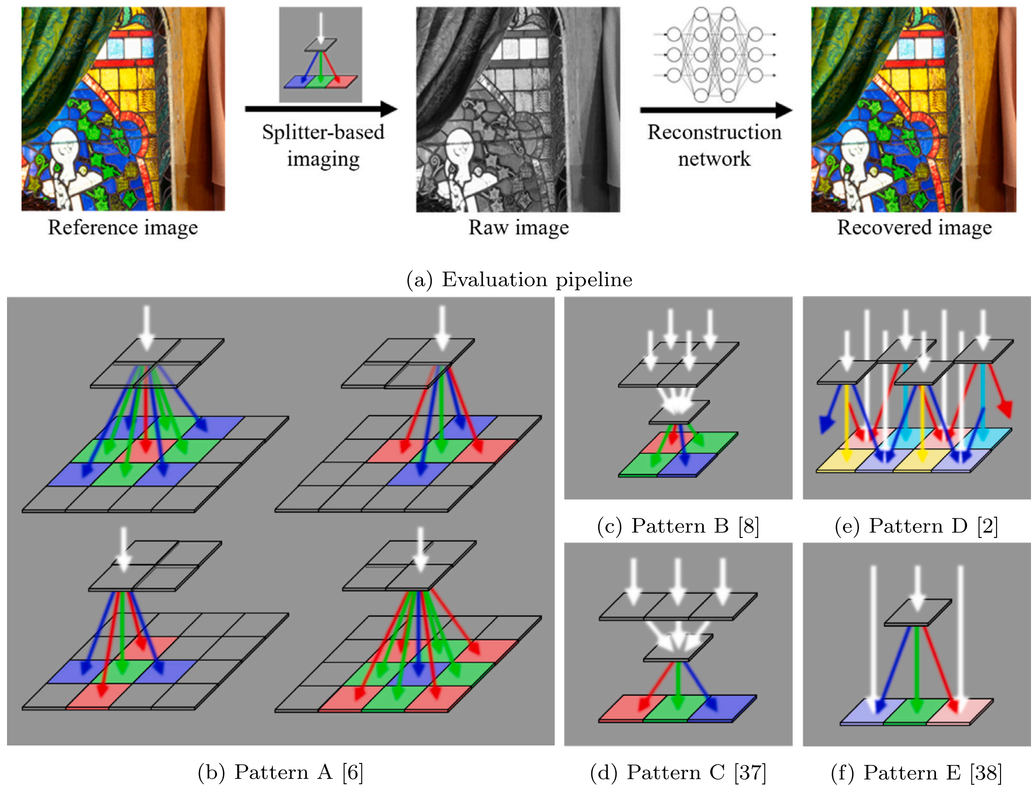


Fig. 3. Illustration of the splitter pattern evaluation experiment, where we use (a) the unified pipeline to evaluate (b)–(f) five existing patterns.

Table 1
Average PSNR (dB) and SSIM of 5 splitter patterns at noise levels $\sigma = 0, 5, 10, 15, 20$ on Kodak24 and McMaster. We rank these patterns based on their average performance. The best one is ranked first.

Pattern	Rank	$\sigma = 0$		$\sigma = 5$		$\sigma = 10$		$\sigma = 15$		$\sigma = 20$	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Pattern A	2	40.02	0.9764	37.58	0.9569	35.72	0.9376	34.44	0.9207	33.47	0.9052
Pattern B	4	32.93	0.9198	32.70	0.9132	32.33	0.9035	31.93	0.8937	31.54	0.8836
Pattern C	5	31.53	0.9159	31.37	0.9088	31.08	0.8986	30.77	0.8881	30.45	0.8773
Pattern D	1	40.35	0.9773	38.56	0.9650	36.62	0.9491	35.17	0.9341	34.02	0.9195
Pattern E	3	35.65	0.9674	34.84	0.9552	33.89	0.9404	33.08	0.9264	32.37	0.9125

relations among neighboring sensor pixels. Owing to this, the blurring effect of patterns not using the sub-pixel structure, such as pattern A, is not as severe as that of the two patterns using that structure, which can be observed by comparing Fig. 5(b)–(d).

