

Spatially Varying Exposure With 2-by-2 Multiplexing: Optimality and Universality

Xiangyu Qu , *Student Member, IEEE*, Yiheng Chi , *Graduate Student Member, IEEE*,
and Stanley H. Chan , *Senior Member, IEEE*

Abstract—The advancement of new digital image sensors has enabled the design of exposure multiplexing schemes where a single image capture can have multiple exposures and conversion gains in an interlaced format, similar to that of a Bayer color filter array. In this article, we ask the question of how to design such multiplexing schemes for *adaptive* high-dynamic range (HDR) imaging where the multiplexing scheme can be updated according to the scenes. We present two new findings. 1) We address the problem of *design optimality*. We show that given a multiplex pattern, the conventional optimality criteria based on the input/output-referred signal-to-noise ratio (SNR) of the independently measured pixels can lead to flawed decisions because it cannot encapsulate the location of the saturated pixels. We overcome the issue by proposing a new concept known as the spatially varying exposure risk (SVE-Risk) which is a pseudo-idealistic quantification of the amount of recoverable pixels. We present an efficient enumeration algorithm to select the optimal multiplex patterns. 2) We report a *design universality* observation that the design of the multiplex pattern can be decoupled from the image reconstruction algorithm. This is a significant departure from the recent literature that the multiplex pattern should be jointly optimized with the reconstruction algorithm. Our finding suggests that in the context of exposure multiplexing, an end-to-end training may not be necessary.

Index Terms—High dynamic range imaging, spatially varying exposure, exposure multiplexing, computational photography.

I. INTRODUCTION

DIGITAL image sensors today, at least for the majority of them, pick and choose a *global* exposure and conversion gain across the entire pixel array to control the amount of photon flux reaching the sensor. For high dynamic range (HDR) scenes, this global configuration requires the camera to capture a bracket of exposures and use post-processing algorithms to fuse an HDR image. However, in the presence of motion and noise, HDR fusion is known to be difficult.

Approximately two decades ago, Nayar and Mitsunaga proposed the idea of spatially multiplexing the exposure and conversion gain [1]. The argument was that we could capture

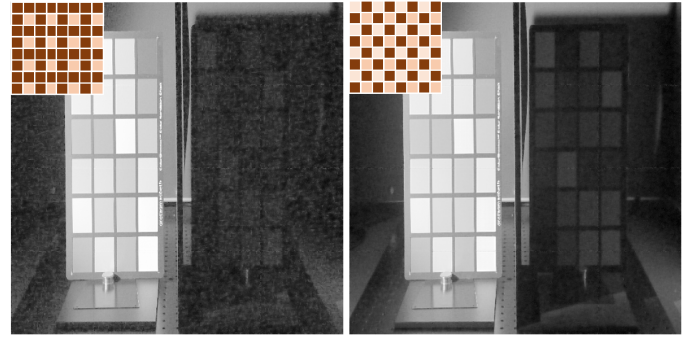


Fig. 1. Given an HDR scene, the choice of spatially-varying exposure pattern significantly influences the quality of the reconstructed image. We propose a new risk estimator to determine the optimal exposure pattern for capturing high dynamic range scenes. [Left] HDR scene reconstruction using an arbitrary spatially-varying exposure pattern. [Right] HDR scene reconstruction using the optimal exposure pattern for the scene.

multiplexed exposures like color filter arrays in a single-shot to avoid the motion problem. The reduction of the spatial resolution can be, in principle, recovered by an appropriately designed interpolation algorithm. Nayar and Mitsunaga's idea led to a series of very interesting work in coded exposures, including adaptive control schemes [2], hardware multi-bucket sensor designs [3], and some of the most recent works in co-optimizing the multiplex pattern and the reconstruction algorithm via deep learning [4], [5].

Despite the large number of prior work, we seldom ask the question of how to *design* the multiplex pattern. This is a meaningful question, because a poorly chosen multiplex pattern can produce a substantially worse image than the optimal one, as illustrated in Fig. 1. However, if we want to choose the optimal pattern, we must first answer the question:

What is the *optimality criteria* for exposure multiplex?

“Optimality criteria” may seem trivial — just pick a pattern that maximizes the signal-to-noise ratio (SNR)! But SNR of what? If it is the SNR of the *measured* pixel values, then we need a way to quantify the locations of the saturated pixels: a group of sparsely located saturated pixels are easy to recover (think about the color filter arrays) whereas a group of densely concentrated pixels are hard to recover (think about inpainting a large hole in an image). How about the SNR of the *reconstructed* image? This

Manuscript received 29 June 2023; revised 12 November 2023; accepted 27 December 2023. Date of publication 22 January 2024; date of current version 19 February 2024. This work was supported by the National Science Foundation under Grants 2030570, 1763896, 2133032, and 2134209. The Associate Editor coordinating the review of this manuscript and approving it for publication was Dr. M. Salman Asif. (Corresponding author: Yiheng Chi.)

The authors are with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907 USA (e-mail: qu27@purdue.edu; chi14@purdue.edu; stanchan@purdue.edu).

This article has supplementary downloadable material available at <https://doi.org/10.1109/TCI.2024.3354426>, provided by the authors.

Digital Object Identifier 10.1109/TCI.2024.3354426

seems more plausible because if a final reconstruction has the highest PSNR, then the pattern is optimal. However, to compare the final reconstruction of every pattern, we need to first capture the scene using every possible pattern. This exhaustive capturing and reconstruction approach defeats the purpose of finding a good multiplex pattern; if one already has all possible captures and associated reconstructions, why bother finding the pattern to be used? Therefore, as we can see, the above seemingly easy task can quickly evolve into a challenging research problem.

The goal of this article is to clarify the difficulties and propose solutions. Our two contributions are:

- We introduce the concept of spatially varying exposure risk (SVE-Risk). SVE-Risk is customized to measure the usefulness of a multiplex pattern. It is universal in the sense that the risk does not require knowledge about a particular image reconstruction algorithm but can still be used to predict what pattern will likely yield the optimal final image for the given scene. We also propose efficient computation methods for SVE-Risk and demonstrate practicality of using our suite of algorithms for multiplex pattern selection.
- We discover, through large-scale experiments, that a multiplex pattern and the image reconstruction algorithm do not need to be co-optimized. This is a departure from prior work that argue the necessity of co-optimization. We hope that our finding can stimulate discussions about the future sensor-algorithm co-design problems.

II. PROBLEM STATEMENT AND RELATED WORK

In this section we define notations, state the problem, and comment on the related work.

A. Problem Statement and Notations

We let $\theta = [\theta_1, \dots, \theta_N] \in \mathbb{R}^N$ be the ground truth image denoting the incoming photon flux, and let $y \in \mathbb{R}^N$ be the observed photoelectric signal produced by the sensor. If the sensor uses a global exposure τ and a global gain α to capture the image y , then y can be simulated according to the equation

$$y(\alpha, \tau) = \text{ADC} \{ \text{Clip} \{ \text{CRF} \{ (\mathcal{P}[\tau\theta \cdot \text{QE}] + \tau\mu_{\text{dark}}) \cdot \alpha \} \} + \mathcal{N}(0, \sigma_{\text{read}}^2) \}. \quad (1)$$

We regard (1) as the first order approximation to the actual image sensor. The symbols we used here are defined in Table I.

The concept of multiplexing is to assign, periodically, an exposure pattern and a gain pattern such that each pixel will be subject to a different exposure and gain. Mathematically, for a 2×2 exposure and gain pattern, we define

$$\tau = \{\tau_\ell\} = \begin{bmatrix} \tau_1 & \tau_2 \\ \tau_3 & \tau_4 \end{bmatrix} \quad \alpha = \{\alpha_\ell\} = \begin{bmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{bmatrix}$$

where each τ_ℓ is sampled from a set of exposure levels, e.g., $\{1, 2, 4, 8, \dots\}$. The same holds for the gain α_ℓ .

The core research question we ask in this article is the choice of τ and α .

TABLE I
IMAGE SENSOR MODEL PARAMETERS

Symbol	Meaning	Typical Value
ADC	Analog-digital	10~12-bit
Clip	Full well limit	5000 e-
τ	Exposure	10ms, 20ms, 40ms
μ_{dark}	dark current	0.002 e-/s
QE	Quantum efficiency	80%
α	Conversion gain	1
σ_{read}	Read noise	0.3 e-
CRF	Camera Response Function	
$\mathcal{P}$	Poisson distribution	
$\mathcal{N}$	Gaussian distribution	

Given the scene radiance θ , how do we select $\tau = \{\tau_\ell\}$ and $\alpha = \{\alpha_\ell\}$ to generate a y such that the reconstructed image $\hat{\theta}(y)$ has the highest PSNR?

Why limit to 2×2 patterns? Readers coming from an algorithmic perspective might think that limiting to 2×2 is too restrictive. Why not analyze 4×4 or $N \times N$? The main reason is that 2×2 patterns are more hardware friendly than other options¹. However, even so, the purpose of this article is not to argue that 2×2 is the best option. Instead, the question we ask is that *given* the problem of using 2×2 , how to determine the optimal one? Having said that, and despite 2×2 patterns are the main subject of interest, we have included experiments on other pattern sizes in Section IV-G.

What other options do we have for exposure control?

The theme of this article is exposure and gains controls. In the literature, there are three mainstream approaches:

- One fixed pattern for all:* Mount a static mask with spatially-varying light transmittance on top of the sensor array. This requires minimal/no additional circuit as compared with traditional CMOS sensors. However, the spatially-varying exposure pattern is then fixed and cannot be changed to adapt to the scene. Most traditional work on spatially-varying exposure image reconstruction explicitly/implicitly use this approach [1], [6], [7], [8], [9], [10], [11], [12].
- (This article) Periodic 2×2 , updated on-the-fly:* Use different signal lines to control the transfer timing (and hence control the exposure time) of different pixels. This approach allows the spatially-varying exposure pattern to be set on-the-fly and adapt to the scene. However, the number of signal lines increases linearly with the number of independent controls, and so the number of signal lines needs to be small. Overall, it is more functionally versatile than mounted static mask at the cost of some additional circuitry.

¹An analogy worth mentioning is the color filter arrays: While we all agree that the 2×2 Bayer pattern is sub-optimal, today we only see a handful of non-Bayer color filters in camera products. Even the latest 4-cell Quad-Bayer patterns by Sony, Samsung and OmniVision are just variants of Bayer.

TABLE II
COMPARISON OF PROBLEM SETTINGS

	Yasuma et al. [10]	Ours
Pattern	fixed	adjustable on-the-fly
Optimization	optimal for the average over all lighting scenarios	optimal for every single scene
Objective	obtain a universal pattern	select best for current scene

C) *Per-pixel or per-block, updated on-the-fly*: Two types of hardware realizations are available, (1) add in-pixel latches/flip-flops/logic circuit (also called digital pixel sensor as compared with traditional active pixel sensor), (2) use digital micromirror device image sensor (DMDIS) where the exposure of each pixel can be independently controlled. The downside of this approach is that, for realization (1), the required additional circuitry is gigantic and that it results in a low fill factor (as low as 6% [13]) as well as higher circuit noise, and for realization (2), the resolution, size, and form factor are limited and it has only been demonstrated in lab environment so far. Most of the focal plane coded exposure work adopt this approach [13], [14], [15], [16], [17], [18], [19], [20], [21].