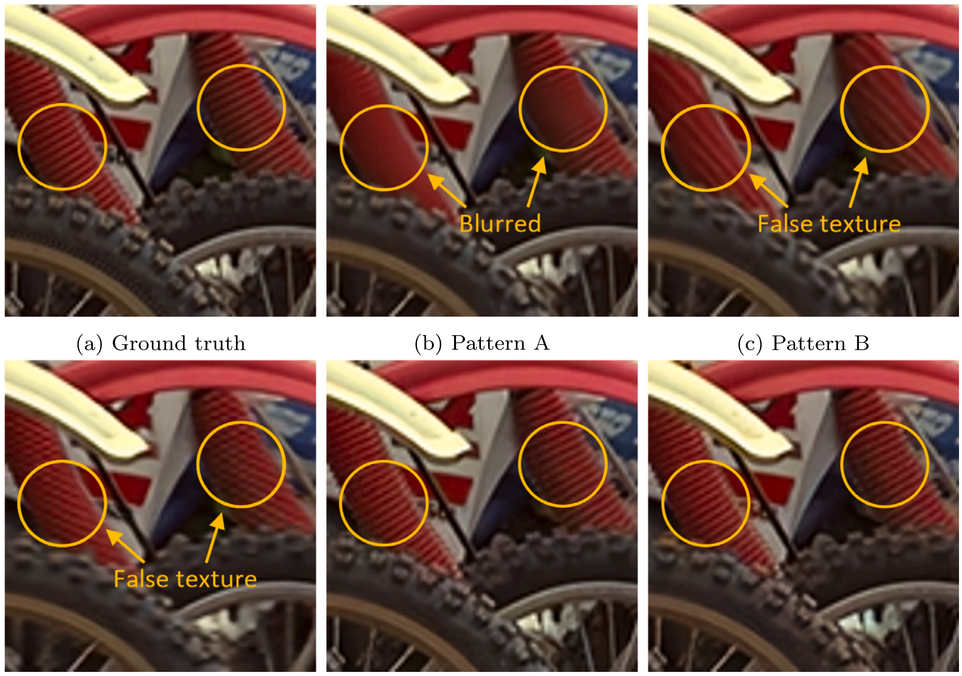
Colors should be split to a region as small as possible. Among patterns without using the sub-pixel structure, it can be observed that there is almost no blurring effect in the results of pattern D [2] and pattern E [38], while that of pattern A is a bit obvious, as shown in marked regions of Fig. 4. The reason behind this phenomenon is that the former two splitter patterns only split colors at a position to 1×2 or 1×3 pixels, while the latter one splits colors to 3×3 pixels, as illustrated in Figs. 3(b), 3(e), 3(f). In other words, for the latter splitter, colors from more positions are mixed up, and thus it is more difficult for the reconstruction method to identify the position from which each part of the colors comes. This indicates that, for the splitter pattern design, colors should be split to a region as small as possible to tackle the potential blurring effect.

The incident light should be fully split. Besides the blurring effect among spatial positions, another potential problem for a splitter-based system is the blurring effect among color channels, which might lead to false colors in the recovered images. Two examples are marked regions in the results of pattern D and pattern E in Fig. 5, where 'red' flowers tend to be 'pink' or even 'gray'. This should be owing to that a great ratio of incident light is not split in these two splitter patterns, making different colors mixed and hard to separate. To tackle this problem, we suppose one intuitive solution should be to use more splitters to fully split incident light so that color information can be better separated during the imaging process.