B. Related Work

Our work stands at the intersection of spatially-varying-exposure (SVE) imaging and HDR imaging exposure control. Some existing works are worth noting.

SVE Imaging: After Nayar and Mitsunaga [1], a number of methods have been proposed to: improve image reconstruction quality [6], improve reconstruction speed as well as robustness against non-uniform noise strength [7], [22], adapt the application of SVE HDR imaging from a single image to videos [8], leverage the higher single frame dynamic range advantage of SVE imaging to tackle motion registration problem of HDR video capturing [9], extend the original concept of SVE to the idea of generalized assorted pixel [10], in which image resolution, dynamic range, and spectral profile can be balanced post-capture by imaging with an optimized complex SVE mask. Among these SVE imaging works, [10] is closest to our problem and we emphasize the difference between their problem setting and ours in Table II.

HDR Exposure Control: Exposure bracketing is a very popular HDR imaging technique, which fuses multiple LDR images exposed at different levels to form one HDR image [23], [24], [25], [26]. Despite its popularity, few studies focus on selecting proper exposure levels for the LDR frames. These exposure control algorithms devised for exposure bracketing mostly fall into two groups: (1) algorithms focusing on design simplicity and efficiency so as to be deployed on imaging devices [27], [28], [29], [30], and (2) image formation modeling based algorithms focusing on optimality, i.e., to find the optimal set of exposures by some metric for reconstructing the scene [31], [32]. A recent work also demonstrates the possibility of using reinforcement learning to train an exposure bracketing selection network [33]. In terms of design, our SVE pattern selection algorithm is an image formation modeling based algorithm; however, methods

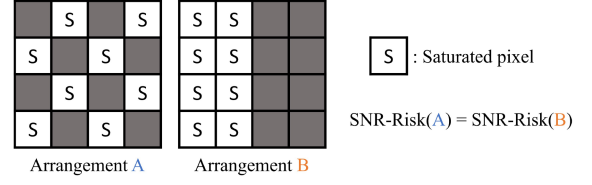


Fig. 2. Two arrangements of saturated pixels share the same SNR-Risk; however, the image captured with arrangement A can be recovered by interpolation while it is more challenging to recover that with arrangement B.

like [31], [32] cannot be migrated to our problem without significant changes, in that spatial multiplexing and pixel interpolation are not within their problem scopes. Imaging with other types of image sensors, e.g., Quanta Image Sensors (QIS) [34], [35], [36], are also candidate solutions to HDR imaging. It is further suggested by [37], [38] that low bit-depth sensors provide wider dynamic range. In this work, we use a general sensor model (1).

III. SVE PATTERN SELECTION FOR HDR

In this section we present the core idea of this article, which is the concept of SVE-Risk and efficient methods to evaluate the SVE-risk.

A. Limitations of SNR

To motivate the definition of the SVE-Risk, we first discuss the limitations of the per-pixel output-referred² signal-to-noise ratio (SNR). In the context of our problem and image formation model, the SNR at pixel i is

$$\text{SNR}_i = \begin{cases} \frac{(\alpha_i \tau_i) \theta_i}{\sqrt{\alpha_i^2 (\tau_i \theta_i + \mu_{\text{dark}}) + \sigma_{\text{read}}^2}}, & (\alpha_i \tau_i) \theta_i \leq V_{\text{max}}, \\ 0, & (\alpha_i \tau_i) \theta_i > V_{\text{max}}, \end{cases} \quad (2)$$

where V_{max} is the maximum voltage allowed by the ADC.

Intuitively, the SNR highlights two aspects of the exposure/gain: (i) If the pixel is not saturated, then the SNR will increase with α_i and τ_i . (ii) If the pixel is saturated (so the signal received by ADC exceeds V_{max}), then the SNR is zero. The two cases are consistent with the classical model in [1].

With the per-pixel SNR_i defined, it is straightforward to define the risk of adopting a particular (α, τ) for capturing a given scene using average SNR of all pixels:

$$\text{SNR-Risk}(\alpha, \tau) = \left(\frac{1}{N} \sum_{i=1}^N \text{SNR}_i \right)^{-1}, \quad (3)$$

where the reciprocal is used to convert the SNR to a risk.

SNR-Risk has two major drawbacks:

- i) SNR-Risk is agnostic to how saturated pixels distribute in the image. Consider the example shown in Fig. 2. While the two patterns will give exactly the same SNR, only pattern A is recoverable because neighboring pixels are

²In this article we are interested in full-well capacities that are sufficiently large. For pixels with extremely small full-well capacity, e.g., Quanta Image Sensors (QIS), one needs to use the more general formula known as the exposure-referred SNR. We refer to the article by Chan [39] for details.

available. Pattern B contains a large region of saturated pixels, which is very difficult to recover.

- ii) SNR is a pixelwise calculation without considering its neighbors. Therefore, if there is a bright region, all four control elements of the 2×2 pattern will try to match the scene without coordinating among themselves. The consequence is that SNR-Risk will choose an all bright or all dark pattern.

The SNR-Risk computes the risk for a pixel depending solely on one pixel. This is in contrast to the fact that practical image reconstruction algorithms almost always aggregate spatial information before predicting a pixel's output value. Hence, SNR-Risk tends to over-estimate the risks associated with easily recoverable pixels. To overcome this issue, we want to design a risk with the neighborhood relations of pixels taken into consideration. To summarize, SNR-Risk is not a good metric because of the following.

SNR-Risk cannot comprehend the local structure of the exposure, hence it cannot tell us whether the acquired image is **recoverable**.

B. SVE-Risk

After explaining why SNR is not a good metric to assess the multiplex patterns, in this subsection we introduce a new concept called spatially varying exposure risk (SVE-Risk).

We first consider the definition of an ideal risk. Let $\mathbf{y}$ be the sensor readout and estimator $\hat{\boldsymbol{\theta}} = [\hat{\theta}_1, \dots, \hat{\theta}_N]$ be the reconstruction mapping that produces an estimate $\hat{\boldsymbol{\theta}}(\mathbf{y}, \boldsymbol{\alpha}, \boldsymbol{\tau})$ of $\boldsymbol{\theta} = [\theta_1, \dots, \theta_N]$. The ideal risk is

$$\text{Risk}(\boldsymbol{\alpha}, \boldsymbol{\tau}) = \inf_{\hat{\boldsymbol{\theta}}} \mathbb{E}_{\mathbf{y}} \left[\frac{1}{N} \sum_{i=1}^N \left(\frac{\hat{\theta}_i(\mathbf{y}, \boldsymbol{\alpha}, \boldsymbol{\tau}) - \theta_i}{\theta_i} \right)^2 \right], \quad (4)$$

where θ_i and $\hat{\theta}_i(\cdot)$ denote the i -th element of the ground truth radiance $\boldsymbol{\theta}$ and the estimate $\hat{\boldsymbol{\theta}}(\cdot)$, respectively. Note that if the estimator $\hat{\boldsymbol{\theta}}$ is a pixel-wise maximum likelihood (ML) estimator using the forward model defined in (1), then the squared ratio in (4) is exactly the inverse of pixel-wise output-referred SNR for non-overflow pixels. From this perspective, the SNR-Risk can be deemed as a special case of the ideal risk, in which the estimator is predetermined to be a pixel-wise ML estimator.

The caveat of (4) is that an oracle estimator $\hat{\boldsymbol{\theta}}$ for a scene is never known and it is also impossible to obtain the infimum by enumerating all possible reconstruction algorithms. To mitigate this issue, we approximate the risk by using a hypothetical ideal *local* estimator. This hypothetical estimator cannot be constructed in practice (any reconstruction algorithm is likely to be worse than the hypothetical estimator), but it can give us a meaningful approximation to the infimum.

Definition 1 (Local estimator): A **local estimator** $\hat{\theta}_i$ at the pixel i is a function that maps the neighborhood observations $\mathbf{y}_i = \{y_j \mid j \in \mathcal{B}_i\}$ to an estimate $\hat{\theta}_i(\mathbf{y}_i)$, where $\mathcal{B}_i$ denotes the neighborhood around pixel i . If y_i is saturated, $\hat{\theta}_i$ uses the neighborhood *without* y_i , i.e., $\mathbf{y}_{-i} = \{y_j \mid j \in \mathcal{B}_i\} \setminus \{y_i\}$.

As we define this hypothetical local estimator, we assume that it has the perfect knowledge about inter-pixel correlations of its neighborhood. Therefore, it allows us to achieve two things:

- When a pixel is not saturated, the estimator will return us the same value as the SNR-Risk.
- When a pixel is saturated, the estimator will make an interpolation. The interpolated pixel will have a risk no higher than the largest risk within the neighborhood.

Based on these properties, we can define the SVE-Risk by considering three situations:

The **SVE-Risk of the i -th pixel** is defined as

$$\begin{aligned} \text{SVE-Risk}_i(\boldsymbol{\alpha}, \boldsymbol{\tau}) & \quad (5) \\ &= \begin{cases} \frac{1}{\theta_i^2 |\mathcal{B}_i^*|} \left[\frac{(\theta_i + \mu_{\text{dark}})}{\tau_i} + \frac{\sigma_{\text{read}}^2}{\alpha_i^2 \tau_i^2} \right], & \text{unsaturated,} \\ \max_{j \in \mathcal{B}_i^* \setminus \{i\}} \frac{1}{\theta_j^2 |\mathcal{B}_j^*|} \left[\frac{(\theta_j + \mu_{\text{dark}})}{\tau_j} + \frac{\sigma_{\text{read}}^2}{\alpha_j^2 \tau_j^2} \right], & \text{sat., } \checkmark \mathcal{B}_i, \\ \left(\frac{V_{\text{max}}}{\alpha_i \tau_i} - \theta_i \right)^2, & \text{sat., } \times \mathcal{B}_i, \end{cases} \end{aligned}$$

where $\mathcal{B}_i^* = \mathcal{B}_i \setminus \mathcal{S}$ is the neighborhood around pixel i minus the set of saturated pixels $\mathcal{S}$.

Let's elaborate on the three cases in the definition:

- *Unsaturated*: If a pixel is unsaturated, the risk is defined as the variance of the measurement (which is the denominator of (2), squared and normalized). The normalization is necessary for preventing bright regions in the image from dominating darker region risks and is crucial for HDR imaging. For unsaturated pixels, the risk calculated in (5) is essentially SNR-Risk value scaled by the number of observations in the neighborhood.
- *Saturated, $\checkmark \mathcal{B}_i$* : The neighborhood contains pixels that can be used for interpolation. In this case, the risk is defined as the worst variance in the neighborhood. The intuition is that since SVE-Risk uses neighboring pixels to determine the risk, it can be deemed as an extension of the SNR-Risk where we combine independent exposures.
- *Saturated, $\times \mathcal{B}_i$* : The neighborhood does not contain any useful pixels. The risk is defined as the squared error between the cutoff $\frac{V_{\text{max}}}{\alpha_i \tau_i}$ and expected radiance θ_i . Note that this risk is worse than the second case because, in the second case, the substitution comes from one of the unsaturated pixels. For the third case, the difference between $\frac{V_{\text{max}}}{\alpha_i \tau_i}$ and θ_i can be very large if θ_i is far from the cutoff.

With pixel-wise SVE-Risk defined, we define overall SVE-Risk by taking sum of each pixel's risk.

The **SVE-Risk of the whole image** is defined as

$$\text{SVE-Risk}(\boldsymbol{\alpha}, \boldsymbol{\tau}) = \sum_{i=1}^N \text{SVE-Risk}_i(\boldsymbol{\alpha}, \boldsymbol{\tau}). \quad (6)$$

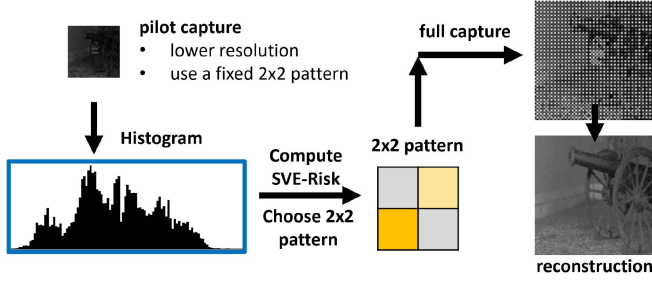


Fig. 3. Our imaging pipeline is for *dynamical* control of the exposure. We capture a low-resolution pilot image using a fixed 2×2 exposure pattern (typically high-medium-medium-low) to construct a histogram. From the histogram we compute the risk, and guide by the risk we capture the full image. Since the SVE-Risk is calculated based on the histogram instead of the image, its computational cost is very low.

C. Computing SVE-Risk

After defining the SVE-Risk, the next big question is how to compute this seemingly “uncivilized” (5). However, before we explain how we calculate the SVE-Risk, we first explain the overall imaging pipeline outlined in Fig. 3.

Since our imaging goal is to dynamically control the exposure, our decision on the 2×2 pattern needs to be fast. As illustrated in Fig. 3, the way we compute the SVE-Risk is based on the histogram of the scene radiance. This histogram does not need to be perfectly precise. Therefore, we can use a lower resolution of the same scene, e.g., a $4 \times$ down-sampled capture for a 1 megapixel sensor at a natural scene, and we can use any average 2×2 pattern such as high-medium-medium-low. The purpose of the pilot capture is to construct a histogram so that we can make decisions for choosing the 2×2 pattern, and the size of it depends on the spatial frequency of image contents and the spatial resolution of the imaging devices. An alternative way to think of this pilot capture is, in video capturing, using the previous frame to guide the pattern selection for the next frame. This saves dedicated captures and potentially allows spatially-varying pattern-based HDR video on the fly at a cost of some SVE-Risk computation and hardware control overhead.

We now discuss the histogram. Assume we have collected a pilot image. We denote the distribution of the radiance as $p(\theta)$. An example is shown in Fig. 4. We stress again that operating on the radiance distribution rather than individual pixels is more efficient, because, once the histogram is built, the complexity of pattern risk estimation becomes proportional to the number of bins in the histogram (order of magnitude is hundreds to thousands) instead of the number of pixels (order of magnitude is hundreds of thousands to millions).

Given a pattern $\{(\tau_\ell, \alpha_\ell) \mid \ell = 0, \dots, 3\}$, without loss of generality, we assume that the exposures/gains are sorted so that $\alpha_0\tau_0 \leq \alpha_1\tau_1 \leq \alpha_2\tau_2 \leq \alpha_3\tau_3$. The pattern gives us four cutoff radiance levels $\frac{V_{\max}}{\alpha_\ell\tau_\ell}$ which partition the entire radiance range into five intervals, as shown in Fig. 4.

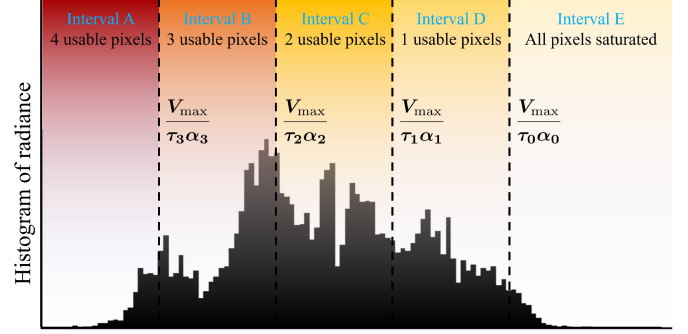


Fig. 4. An example radiance histogram. The overall radiance histogram is partitioned by the 4 exposures/gains into 5 intervals. For the unsaturated pixels in intervals A to D, the risk associated with them corresponds to the unsaturated case in (5). Pixels in the interval above the threshold $V_{\max}/\tau_0\alpha_0$ (interval E) are all saturated and unlikely to be recovered. Their risks are computed as the empirical mean squared error as the saturated, $\times B_i$ case in (5). Risks of saturated pixels in intervals below $V_{\max}/\tau_0\alpha_0$ (intervals A to D) can be substituted with the worst risks of their neighbors, which is the saturated, $\checkmark B_i$ case in (5).

As shown in Fig. 4, a radiance level θ saturates an exposure/gain element (τ_ℓ, α_ℓ) if it is greater than the corresponding cutoff $\frac{V_{\max}}{\alpha_\ell\tau_\ell}$. For $\theta > \frac{V_{\max}}{\alpha_0\tau_0}$ (interval E), they saturate all exposure/gain elements and it is very likely that the pixels corresponding to these radiance levels are not recoverable by any reconstruction algorithm. In this case, risk associated with these radiance levels is, according to Case 3 in (5):

$$\overline{\text{SVE-Risk}}_{\text{nonrecoverable}}(\alpha, \tau, \theta) = \left(\frac{V_{\max}}{\alpha_0\tau_0} - \theta \right)^2. \quad (7)$$

Now consider a radiance level between two cutoffs $\frac{V_{\max}}{\alpha_\ell\tau_\ell} < \theta < \frac{V_{\max}}{\alpha_{\ell-1}\tau_{\ell-1}}$ (i.e., a radiance level in intervals B, C, D). The radiance level θ saturates all exposure/gain elements below it; however, since there are neighboring elements not saturated by θ , it is likely that saturated pixels at this radiance level can be recovered by exploring their neighbor pixels. The risk associated with this radiance level is then, according to Case 1 and Case 2 in (5):

$$\overline{\text{SVE-Risk}}_{\text{recoverable}}(\alpha, \tau, \theta) = \frac{1}{4} \left[\underbrace{\sum_{j=0}^{\ell-1} \frac{1}{B_j\theta^2} \left(\frac{(\theta + \mu_{\text{dark}})}{\tau_j} + \frac{\sigma_{\text{read}}^2}{\alpha_j^2\tau_j^2} \right)}_{\text{contribution from non-saturated elements}} + \underbrace{\sum_{j=\ell}^3 \frac{1}{B_0\theta^2} \left(\frac{(\theta + \mu_{\text{dark}})}{\tau_0} + \frac{\sigma_{\text{read}}^2}{\alpha_0^2\tau_0^2} \right)}_{\text{saturated, use risk of worst non-saturated neighbor}} \right], \quad (8)$$

where B_j is the number of non-saturated pixels within the neighborhood (with predetermined size) of exposure/gain element (α_ℓ, τ_ℓ) . Note that, since we know how pattern is tiled across the entire sensor array, B_ℓ can be calculated for each saturation

TABLE III
COMPARISON OF RISK ESTIMATOR RUN TIME FOR SENSORS OF DIFFERENT RESOLUTIONS

Resolution	Runtime for calculating SVE-Risk	Runtime for calculating SNR-Risk
512x896	0.20±0.01 sec	0.69±0.01 sec
1024x1792	0.27±0.01 sec	6.13±0.22 sec
2048x3584	0.52±0.01 sec	27.35±0.12 sec

Numbers shown are run time of evaluating the corresponding risk of 495 patterns averaged over 100 executions.

scenario (i.e., 0, ..., 3 exposure/gain elements saturated) before-hand and be stored in memory (see supplementary).

The risk is therefore the sum of (7) and (8).

To **numerically compute the SVE-Risk**, we construct the radiance histogram (like Fig. 4), and calculate

$$\begin{aligned} \text{SVE-Risk}(\alpha, \tau) &= \int_{\frac{V_{\max}}{\alpha_0 \tau_0}}^{\infty} \overline{\text{SVE-Risk}}_{\text{nonrecoverable}}(\alpha, \tau, \theta) p(\theta) d\theta \\ &\quad + \int_0^{\frac{V_{\max}}{\alpha_0 \tau_0}} \overline{\text{SVE-Risk}}_{\text{recoverable}}(\alpha, \tau, \theta) p(\theta) d\theta, \end{aligned}$$

where $p(\theta)$ is the radiance histogram of the pilot capture.

We show a comparison of the run time of calculating the SVE-Risk and the SNR-Risk in Table III. Since SVE-Risk is calculated using the histogram, it is insensitive to the image resolution. In contrast, since SNR needs to evaluate every single pixel, its computation grows with the number of pixels.

D. Efficient Pattern Enumeration

There is one final design question we need to answer before using the SVE-Risk. It is the problem of candidate patterns to evaluate. Suppose we have 9 exposure levels for a 2×2 pattern, we will have a total of $9^4 = 6561$ candidates. It will be too much computation if we need to calculate the SVE-Risk for each candidate pattern. Therefore, in this subsection, we present a method to eliminate low priority patterns.

To remove the low-priority patterns, we made an observation that the majority of all candidate patterns are redundant. For example, from the reconstruction point-of-view, an exposure pattern $\tau = \{1, 10, 10, 10\}$ is almost identical to patterns $\{10, 1, 10, 10\}$, $\{10, 10, 1, 10\}$, and $\{10, 10, 10, 1\}$. To justify this claim, we experiment with three representative reconstruction algorithms on sensor readouts synthesized with these four patterns. We test on a dataset containing 46 images and show the results in Table IV. Across the four exposure patterns, the reconstruction results are identical for any fixed algorithm.