3.2. Joint optimization of splitter pattern and image reconstruction

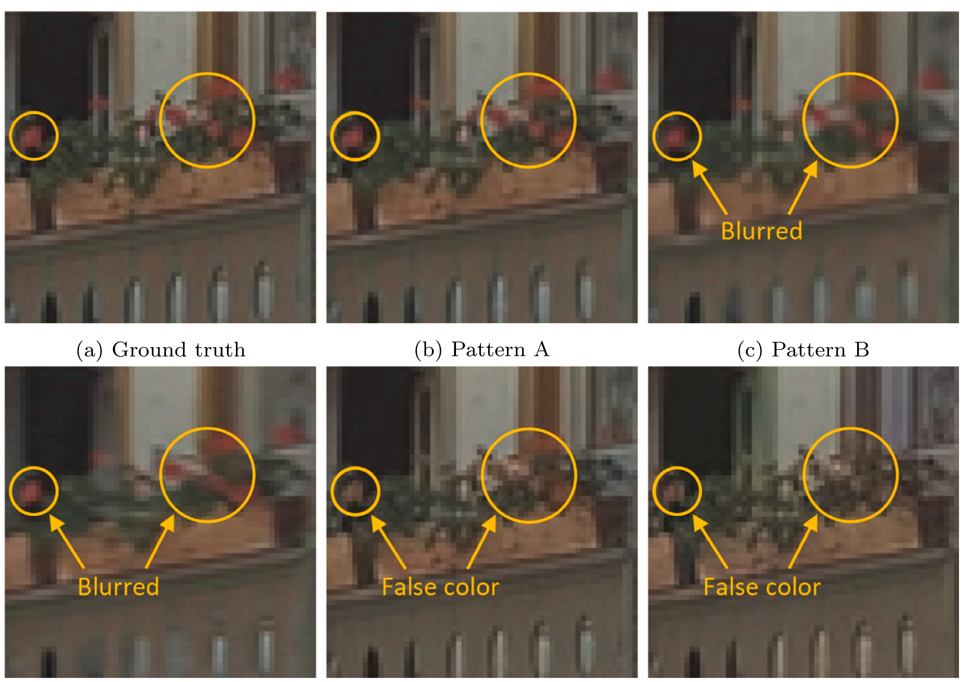
To jointly optimize the splitter pattern and the reconstruction network, the joint network illustrated in Fig. 6(a) is adopted. The splitter network is implemented based on our proposed differentiable forward imaging model shown in Eq. (5). The reconstruction network, together with the experimental setting including datasets used for training and testing, are just the same as the ones we use for the pattern evaluation experiments. Here, we initialize the learnable parameters $\boldsymbol{z}$ based on two patterns, i.e., pattern A and pattern D, to test whether the proposed joint optimization method can improve existing patterns. The comparisons between these two patterns and their optimized patterns are provided as follows:

Optimization of pattern A. As Table 2 shows, on average, the optimized pattern achieves better performance than the original pattern in terms of both PSNR and SSIM in all cases. The PSNR improvement is up to 0.78 dB. Besides, as Fig. 7 shows, the optimized pattern successfully tackles the blurring problem of the original pattern. By comparing Figs. 3(b) and 6(b), it can be observed that the modification



(a) Ground truth (b) Pattern A (c) Pattern B
(d) Pattern C (e) Pattern D (f) Pattern E

Fig. 4. Visual results of the splitter pattern evaluation experiment on zoomed regions of the image 'kodim05' at the noise level $\sigma = 10$.