Based on the observations above, we define the concept of pattern equivalence: *two patterns are equivalent for capturing a scene if they are permutations of each other*. Our proposed strategy is to enumerate on **pattern equivalence classes** and compute canonical form of each class. The canonical pattern is the one

TABLE IV
 μ PSNR (dB) OF THE RECONSTRUCTED IMAGES FOR THE FOUR EQUIVALENT MULTIPLEX PATTERNS

	1	10	10	1	10	10	10	10
	10	10	10	10	1	10	10	1
ADMM-TV [40]	31.1dB	31.1dB	31.1dB	31.1dB	31.1dB	31.1dB	31.1dB	31.1dB
Restormer [41]	30.0dB	30.0dB	30.0dB	30.0dB	30.0dB	30.0dB	30.0dB	30.0dB
LPA [7]	29.7dB	29.7dB	29.7dB	29.7dB	29.7dB	29.7dB	29.7dB	29.7dB

The PSNR is averaged across 46 test images and rounded to first decimal place.

with maximum variation in the 2×2 grid. Specifically, if the input pattern is $\tau = [\tau_1, \tau_2, \tau_3, \tau_4]$, we sort the sequence to obtain $\text{sort}(\tau) = \{\tau_{[1]}, \tau_{[2]}, \tau_{[3]}, \tau_{[4]}\}$ where $\tau_{[1]} \leq \tau_{[2]} \leq \tau_{[3]} \leq \tau_{[4]}$. The canonical pattern is then defined as

$$\text{canon}(\tau) = \begin{bmatrix} \tau_{[1]} & \tau_{[3]} \\ \tau_{[4]} & \tau_{[2]} \end{bmatrix}.$$

The intuition here is that since $\text{sort}(\tau)$ is already sorted, the alternating allocation of the exposure values will maximize the variation within the 2×2 grid. For example, the canonical forms of the patterns $\tau = [8, 32, 1, 10]$ and $\tau = [1, 10, 1, 10]$ are

$$\text{canon}(\tau) = \begin{bmatrix} 1 & 10 \\ 32 & 8 \end{bmatrix} \quad \text{and} \quad \text{canon}(\tau) = \begin{bmatrix} 1 & 10 \\ 10 & 1 \end{bmatrix},$$

respectively. Once all the patterns are converted to their canonical forms, checking the equivalence is simplified to check whether the two canonical forms are identical.

By enumerating on equivalence class instead of all realizable patterns, we reduce the complexity to $\sum_{k=1}^{\min(L,m)} \binom{L}{k} \binom{m-1}{k-1}$ while maintain same coverage of realizable pattern space.³ When $L = 9$ and $m = 4$, this reduces enumeration size from 6561 to 495.

Remark: Readers may ask: The proposed algorithm is largely an exhaustive search and it requires knowledge about the radiance distribution. Is it possible to improve the search? We note that the pilot estimate is designed to be *coarse*. As long as the shape of the radiance distribution is obtained, we can perform the histogram-based calculation. For faster algorithms, we do not think the typical gradient-based algorithm would work here because our problem is discrete with many stationary points. There might be some discrete optimization methods. We are open to explore them in our future work.

IV. OPTIMALITY AND UNIVERSALITY

In the beginning of the paper, we mentioned two key findings of this article. Firstly, we claim that for exposure multiplexing, there exists a better *optimality* criteria than the SNR. We have elaborated on the SVE-Risk in the previous section. In this section, we evaluate SVE-Risk by justifying the following statement.

³See supplementary for the derivation and pseudo-code.

Optimality: The SVE-Risk optimal exposure multiplex pattern can generate a raw image, if passed through a reconstruction algorithm, with nearly the highest PSNR.

The second claim we made is *universality*. In this section, we justify the following statement.

Universality: The optimal exposure multiplex pattern is universally good for *all* image reconstruction algorithms.

Because of the empirical nature of both statements, we answer them through experiments. Our experiments involve large-scale datasets, several new ways of visualizing the results, and a collection of real data.

A. Datasets

Before diving into experiment design and results, we describe datasets and synthesis parameters in more details. We use linearized ground truth 16-bit HDR images from NTIRE [42], HDR-Eye [43], and SIGGRAPH17 [44] datasets as scene radiance maps. These radiance maps are normalized such that, under minimal exposure and unit gain, the photon flux corresponding to 99 percentile is within the ADC range. This setting is practical and realizable on hardware using modern auto-exposure control. We do not attempt to accommodate for the brightest 1% of pixels, because these pixels usually are light sources that directly shine on the sensor. Each normalized radiance map is resized such that the short edge has a length of 512 pixels. The radiance maps are then used to synthesize raw sensor readout using image formation model in Section II-A. We use HDR-Eye data for reconstruction algorithms hyper-parameter tuning and model training, SIGGRAPH17 data for hyper-parameter validation, and NTIRE data for testing and empirical study.

B. Reconstruction Algorithms

An important task of our verification is to evaluate the quality of the *reconstructed* image. Thus, it is necessary to consider image reconstruction algorithms. However, given the sheer volume of reconstruction methods, it would be impossible to evaluate everyone. A surrogate we take here is to consider three representative classes of algorithms. Within each class, we consider a representative method which we can either implement or we have access to the original source code.

- i) *Non-data-driven non-iterative:* Algorithms in this class do not require any training and usually make relatively simple or even no assumptions about image structure. During reconstruction, the estimation for each pixel is only carried out once (hence non-iterative). Traditional bi-linear/quadratic/cubic interpolation, median filters, filter banks and more fall into this class. We adopt and implement a local polynomial approximation (LPA) [7] as a representative of this class.
- ii) *Non-data-driven iterative:* These are classical tools for solving an inverse problem. Compared with non-iterative approaches, algorithms in this class usually model both

the forward imaging process and underlying image structure/prior. This class of algorithms alternate between a forward step to handle the data fidelity, and a backward step that integrates the scene prior. Typical examples include Plug-and-Play [40] and Piecewise Linear Estimators [45], etc. We implement a Plug-and-Play ADMM with total variation prior (ADMM-TV) to represent this category.

- iii) *Data-driven:* Dictionary learning [46] and neural networks [41], [47], [48] based image reconstruction/restoration methods fall into this category. It should be noted that although some non-data-driven iterative approaches may also adopt a dictionary or network as a sub-component (e.g., one may use a denoiser network in Plug-and-Play framework as prior step), and there has been numerous efforts [49], [50] to try to bring together the best of both tools, we do not consider them as purely data-driven approaches. We limit the scope of this category to methods that are one-pass (i.e., non-iterative) and trained directly for the inference task. We use Restormer⁴ [41] as an example of this category.

C. Metrics

Because of the unique problem setting we have, there is no prior standardized evaluation criterion. To this end, we consider a few known metrics and introduce a few new ones.

- i) μ PSNR, μ SSIM, μ LPIPS: In high dynamic range images, high exposure regions can easily dominate losses or metrics over low exposure regions; therefore, evaluating reconstruction quality in linear scale is usually less meaningful. Similar to other HDR related works [33], we evaluate reconstruction quality on μ -tone-mapped images, which is defined as

$$\mathbf{x}_\mu = \frac{\log(1 + \mathbf{x} \cdot \mu)}{\log(1 + \mu)},$$

where x is a linear scale image normalized to $[0, 1]$ and μ is a hand-picked hyper-parameter controlling the strength of dynamic range compression. In our experiment, we set μ to the maximum reference level of ADC (see Section II-A).

In tone-mapped space, we measure the PSNR (μ PSNR, higher is better), structural similarity (μ SSIM, higher is better) [51], and perceptual distance (μ LPIPS, lower is better) [52] between reconstructions and ground truth images.

- ii) **SNR-Risk, SVE-Risk, and their variants:** Since one main objective of this article is to propose SVE-Risk, it is necessary to compare it with SNR. In addition to the

⁴We discovered in our experiment that networks cannot be trained well when the input to a network has a very high dynamic range, and this training failure cannot be saved by input normalization. Therefore, instead of operating on raw sensor readout in linear scale and predict a linear/log scale output (as most of network-based HDR works do. Their inputs are usually LDR images and their task is to combine LDR images into HDR images, so the domain of their problems does not align exactly with ours), we take a log scale normalized sensor readout as input and predict a log scale radiance map.

standard SNR-Risk and SVE-Risk described in previous sections, we also evaluate following two variants of the risks, which arise naturally as one contemplates why one risk works while the other does not. Thus, we have four risk terms to consider:

- SNR-Risk, as defined in (3).
- SVE-Risk, as defined in (6).
- SNR_{MSE}: SVE-Risk assigns coarse estimates of mean squared error (MSE) as risk to unrecoverable pixels. Given the close relationship between MSE and PSNR/SSIM, one may wonder if this assignment gives SVE-Risk an unfair advantage over SNR-Risk, as SNR-Risk is purely forward model based. To answer this, we modify SNR-Risk by assigning the actual MSE between the normalized sensor readout (i.e., $y/(\alpha\tau)$) and the corresponding ground truth to overflowing pixels. We denote this variant as SNR_{MSE} in results.
- SVE_{w/o}: The idea of our SVE-Risk design is that it penalizes neighborhoods with too many saturated pixels through an auto-tuned parameter $|\mathcal{B}|$. We evaluate SVE-Risk without the penalty term $|\mathcal{B}|$, denoted SVE_{w/o} in results, to show that this idea is indeed imperative for achieving optimal performance instead of being a dubious add-on.

D. Verify the Optimality of SVE-Risk

In this subsection we discuss our experiments to assess the optimality of the SVE-Risk. We first discuss the protocol of the experiment, and then the results.

Protocol of experiment: We would like to compare SNR-Risk and SVE-Risk. The evaluation of SNR-Risk requires access to the radiance map. For convenience, we directly use ground truth radiance map as the input. The evaluation of SVE-Risk is easier because we only need the histogram. We assume we have access to 4 exposures evenly distributed across the total exposure levels. The histogram is then built using all non-saturating pixels.

Given a set of candidate patterns (495 patterns in our experiment), we define the **oracle pattern** as the one that gives the highest PSNR. We do not have access to this oracle pattern. We want to use a risk to estimate the best pattern and to rank all patterns. Note that the ranked top-1 pattern by a risk is usually *not* the oracle pattern.

To gauge the ranking power of a risk, it is not informative to compare the rank of the oracle pattern rated by different estimators, because the ranks do not reflect reconstruction quality differences. It is also not enough to only look at the reconstruction quality difference between the top-1 pattern and the oracle pattern for two reasons. Firstly, a small quality difference can be a coincidence due to specific textures or scenes being insensitive to choices of pattern. Secondly, even a risk that ranks patterns poorly may find an acceptable pattern once in a while (as illustrated in Fig. 6(d)).

Therefore, in our evaluation protocol, we propose to measure two descriptive statistics of risk estimators:

- Average quality difference* between using the oracle and the top- K patterns ranked by an estimator. This statistic evaluates the absolute reconstruction quality drop when



Optimal patterns (oracle scheme)			
Image	How bright?	Optimal exposure τ	Optimal gain α
Scene 1	All dark	{10,1,10,10}	{4,2,4,4}
Scene 2	All bright	{1,1,1,1}	{2,1,2,2}
Scene 3	Half-half	{1,10,10,1}	{1,2,4,2}
Scene 4	More dark	{1,10,10,10}	{1,4,4,1}

Fig. 5. Scenes and their corresponding oracle patterns.

one uses top patterns selected by a risk estimator compared to using the oracle pattern. Furthermore, if top- K average difference is an increasing function of K , then the risk estimator likely has good ranking power on patterns. Mathematically, we define

$$\Delta_K = \frac{1}{NK} \sum_{n=0}^{N-1} \sum_{i=0}^{K-1} \left(\underbrace{s_n^*}_{\text{oracle score}} - \underbrace{s_{n,i}}_{i\text{-th score}} \right), \quad (9)$$

where Δ_K is the top- K average difference, s_n^* is the score of the oracle pattern on the n -th radiance map ($n = 0, 1, \dots, N-1$, with $N = 1494$ in this article), $s_{n,i}$ is the score of the i -th top pattern as ranked by the risk.