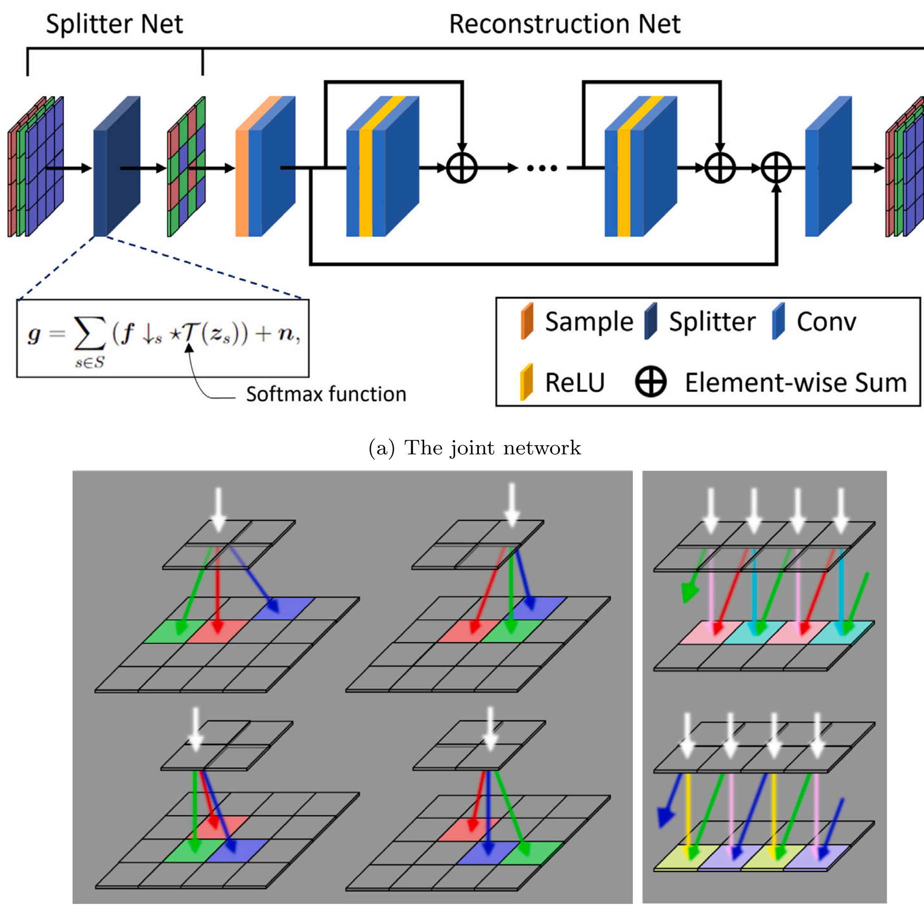
(a) Ground truth (b) Pattern A (c) Pattern B
(d) Pattern C (e) Pattern D (f) Pattern E

Fig. 5. Visual results of the splitter pattern evaluation experiment on zoomed regions of the image 'kodim24' at the noise level $\sigma = 5$.

of the optimized pattern over the original pattern follows the second design principle we establish above, i.e., colors should be split to a region as small as possible. In Fig. 3(b), a single color at a position is equally split to a set of sensor pixels. By contrast, in Fig. 6(b), a single color at a position is only diverted to a single sensor pixel.

Optimization of pattern D. As Table 3 shows, on average, the optimized pattern achieves better performances than the original pattern in terms of both PSNR and SSIM in all cases. The PSNR improvement is up to 0.75 dB. Besides, as Fig. 8 shows, the optimized pattern is capable of correcting the false color caused by the original pattern. The comparison between Figs. 3(e) and 6(c) indicates that the modification of the optimized pattern over the original pattern follows both the second and third design principles. On the one side, the red color and blue color are diverted to single sensor pixels, respectively. On the other hand, parts of green color are diverted to the neighboring pixel from the incident light that is originally not split.

In conclusion, our proposed joint optimization method successfully improves existing splitter patterns, validating its effectiveness.



(a) The joint network

(b) Optimized pattern A (c) Optimized pattern D

Fig. 6. Illustration of the joint optimization experiment, where (a) the joint network is trained to automatically optimize a splitter pattern. (b) Optimized result of pattern A. (c) Optimized result of pattern D. The parameters of these optimized patterns can be found in our publicly available trained models.

Table 2
Average PSNR (dB) and SSIM of pattern A and its optimized pattern at noise levels $\sigma = 0, 5, 10, 15, 20$ on Kodak24 and McMaster.

Pattern	$\sigma = 0$		$\sigma = 5$		$\sigma = 10$		$\sigma = 15$		$\sigma = 20$	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Pattern A	40.01	0.9764	37.57	0.9569	35.72	0.9376	34.44	0.9206	33.47	0.9052
Optimized	40.33	0.9772	38.38	0.9631	36.57	0.9468	35.27	0.9319	34.25	0.9178

Table 3
Average PSNR (dB) and SSIM of pattern D and its optimized pattern at noise levels $\sigma = 0, 5, 10, 15, 20$ on Kodak24 and McMaster.