- Probability* that the reconstruction quality difference between the oracle pattern and the top-1 pattern is above certain pre-determined threshold. This statistic measures: given a threshold that one considers as critical, what is the probability that using a particular risk estimator will not yield satisfactory results. Formally, we define

$$Q(\eta) = \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{I} \left\{ \frac{s_n^* - s_{n,1}}{s_n^*} > \eta \right\}, \quad (10)$$

where Q is the probability of having a difference above a threshold η , $\mathbb{I}[\cdot]$ is an indicator function.

Results: We show in Table V the top- K average differences and Q scores at two thresholds across the pattern ranking dataset. The full SVE-Risk is capable of selecting a pattern that will yield a reconstruction with close to oracle performance, with an average μ PSNR drop for the top-1 pattern around 1 dB. SVE-Risk without the neighborhood penalty term can still pick a reasonably good pattern, but is almost always subpar compared to pattern selected by full SVE-Risk. SNR-Risk as well as its variant SNR_{MSE} are incapable of picking a good pattern in almost all scenarios, and equipping SNR-Risk with MSE for overflowing pixels does not help SNR-Risk. This experiment shows the significance of properly assigning surrogate risk to recoverable overflowing pixels and exposure/gain control element binding.

How do the Optimal Patterns Look Like? To give readers an idea of how the optimal exposure/gain patterns look like, we show in Fig. 5 four randomly selected scenes and their

TABLE V
TOP-K AVERAGE DIFFERENCES AND Q SCORES AT THRESHOLDS

	μPSNR															
	Δ_1				Δ_5				$Q(1\%)$				$Q(5\%)$			
	SVE	$\text{SVE}_{w/o}$	SNR	SNR_{mse}	SVE	$\text{SVE}_{w/o}$	SNR	SNR_{mse}	SVE	$\text{SVE}_{w/o}$	SNR	SNR_{mse}	SVE	$\text{SVE}_{w/o}$	SNR	SNR_{mse}
LPA	0.66	1.03	10.32	9.66	1.04	0.98	8.5	8.81	53.8	71.4	99.9	98.4	10	18.6	99.9	96.5
ADMM-TV	1.3	1.7	8.43	7.99	1.43	1.8	7.44	7.38	78.9	81.9	99.9	98.4	26.5	34.5	99.5	92.5
Restormer	0.49	1.54	6.19	5.38	0.62	1.31	4.57	4.49	49.3	87.4	99.9	98.7	4.3	29.8	97.7	96
	μSSIM															
	Δ_1				Δ_5				$Q(1\%)$				$Q(5\%)$			
	SVE	$\text{SVE}_{w/o}$	SNR	SNR_{mse}	SVE	$\text{SVE}_{w/o}$	SNR	SNR_{mse}	SVE	$\text{SVE}_{w/o}$	SNR	SNR_{mse}	SVE	$\text{SVE}_{w/o}$	SNR	SNR_{mse}
LPA	0.0048	0.0050	0.1056	0.1255	0.0093	0.0065	0.0887	0.1096	14.3	17.4	95.9	93.6	0.3	0.1	62.2	81.3
ADMM-TV	0.0080	0.0078	0.1099	0.1070	0.0103	0.0098	0.0829	0.0892	27.2	29	84	85.5	1.9	0.7	45.9	58.5
Restormer	0.0063	0.0106	0.0469	0.0573	0.0075	0.0103	0.0391	0.0443	21.6	37.4	69.5	92.2	0	0	22.5	54.9
	μLPIPS															
	Δ_1				Δ_5				$Q(20\%)$				$Q(80\%)$			
	SVE	$\text{SVE}_{w/o}$	SNR	SNR_{mse}	SVE	$\text{SVE}_{w/o}$	SNR	SNR_{mse}	SVE	$\text{SVE}_{w/o}$	SNR	SNR_{mse}	SVE	$\text{SVE}_{w/o}$	SNR	SNR_{mse}
LPA	0.007	0.018	0.179	0.218	0.019	0.020	0.147	0.191	10.4	40.1	99.6	97.1	0.1	2.3	83.5	94.6
ADMM-TV	0.026	0.031	0.169	0.169	0.032	0.035	0.128	0.140	58.4	71.6	87.9	97.5	8.6	13.5	72.5	78
Restormer	0.03	0.046	0.08	0.112	0.035	0.044	0.070	0.093	75.8	92.6	92.2	94.5	6.4	26.2	42.6	60.5

Right eight columns: probability of the reconstruction quality difference between the oracle pattern and the top-1 pattern selected by a risk being greater than pre-determined thresholds. For all categories, smaller numbers correspond to better selections of exposure patterns. The best score in each comparison category is colored in blue. We use different Q-score threshold for μLPIPS because its scale and spread are much wider than those of the other two metrics.

Left eight columns: average reconstruction quality drop (measured by μPSNR , μSSIM , μLPIPS) between using the oracle pattern and the top-1/top-5 pattern(s) for various reconstruction algorithms.

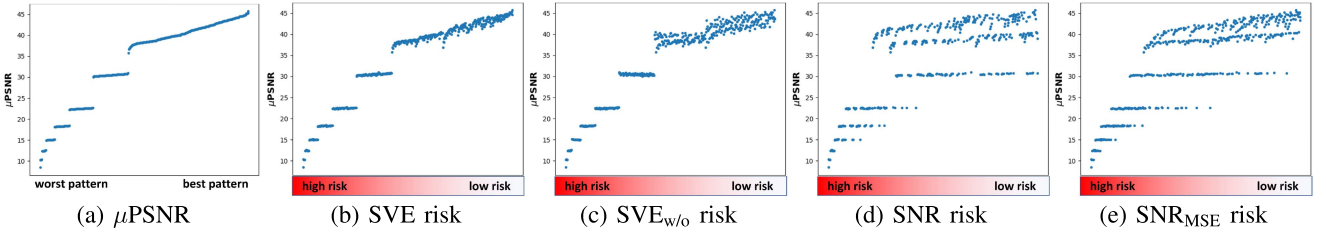


Fig. 6. Scatter plots of μPSNR of exposure patterns with their x -coordinates sorted. Each dot represents a pattern. (a) Exposure patterns sorted by their corresponding μPSNR scores from low to high, representing what an ideal risk estimator should achieve. An ideal risk estimator always assigns lower risks to patterns yielding higher μPSNR . (b)–(e) Exposure patterns sorted by their estimated risk values using various risk estimators. The y -coordinate of a dot is the μPSNR of reconstruction using that pattern.

corresponding optimal exposure/gain patterns. As the radiance of the scenes change from all dark to all bright, the optimal patterns change from all-high to all-low. This variety of scenes with the experimental results suggest that our proposed scheme is able to adaptively select the exposure and gain based on the radiance.

Visualize Ranking Power: To illustrate the ranking power of SNR-Risk and SVE-Risk, we show a scatter plot of μPSNR of a scene sorted by risk values in Fig. 6. We remark that this is a novel visualization of the performance, as we have not seen a similar plot in the literature.

To interpret the results of this plot, we note that the x -axis of the plots is the risk ranked from high to low. There are four risks: SNR, SVE, $\text{SVE}_{w/o}$ and SNR_{mse} . The ideal risk for our task is μPSNR . If we use μPSNR as the metric to rank the patterns, we will have a scatter plot shown in Fig. 6(a). A better pattern ranked by μPSNR will, of course, give a higher μPSNR . When we evaluate other risks, we see that the proposed SVE-Risk has the closest behavior to the ideal risk. In contrast, SNR-based risks show an overlapped behavior. This means that if we use SNR to pick the pattern, we will not be able to tell which pattern is the best because for the same SNR (x -axis), we have multiple patterns on the y -axis.

Staircase μPSNR Behavior: Readers may wonder about the step-wise behavior. This is due to the image histogram, as the scene may contain large flat regions of similar radiance values. As the minimum exposure and gain go above certain thresholds, the brightest large region of scene becomes completely saturated and irrecoverable, causing significant quality drop. Such drop is intrinsic to the scene itself and related to values that local exposure and gain can take, but no pattern is guaranteed to be in any particular cluster as the scene varies (i.e., no intrinsically bad pattern).

Visualize the Patterns: A visualization of adopting the top pattern selected by different risks for capturing is shown in Fig. 7. In Fig. 8, we show an example of reconstructing simulated readouts captured with SVE-Risk top pattern and SNR-Risk top pattern using different reconstruction algorithms.

Can SNR-Risk Work if We Discard “Bad” Patterns? A common question people ask is that would SNR-Risk perform better if we throw away the bad patterns. Our answer is no. Firstly, we simply cannot discard bad patterns when there is no intrinsically bad pattern. Secondly, even if we analyze current scene and discard all patterns that may yield large saturated region, pixel-wise SNR will still not use nonuniform exposure

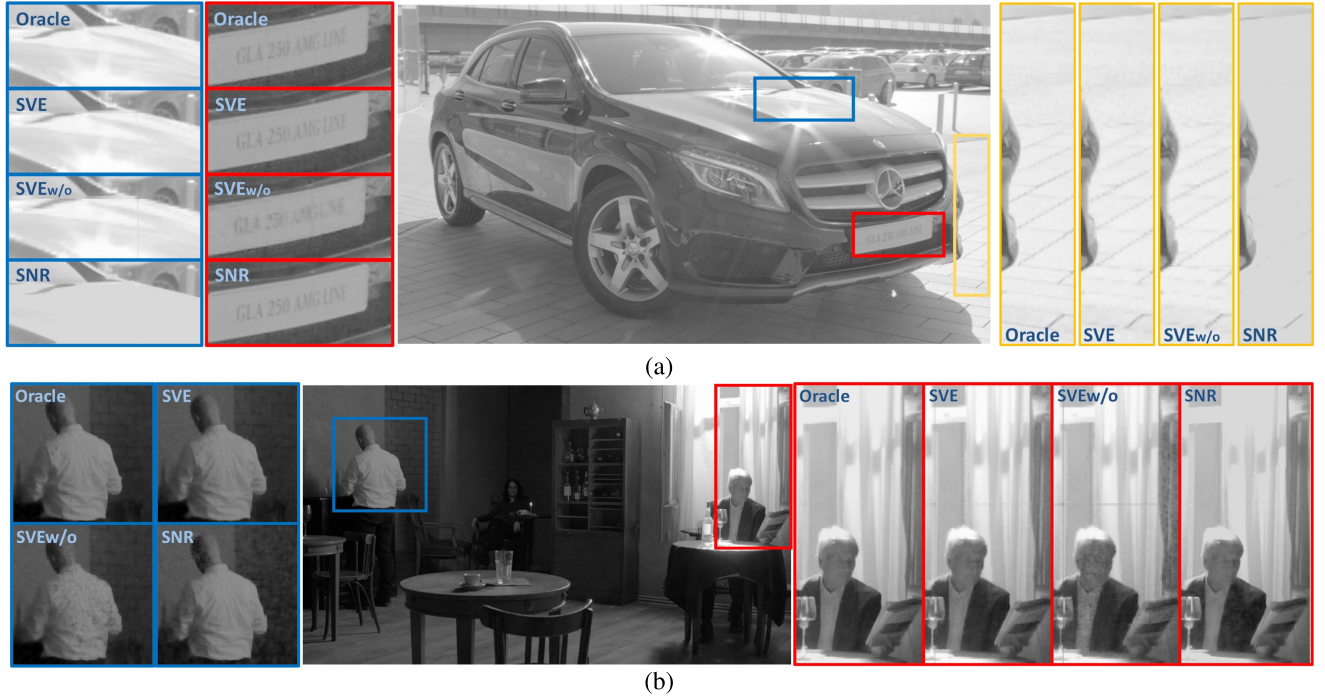


Fig. 7. Qualitative comparisons of risk estimators. Images in the middle are the ground truth. Images on the sides are reconstructions using simulated readouts captured with spatially-varying exposure patterns selected by different risk estimators. Oracle denotes the best possible pattern through exhaustive search. (a) SNR-Risk is dominated by relatively darker regions, so it selects an exposure pattern such that the high-light textures are irrecoverable. SVE-Risk_{w/o} fails to retain low-light details, e.g., characters on the license plate. Oracle and SVE-Risk selections preserve most information across high dynamic ranges. (b) SNR-Risk is unable to pick a spatially-varying exposure/gain, and therefore compromises texture in high-light regions and introduces irrecoverably heavy noise in low-light region. SVE-Risk_{w/o} loses details in medium exposure regions such as the human face and the white shirt. Oracle and SVE-Risk selections trade off high/low flux regions differently: the oracle pattern better preserves the face but loses some details of the curtain, while SVE-Risk's top pattern preserves all curtain details but results in a slightly more blurry face. The reconstruction algorithm is restormer for all images. All images are tone-mapped.

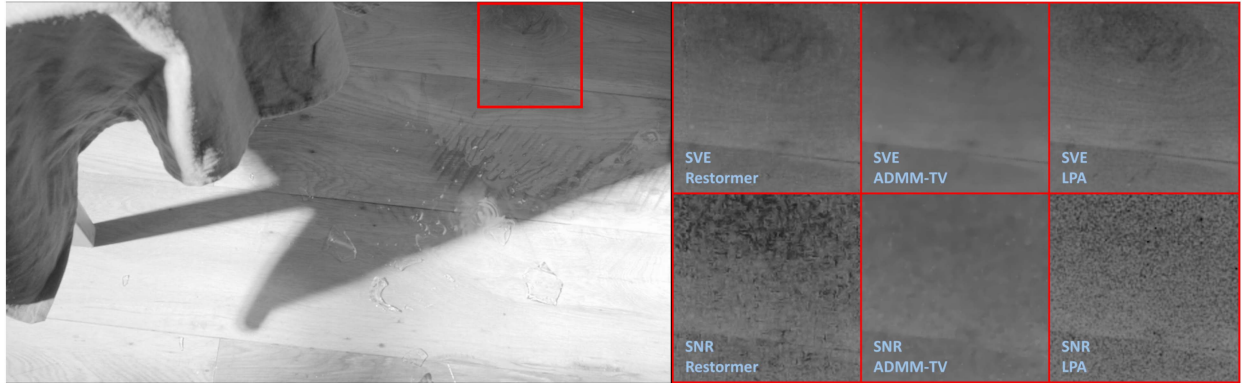


Fig. 8. Images reconstructed by different algorithms. Using the sub-optimal spatially-varying exposure pattern selected by SNR-Risk, all three reconstruction algorithms fail to recover the texture of the hardwood flooring.

levels. This can be seen in the μ PSNR v.s. risk rank scatter plots as shown Fig. 9.

E. Verify the Universality of Patterns

In this subsection, we describe our discovery that an exposure/gain pattern is universal for *many* image reconstruction algorithms. This is a significant departure from the recent trend of camera-algorithm co-optimization where people have been

arguing that jointly optimizing the pattern and the algorithm is essential (e.g., [53] finds better color filter patterns by back-propagating through reconstruction networks). Our experiments in this subsection show the opposite. We find that the design of the exposure/gain pattern can be completely decoupled from the design of the image reconstruction algorithm.

Spearman's ranking correlation: To evaluate the dependency of the pattern and the algorithm, we need some notion of correlation between the two factors. The metric we consider in this

TABLE VI
AVERAGE (STANDARD DEVIATION) AND MEDIAN CORRELATION COEFFICIENTS AND AVERAGE P-VALUE OF EACH CORRELATION COEFFICIENT ACROSS NTIRE DATASET

	μ PSNR			μ SSIM			μ LPIPS		
	average ρ	median ρ	average p-value	average ρ	median ρ	average p-value	average ρ	median ρ	average p-value
ADMM-TV v.s. LPA	0.937 ± 0.035	0.952	$< 10^{-7}$	0.805 ± 0.116	0.830	$< 10^{-7}$	0.740 ± 0.195	0.775	0.008 ± 0.079
LPA v.s. Restormer	0.877 ± 0.070	0.878	$< 10^{-7}$	0.826 ± 0.053	0.826	$< 10^{-7}$	0.773 ± 0.112	0.770	$< 10^{-7}$
Restormer v.s. ADMM-TV	0.891 ± 0.042	0.885	$< 10^{-7}$	0.701 ± 0.130	0.719	$< 10^{-7}$	0.564 ± 0.306	0.625	0.071 ± 0.252

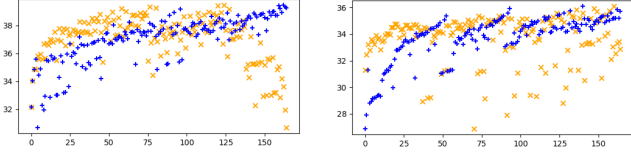


Fig. 9. μ PSNR versus risk rank scatter plots. Even if we throw away bad pixels, SNR-Risk still cannot generate a smooth curve. In contrast, SVE-Risk produces a largely smooth curve. Color code: (Blue: SVE, Orange: SNR).

article is the Spearman's ranking correlation [54], although other types of correlations can also be used.

For each ground truth radiance image θ , we synthesize 495 raw sensor readouts $\{x_0, \dots, x_{494}\}$, one for each pattern equivalence class. For each readout, we use three distinct algorithms $\{f_0, f_1, f_2\}$ to reconstruct the radiance image and evaluate the reconstruction quality using three different metrics $\mathcal{M}_{\mu\text{PSNR}}, \mathcal{M}_{\mu\text{SSIM}}, \mathcal{M}_{\mu\text{LPIPS}}$. We define a score as

$$s_{i,j,k} = \mathcal{M}_k(f_j(x_i), \theta), \quad (11)$$

where $s_{i,j,k}$ denotes the score using i th pattern, j th algorithm, and metric k . The exhaustive evaluation results are collated to create a pattern ranking data set. We assess whether a monotonic relationship exists between a pair of reconstruction algorithms as pattern varies by computing Spearman's ranking correlation coefficient [54] over scores of a metric

$$\rho_{j,j',k} = \text{Spearman}(\{(s_{i,j,k}, s_{i,j',k}) \mid i = 0, \dots, 494\}),$$

where a tuple of scores $(s_{i,j,k}, s_{i,j',k})$ is treated as an observation, associated with which a p-value describes the probability that no monotonic relationship exists between them. The correlation coefficients are computed for every pair of algorithms over all 1494 images from NTIRE dataset.

Experiment Protocol: The overall procedure of the experiments is as follows. Given a scene radiance map, we exhaustively evaluate reconstruction quality of distinct reconstruction algorithms on synthesized sensor readouts for all patterns returned by our enumeration algorithm. Taking reconstruction quality scores as data samples, we calculate Spearman's ranking correlation coefficient for every pair of reconstruction algorithms. We conduct hypothesis testing

- H_0 : the reconstruction quality between two algorithms are uncorrelated as exposure pattern varies,
- H_a : the reconstruction quality are positively correlated.

This procedure is repeated on every sample of NTIRE [42] dataset. We report the average and median of correlation coefficients across 1494 images for each metric in Table VI.

Results: To give a sense of the reported value, we also show scatter plots of μ PSNR of pairs of reconstruction algorithms in Fig. 10, with detailed numbers shown in Table VI.

The three scattered plots in Fig. 10 are worth discussing. These three subplots are the μ PSNR comparison between LPA, ADMM-TV, and Restormer. The scattered plot shows a surprisingly strong correlation between any pair of the methods: If a 2×2 pattern favors LPA, it also favors ADMM-TV, and similarly for other pairs. Therefore, at least based on this limited set of experiments, we find that if a pattern is good, it is good for all reconstruction algorithms; if it is bad, it is bad for all reconstruction algorithms. By inspecting the numbers in Table VI, we further note that the Spearman's correlation coefficients are all in the range of 0.87 or above (for μ PSNR). For other evaluation metrics μ SSIM and μ LPIPS, we also see a high correlation coefficient.

We believe that this finding is new and perhaps less expected. The implication is that if we need to design the multiplexing pattern, there is no need to consider the image reconstruction algorithm. This is a good news from the point of a designer's perspective. Co-optimization is not always preferred because we do not want the patterns to be dependent on a particular algorithm. If we can modularize the designs of the two, the debugging and analysis of the methods will be significantly easier.

F. Real Experiments

In this section, we test the SVE-Risk and SNR-Risk on real camera raw readouts and show the feasibility of the proposed risk on real hardware for exposure pattern selection. Since no SVE sensor is available to us, we interlace real camera raw readouts to synthesize images captured with SVE patterns.

1) Experiment Settings and Procedure: For each high dynamic range scene, we use a Sony Alpha 7 II camera to capture 9 differently exposed LDR frames. We capture 5 HDR scenes in total. For ease of camera parameter calibration, we keep the camera ISO at 100 and gradually increased the exposure time from 1/80 sec to 3.2 sec, doubling from one frame to the next. Our imaging model based risks require knowledge of dark current and read noise level. For these two parameters, we estimate them by capturing three dark frames (ISO 100, exposure 1/80 sec, 0.2 sec, 1.6 sec). We generate a pseudo ground truth radiance map by fusing all 9 frames. We synthesize all possible SVE captures using following procedure:

- Enumerate all possible SVE patterns with 9 different exposures.
- For each pattern, pick the corresponding frames from the 9 raw frames.