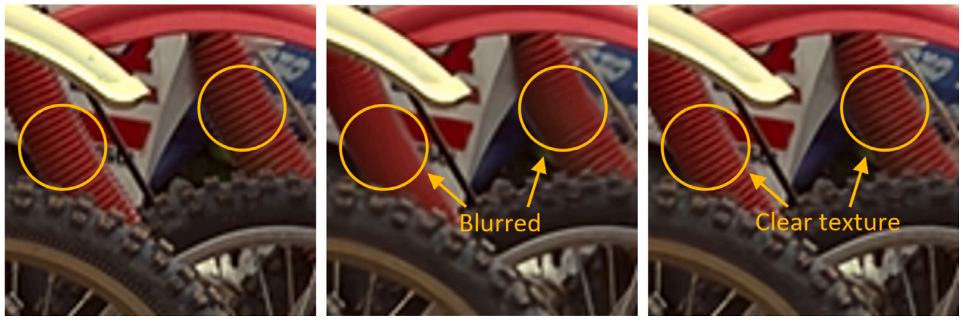
Pattern	$\sigma = 0$		$\sigma = 5$		$\sigma = 10$		$\sigma = 15$		$\sigma = 20$	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Pattern D	40.35	0.9772	38.56	0.9650	36.61	0.9491	35.16	0.9341	34.02	0.9195
Optimized	41.03	0.9797	38.98	0.9668	36.96	0.9508	35.49	0.9361	34.34	0.9219

Interestingly, although not specifically designed, the improvements of the optimized patterns over the original patterns follow our proposed design principles, which can demonstrate that these principles are reasonable.

3.3. More studies

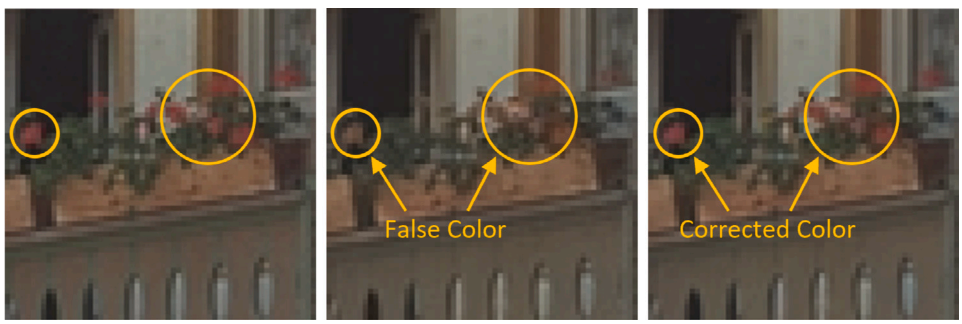
To evaluate our proposed method from more aspects, this subsection provides more studies, including comparison with the filter-based color imaging system, ablation study on the initialization of z, and evaluation with the quantitative indicator CIEDE2000. Here, pattern D and its optimized pattern are adopted as representatives to participate in the evaluation. Besides these two, the filter-based imaging system and the other optimized pattern, z of which is randomly initialized, are also evaluated. The comparison results are visualized in Fig. 9 and summarized as follows:

Evaluation with CIEDE2000. The CIEDE2000 is an indicator used to measure the color difference between two RGB images. It can be implemented with the MATLAB function *imcolordiff*. As Fig. 9(b) shows, the optimized pattern achieves smaller CIEDE2000 values than pattern D, implying that the optimized pattern can better recover the color information. This conclusion is consistent with our observation from



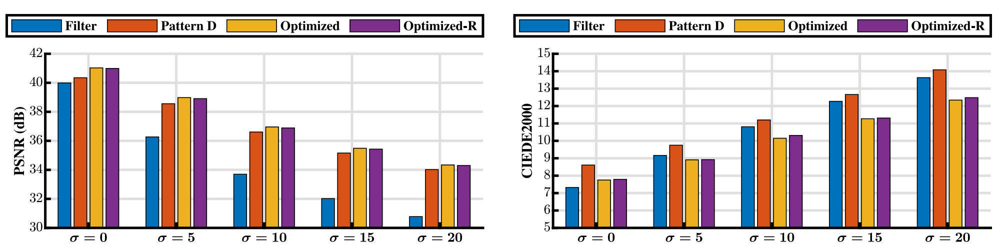
(a) Ground truth (b) Pattern A (c) Optimization for pattern A

Fig. 7. Visual results of pattern A and its optimized pattern on zoomed regions of the image 'kodim05' at the noise level $\sigma = 10$.



(a) Ground truth (b) Pattern D (c) Optimization for pattern D

Fig. 8. Visual results of pattern D and its optimized pattern on zoomed regions of the image 'kodim24' at the noise level $\sigma = 5$.



(a) PSNR comparison (b) CIEDE2000 comparison

Fig. 9. Performance comparison among a filter-based system, splitter-based systems with pattern D, and two optimized patterns with different initialization methods. The 'Optimized' is initialized with pattern D, while 'Optimized-R' is with random initialization. For the indicator CIEDE2000, a smaller value indicates a better performance.

Fig. 8 that the optimized pattern can correct the false color caused by its original pattern.

Comparison with filter-based system. To incorporate the filter-based system into comparison, the widely adopted Bayer color filter array [41] is used as a representative. From Fig. 9, it can be observed that pattern D performs better than the Bayer pattern in terms of PSNR, while worse in terms of CIEDE2000. However, the drawback of pattern D is greatly overcome after optimization with our proposed method. Its optimized pattern outperforms the Bayer pattern in terms of both PSNR and CIEDE2000.

Ablation study on the initialization of z. In our default setting, z is initialized with existing patterns. To study the effect of initialization, a randomly initialized z is evaluated here. Note that the dimensions of z are kept the same in two initialization methods. From Fig. 9, it can be observed that initializing z with pattern D and with random values achieve similar performances, implying that the initialization of z only has weak effects on our proposed method.

4. Conclusion and outlook

In summary, this work is to tackle the splitter pattern design problem. We first establish a differentiable mathematical model for splitter-based image sensors, showing that the detected raw image is degraded by noise, incomplete sampling, and blurring. Because of this, we develop a deep neural network capable of simultaneously tackling multiple image-processing tasks to achieve image reconstruction. With this reconstruction network, we conduct an evaluation of five existing patterns, through which we establish three splitter pattern design principles. Furthermore, we implement our proposed differentiable model with a splitter network and concatenate it with the reconstruction network so that we can achieve an automatic optimization of a splitter pattern via jointly training these two networks. We apply this method to two existing splitters. In both cases, the optimized pattern achieves better performance than the original pattern, demonstrating the effectiveness of our proposed joint optimization method.

This work can be further extended. Here, we have proposed a method to jointly optimize the splitter pattern and the reconstruction network, but there is still a gap between the splitter pattern design and the metasurface design. A more interesting and more challenging extension is to close this gap by directly optimizing the desired phase profiles, or even the subwavelength structures of metasurface-based splitters, together with the reconstruction method. To achieve this, functions mapping from them to splitter patterns or detected raw images have to be established. Also, more constraints should be incorporated based on physical requirements.

CRediT authorship contribution statement

Haosen Liu: Conceptualization, Investigation, Formal analysis, Methodology, Software, Writing – original draft, Writing – review & editing. **Edmund Y. Lam:** Conceptualization, Funding acquisition, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Our source codes and trained models are publicly available at the linkage https://github.com/haosennn/Splitter.

Acknowledgments

This work is supported in part by the Research Grants Council of Hong Kong (GRF 17201620, 17200321). The authors would like to thank Mr. Chutian Wang, Dr. Li Song and Dr. Shuo Zhu for useful discussions about this paper.