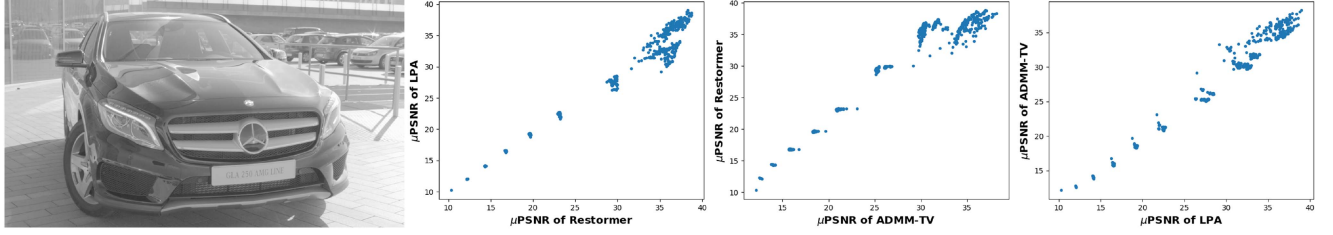


Fig. 10. Tone-mapped ground truth HDR scene and cross comparisons between reconstruction algorithms in μ PSNR. Each dot in the scatter plots represents an exposure pattern. The x and y coordinates are μ PSNR scores of the reconstructions by different algorithms on the image captured using that exposure pattern. The average correlation coefficient of the three scatter plots is approximately $\rho = 0.86$.

TABLE VII
MEASURED PARAMETERS IN TERMS OF EQUIVALENT CHARGES FOR REAL CAMERA EXPERIMENT

Gain α	5.25 e- per ADC Unit
Exposure τ	1/80 sec, 1/40 sec, ..., 1.6 sec
Read noise σ_{read}	1.3 e-
Camera response function	Linear
dark current	1.4e-/sec
Clip threshold	84667e-

TABLE VIII
 μ PSNR OF IMAGES RECONSTRUCTED FROM RAW IMAGES INTERLACED ACCORDING TO BEST POSSIBLE PATTERNS (ORACLE), PATTERNS SELECTED BY SVE-RISK, AND PATTERNS SELECTED BY SNR-RISK, COMPARED TO THE PSEUDO GROUND TRUTH

	scene 1	scene 2	scene 3	scene 4	scene 5
oracle	44.5	44.4	35.5	40.3	49.5
SVE-Risk top-1	44.5	42.3	33.9	38.8	49.3
SNR-Risk top-1	41.8	38.1	30.0	27.5	47.6

C) Interlace picked frames to generate an SVE frame by taking every other pixel.

Then, we use our trained Restormer to reconstruct scene radiance for every SVE captures and compare the reconstruction against the pseudo ground truth. We build the empirical radiance histogram of a scene by treating the SVE capture with exposure (1/80 sec, 1/20 sec, 1/5 sec, 4/5 sec) as the pilot frame. We use this histogram to calculate SVE-Risk for every pattern. For SNR-Risk, we used the pseudo ground truth of a scene for calculation.

The camera response function (CRF) of Sony Alpha 7 II camera raw readout is close to linear at ISO 100 except when the photon charge accumulated is near full well capacity. Therefore, we use a linear CRF and aggregate the entire conversion from charge to final ADC readout (voltage follower gain, column amplifier gain, and output amplifier gain) into conversion gain term α in (1). The final parameters used in this experiment are listed in Table VII.

2) *Results*: We show the quantitative reconstruction quality in terms of μ PSNR in Table VIII and two sets of qualitative results in Fig. 11. As evidenced in both quantitative and qualitative results, using top pattern selected by SVE-Risk can yield near-optimal final reconstruction.

G. Generalization to $N \times N$ and Color

In this section, we demonstrate the application of SVE-Risk to larger multiplexing patterns and color imaging.

Experiment Settings. Larger multiplexing patterns: Our previous discussions and analyses are based on a fixed pattern size of 2×2 . Here, we further test on 3×3 and 4×4 exposure/gain patterns using the same procedure in Section IV-D. To compute SVE-Risk, we previously use a 3×3 neighborhood $\mathcal{B}_i$ for 2×2 patterns, and here we use 5×5 and 7×7 neighborhoods for 3×3 and 4×4 pattern cases, respectively. For comparison purposes, we used the same number of exposure-gain levels as previous, resulting in a total of 24,310 and 735,471 possible patterns in 3×3 and 4×4 cases. We use Restormer as the reconstruction algorithm. We again illustrate the ranking power of SVE-Risk with scatter plots and visualize the reconstruction results for all pattern sizes. We also report the average performance measured by Δ_1 and $Q(1\%)$ metrics calculated for μ PSNR for each size. Due to the volume of the possible patterns, we do not carry out evaluation for all patterns. Instead, after ranking the patterns with SVE-Risk, we sample 100 patterns uniformly from the best to the worst in terms of their ranks. If the SVE-Risk works for larger multiplexing patterns, we should observe similar trend in the scatter plots as in Fig. 6.

Color imaging: Capturing and reconstructing color images is one of the most critical and challenging problems in imaging today. It usually requires both a special design of hardware such as color filter array and a corresponding software algorithm such as demosaicking. Exploring these color imaging pipelines or finding the optimal color reconstruction methods is beyond the scope of this article. Yet, we demonstrate the possibility of combining color filters and exposure multiplexing to reconstruct HDR color images. We start with the Quad-Bayer coding, where the four adjacent pixels are clustered with the same color using one color filter of red, green, or blue. With this color filter design, the periodic 2×2 exposure and gain pattern control can fit in such that each color channel has access to the full sets of exposure and gain levels. Next, image reconstructions can be applied on the captured signal to recover the color and HDR information. In our experiment, we synthesize the Quad-Bayer exposure-gain-multiplexed images as input, and we train a Restormer model to reconstruct both the color and the dynamic range. We demonstrate that this naive design is capable of producing reasonable results, and we consider developing

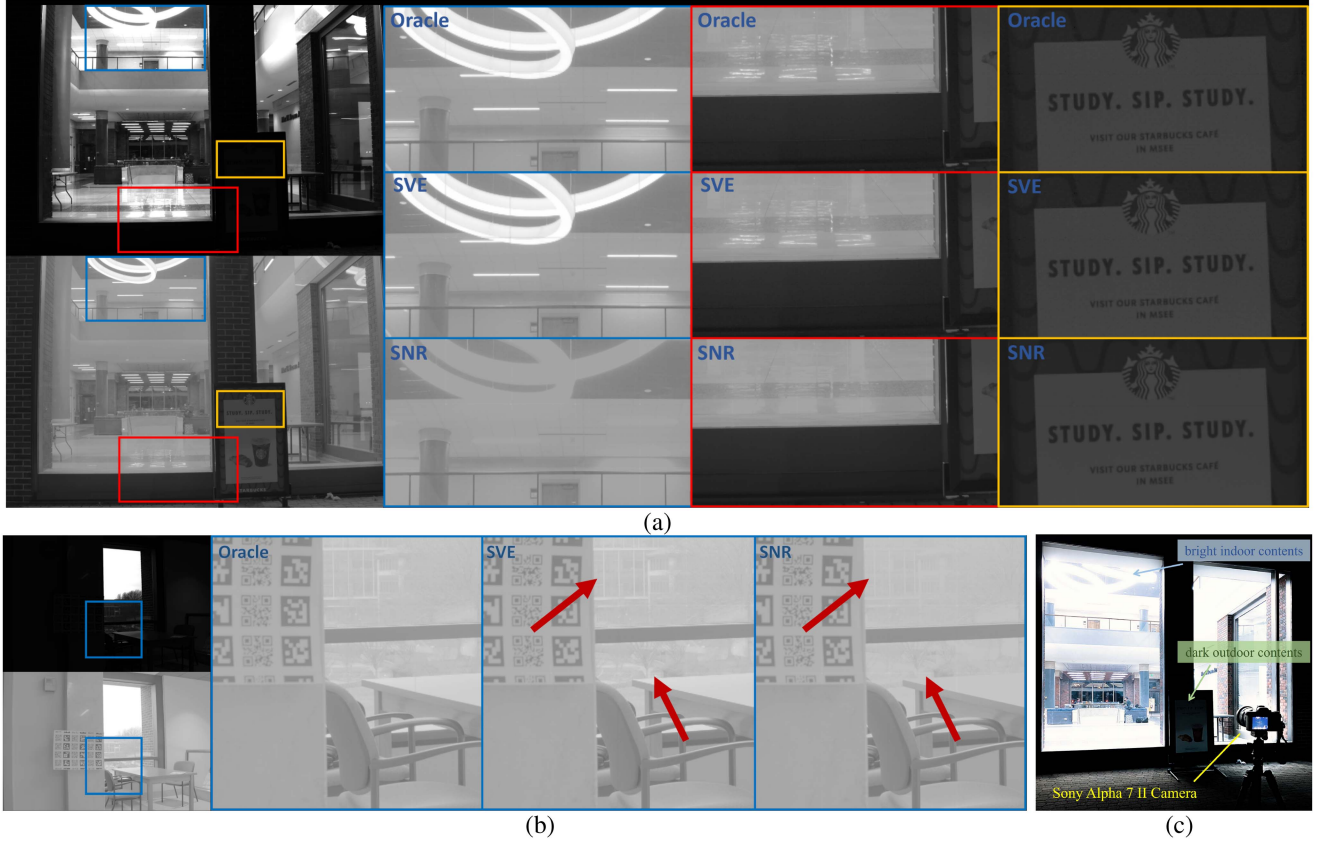


Fig. 11. Reconstruction results on real images and experimental setup. Input images with spatially-varying patterns are generated by interlacing raw images captured at various exposures. For (a) and (b), we show full frames of pseudo ground truth in both linear scale and tone-mapped scale (top-left and bottom-left, respectively). The linear scales are clipped at 95 percentile radiance level and are normalized to $[0, 255]$. We also show zoomed-in regions of tone-mapped reconstructions on the best possible spatially-varying pattern via exhaustive search (oracle), the best pattern selected by SVE-Risk (SVE), and the best by SNR-Risk (SNR). The reconstruction algorithm is Restormer for all images. (a) The texts and textures at low light shown in the orange box are preserved equally well in all three reconstructions, but SNR-Risk reconstruction loses all details of high-light regions. (b) The dark regions (QR code, wall, etc.) of all three reconstructions are indistinguishable, but SNR-Risk reconstruction loses outdoor high-light details (red arrows). (c) Experimental setup. We capture real images of HDR scenes using a Sony Alpha 7 II camera.

better exposure multiplexed color reconstruction a direction for future works.

The details of risk computation in our color reconstruction are as follows. We start by estimating the radiance distribution and calculating SVE-Risk for each color channel independently. Next, within each channel, all pixels of other colors are regarded as pixels to be interpolated, and their risks are calculated by (8) with usable non-saturated neighbors coming from other color channels. Finally, we sum the risks across all color channels to obtain the total risk.

Results. Larger multiplexing patterns: We show in Fig. 12 the reconstruction results and scatter plots of ranking power of SVE-Risk. We also show make a comparison of the average performance measured by Δ_1 and $Q(1\%)$ in Table IX. Generally, SVE-Risk is capable of ranking SVE patterns for all tested pattern sizes.