References

[1] Y. Zhang, G. Wang, J. Xu, Z. Shi, D. Dong, The modified gradient edge detection method for the color filter array image of the CMOS image sensor, Opt. Laser Technol. 62 (2014) 73–81.
[2] S. Nishiwaki, T. Nakamura, M. Hiramoto, T. Fujii, M.-a. Suzuki, Efficient colour splitters for high-pixel-density image sensors, Nat. Photonics 7 (3) (2013) 240–246.
[3] P. Wang, R. Menon, Ultra-high-sensitivity color imaging via a transparent diffractive-filter array and computational optics, Optica 2 (11) (2015) 933–939.
[4] J. Hu, S. Bandyopadhyay, Y.-h. Liu, L.-y. Shao, A review on metasurface: From principle to smart metadevices, Front. Phys. 8 (2021) 586087.
[5] A.C. Overvig, S. Shrestha, S.C. Malek, M. Lu, A. Stein, C. Zheng, N. Yu, Dielectric metasurfaces for complete and independent control of the optical amplitude and phase, Light: Sci. Appl. 8 (1) (2019) 92.
[6] M. Miyata, N. Nemoto, K. Shikama, F. Kobayashi, T. Hashimoto, Full-color-sorting metalenses for high-sensitivity image sensors, Optica 8 (12) (2021) 1596–1604.
[7] B.H. Chen, P.C. Wu, V.-C. Su, Y.-C. Lai, C.H. Chu, I.C. Lee, J.-W. Chen, Y.H. Chen, Y.-C. Lan, C.-H. Kuan, et al., GaN metalens for pixel-level full-color routing at visible light, Nano Lett. 17 (10) (2017) 6345–6352.
[8] C. Kim, J. Hong, B. Lee, Visible color routing metasurface using dual focal phase profile in green light, in: Holography, Diffractive Optics, and Applications XI, Vol. 11898, SPIE, 2021, pp. 224–228.
[9] P. Camayd-Muñoz, C. Ballew, G. Roberts, A. Faraon, Multifunctional volumetric meta-optics for color and polarization image sensors, Optica 7 (4) (2020) 280–283.
[10] N. Zhao, P.B. Catrysse, S. Fan, Perfect RGB-IR color routers for sub-wavelength size CMOS image sensor pixels, Adv. Photonics Res. 2 (3) (2021) 2000048.
[11] E. Johlin, Nanophotonic color splitters for high-efficiency imaging, Iscience 24 (4) (2021) 102268.
[12] S. Molesky, Z. Lin, A.Y. Piggott, W. Jin, J. Vucković, A.W. Rodriguez, Inverse design in nanophotonics, Nat. Photonics 12 (11) (2018) 659–670.
[13] J. Cheng, R. Li, Y. Wang, Y. Yuan, X. Wang, S. Chang, Inverse design of generic metasurfaces for multifunctional wavefront shaping based on deep neural networks, Opt. Laser Technol. 159 (2023) 109038.
[14] T. Zeng, E.Y. Lam, Robust reconstruction with deep learning to handle model mismatch in lensless imaging, IEEE Trans. Comput. Imaging 7 (2021) 1080–1092.
[15] N. Chen, T.-C. Poon, L. Cao, B. Lee, E.Y. Lam, Differentiable imaging: A new tool for computational imaging, Adv. Phys. Res. 2 (6) (2023) 2200118.
[16] J. Page, P. Favaros, Learning to model and calibrate optics via a differentiable wave optics simulator, in: IEEE International Conference on Image Processing, IEEE, 2020, pp. 2995–2999.
[17] A. Klauss, F. Conrad, C. Hille, Binary phase masks for easy system alignment and basic aberration sensing with spatial light modulators in STED microscopy, Sci. Rep. 7 (1) (2017) 15699.
[18] T. Zhao, T. Mauger, G. Li, Optimization of wavefront-coded infinity-corrected microscope systems with extended depth of field, Biomed. Opt. Express 4 (8) (2013) 1464–1471.
[19] T. Liaudat, J.-L. Starck, M. Kilbinger, P.-A. Frugier, Rethinking data-driven point spread function modeling with a differentiable optical model, Inverse Problems 39 (3) (2023) 035008.
[20] H. Haim, S. Elmalem, R. Giryes, A.M. Bronstein, E. Marom, Depth estimation from a single image using deep learned phase coded mask, IEEE Trans. Comput. Imaging 4 (3) (2018) 298–310.
[21] Y. Wu, V. Boominathan, H. Chen, A. Sankaranarayanan, A. Veeraraghavan, PhaseCam3D—Learning phase masks for passive single view depth estimation, in: IEEE International Conference on Computational Photography, IEEE, 2019, pp. 1–12.
[22] J. Chang, G. Wetzstein, Deep optics for monocular depth estimation and 3D object detection, in: IEEE/CVF International Conference on Computer Vision, 2019, pp. 10193–10202.
[23] H. Ikoma, C.M. Nguyen, C.A. Metzler, Y. Peng, G. Wetzstein, Depth from defocus with learned optics for imaging and occlusion-aware depth estimation, in: IEEE International Conference on Computational Photography, IEEE, 2021, pp. 1–12.
[24] C. Wang, N. Chen, W. Heidrich, Towards self-calibrated lens metrology by differentiable refractive deflectometry, Opt. Express 29 (19) (2021) 30284–30295.
[25] C. Wang, N. Chen, W. Heidrich, dO: A differentiable engine for Deep Lens design of computational imaging systems, IEEE Trans. Comput. Imaging 8 (2022) 905–916.
[26] Q. Sun, C. Wang, Q. Fu, X. Dun, W. Heidrich, End-to-end complex lens design with differentiable ray tracing, ACM Trans. Graph. 40 (4) (2021) 1–13.
[27] E. Tseng, A. Mosleh, F. Mannan, K. St-Arnaud, A. Sharma, Y. Peng, A. Braun, D. Nowrouzezahrai, J.-F. Lalonde, F. Heide, Differentiable compound optics and processing pipeline optimization for end-to-end camera design, ACM Trans. Graph. 40 (2) (2021) 1–19.
[28] N. Chen, C. Wang, W. Heidrich, ∂H: Differentiable holography, Laser Photonics Rev. 17 (9) (2023) 2200828.
[29] M. Gharbi, G. Chaurasia, S. Paris, F. Durand, Deep joint demosaicking and denoising, ACM Trans. Graph. 35 (6) (2016) 1–12.
[30] F. Kokkinos, S. Lefkimmiatis, Deep image demosaicking using a cascade of convolutional residual denoising networks, in: European Conference on Computer Vision, 2018, pp. 303–319.
[31] F. Kokkinos, S. Lefkimmiatis, Iterative joint image demosaicking and denoising using a residual denoising network, IEEE Trans. Image Process. 28 (8) (2019) 4177–4188.
[32] L. Liu, X. Jia, J. Liu, Q. Tian, Joint demosaicing and denoising with self guidance, in: IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 2240–2249.
[33] W. Xing, K. Egiazarian, End-to-end learning for joint image demosaicing, denoising and super-resolution, in: IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 3507–3516.
[34] V. Dumoulin, F. Visin, A guide to convolution arithmetic for deep learning, 2016, arXiv preprint arXiv:1603.07285.
[35] B. Lim, S. Son, H. Kim, S. Nah, K. Mu Lee, Enhanced deep residual networks for single image super-resolution, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2017, pp. 136–144.
[36] D.P. Kingma, J. Ba, Adam: A method for stochastic optimization, 2014, arXiv preprint arXiv:1412.6980.
[37] M. Chen, L. Wen, D. Pan, D.R. Cumming, X. Yang, Q. Chen, Full-color nanorouter for high-resolution imaging, Nanoscale 13 (30) (2021) 13024–13029.
[38] M. Miyata, M. Nakajima, T. Hashimoto, High-sensitivity color imaging using pixel-scale color splitters based on dielectric metasurfaces, ACS Photonics 6 (6) (2019) 1442–1450.
[39] E. Agustsson, R. Timofte, NTIRE 2017 challenge on single image super-resolution: Dataset and study, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2017, pp. 126–135.
[40] L. Zhang, X. Wu, A. Buades, X. Li, Color demosaicking by local directional interpolation and nonlocal adaptive thresholding, J. Electron. Imaging 20 (2) (2011) 023016.
[41] B. Bayer, Color Imaging Array, United States Patent, No. 3971065, 1976.