Theoretically, larger patterns correspond to better performance, because the best 4×4 pattern should be at least as good as the best 2×2 patterns as the set of all possible 4×4 patterns forms a superset of all possible 2×2 patterns. However, the improvement observed in the experiment does not fully support

TABLE IX
QUANTITATIVE RESULTS OF SVE-RISK ON DIFFERENT SIZES OF PATTERNS

μPSNR								
N	Δ_1	$Q(1\%)$	N	Δ_1	$Q(1\%)$	N	Δ_1	$Q(1\%)$
2	0.49	49.3	3	0.07	60.0	4	2.08	87.0

TABLE X
QUANTITATIVE RESULTS OF SVE-RISK ON GRAYSCALE AND COLOR IMAGE RECONSTRUCTION

μPSNR		
	Δ_1	$Q(1\%)$
2×2 Grayscale	0.49	49.3
$2 \times 2 + \text{Quad-Bayer}$	1.18	33.3

this theoretical correspondence. We observe an improvement of Δ_1 from 2×2 to 3×3 but a drop going to 4×4 . In addition, we observe $Q(1\%)$ gradually decreases as the size increases. We believe this is due to the limitation of learning-based reconstruction instead of the risk. A larger pattern size means a much larger set of possible patterns; therefore, to fit all of the patterns, the

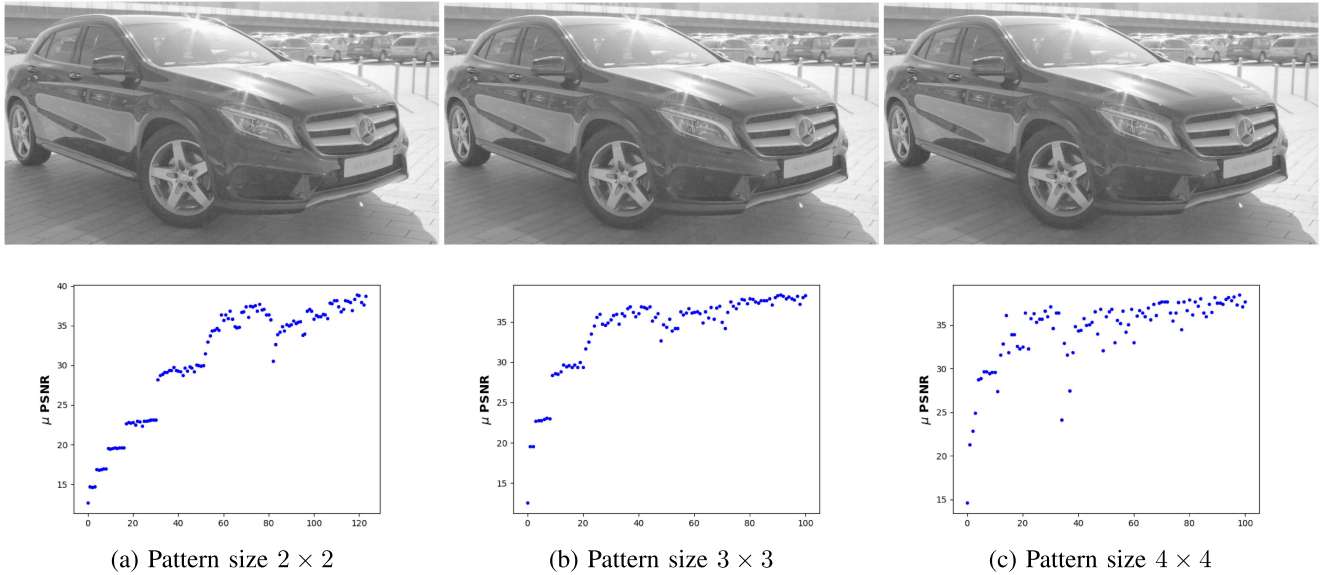


Fig. 12. [Top] Tone-mapped reconstructions for different SVE pattern size. [Bottom] Scatter plots of μ PSNR ranked from high risk to low risk for different SVE pattern size.



Fig. 13. Experimental results of SVE-Risk on color images. (a) tone-mapped ground truth, (b) tone-mapped reconstruction from the image captured with best pattern selected by SVE-Risk, (c) scatter plot of μ PSNR as sorted from highest risk to lowest risk ranked by SVE-Risk.

reconstruction model requires more training samples, a longer training time, and a larger model capacity. Due to limited training budget, we observe the model not fully converged in the cases of larger patterns, especially for 4×4 .

Color imaging: The color image reconstruction results are shown in Fig. 13, and the quantitative results evaluated with μ PSNR metric are reported in Table X. We find SVE-Risk also works effectively for color images. Quantitative results show better $Q(1\%)$ score for the color case than that in the grayscale case, while the reconstruction quality drop Δ_1 is higher.

V. CONCLUSION

In this article, we report two findings about the design of a spatially varying exposure multiplexing scheme. Firstly, we show that the pixel-wise SNR is a poor metric to quantify the performance of a multiplex pattern because it fails to differentiate the recoverable cases and the non-recoverable cases. We circumvent the difficulty by proposing the SVE-Risk. Our experiments show that the pattern ranking provided by the SVE-Risk correlates extremely well with the ideal ranking. Secondly,

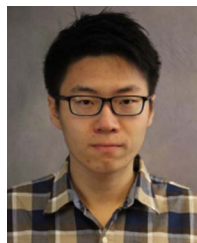
through a large-scale experiment, we find that for spatially-varying-exposure imaging with tiled exposure patterns, it is not necessary to design a pattern selection algorithm tailored for specific reconstruction algorithm; the margin of improvement for using tailored/co-designed pattern selection algorithm is limited. Our finding is a significant departure from recent work in computational photography that advocates for sensor-algorithm co-optimization. To sensor designers, this could be good news because sensor-algorithm co-design is significantly more costly for production. However, our bigger hope is that this counterexample can stimulate more discussions about the necessity of sensor-algorithm co-optimization, and under what context would it become beneficial *not* to co-optimize.

REFERENCES

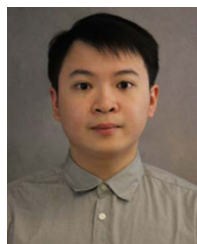
- [1] S. Nayar and T. Mitsunaga, "High dynamic range imaging: Spatially varying pixel exposures," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2000, pp. 472–479.
- [2] S. K. Nayar and V. Branzoi, "Adaptive dynamic range imaging: Optical control of pixel exposures over space and time," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2003, pp. 1168–1175.

- [3] G. Wan, X. Li, G. Agranov, M. Levoy, and M. Horowitz, "CMOS image sensors with multi-bucket pixels for computational photography," *IEEE J. Solid-State Circuits*, vol. 47, no. 4, pp. 1031–1042, Apr. 2012.
- [4] C. M. Nguyen, J. N. Martel, and G. Wetzstein, "Learning spatially varying pixel exposures for motion deblurring," in *Proc. IEEE Int. Conf. Comput. Photography*, 2022, pp. 1–11.
- [5] T. Klinghoffer, S. Somasundaram, K. Tiwary, and R. Raskar, "Physics vs. learned priors: Rethinking camera and algorithm design for task-specific imaging," in *Proc. IEEE Int. Conf. Comput. Photography*, 2022, pp. 1–12.
- [6] C. Aguerrebere, A. Almansa, Y. Gousseau, J. Delon, and P. Musé, "Single shot high dynamic range imaging using piecewise linear estimators," in *Proc. IEEE Int. Conf. Comput. Photography*, 2014, pp. 1–10.
- [7] S. Hajisharif, J. Kronander, and J. Unger, "Adaptive dual ISO HDR reconstruction," *EURASIP J. Image Video Process.*, vol. 2015, pp. 1–13, 2015.
- [8] M. Schobert, A. Belz, A. Nowak, J. Seiler, A. Kaup, and S. Foessel, "Building a high dynamic range video sensor with spatially nonregular optical filtering," *Proc. SPIE*, vol. 8499, pp. 102–112, 2012.
- [9] U. Cogalan, M. Bemana, K. Myszkowski, H. P. Seidel, and T. Ritschel, "Learning HDR video reconstruction for dual-exposure sensors with temporally-alternating exposures," *Comput. Graph.*, vol. 105, pp. 57–72, 2022.
- [10] F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar, "Generalized assorted pixel camera: Postcapture control of resolution, dynamic range, and spectrum," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2241–2253, Sep. 2010.
- [11] M. Alghamdi, Q. Fu, A. Thabet, and W. Heidrich, "Reconfigurable snapshot HDR imaging using coded masks and inception network," in *Proc. Int. Symp. Vis. Model. Vis.*, 2019.
- [12] Y. Jiang, I. Choi, J. Jiang, and J. Gu, "HDR video reconstruction with tri-exposure quad-Bayer sensors," 2021, *arXiv:2103.10982*.
- [13] S. J. Carey, D. R. Barr, B. Wang, A. Lopich, and P. Dudek, "Mixed signal SIMD processor array vision chip for real-time image processing," *Analog Integr. Circuits Signal Process.*, vol. 77, pp. 385–399, 2013.
- [14] Y. Luo, D. Ho, and S. Mirabbasi, "Exposure-programmable CMOS pixel with selective charge storage and code memory for computational imaging," *IEEE Trans. Circuits Syst. I: Regular Papers*, vol. 65, no. 5, pp. 1555–1566, May 2018.
- [15] H. Ke et al., "Extending image sensor dynamic range by scene-aware pixelwise-adaptive coded exposure," in *Proc. Int. Image Sensor Workshop*, 2019, pp. 111–114.
- [16] H. Reyserhove, A. S. Berkovich, and X. Liu, "Programmable pixel array," U.S. Patent 20200195828A1, Jun. 2020.
- [17] J. Zhang, R. Etienne-Cummings, T. Xiong, T. D. Tran, and S. H. Chin, "Flexible pixel-wise exposure control and readout," U.S. Patent 20180115725A1, Nov. 2019.
- [18] N. Sarhangnejad et al., "5.5 dual-tap pipelined-code-memory coded-exposure-pixel CMOS image sensor for multi-exposure single-frame computational imaging," in *Proc. IEEE Int. Solid-State Circuits Conf.*, 2019, pp. 102–104.
- [19] J. N. P. Martel, L. K. Müller, S. J. Carey, P. Dudek, and G. Wetzstein, "Neural sensors: Learning pixel exposures for HDR imaging and video compressive sensing with programmable sensors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 7, pp. 1642–1653, Jul. 2020.
- [20] J. Zhang et al., "A closed-loop, all-electronic pixel-wise adaptive imaging system for high dynamic range videography," *IEEE Trans. Circuits Syst. I: Reg. Papers*, vol. 67, no. 6, pp. 1803–1814, Jun. 2020.
- [21] W. Feng, F. Zhang, W. Wang, W. Xing, and X. Qu, "Digital micromirror device camera with per-pixel coded exposure for high dynamic range imaging," *Appl. Opt.*, vol. 56, pp. 3831–3840, 2017.
- [22] J. Kronander, S. Gustavson, G. Bonnet, and J. Unger, "Unified HDR reconstruction from raw CFA data," in *Proc. IEEE Int. Conf. Comput. Photography*, 2013, pp. 1–9.
- [23] Y. Deng, Q. Liu, and T. Ikenaga, "Multi-scale contextual attention based HDR reconstruction of dynamic scenes," in *Proc. 12th Int. Conf. Digit. Image Process.*, 2020, pp. 413–419.
- [24] Q. Yan et al., "A lightweight network for high dynamic range imaging," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2022, pp. 824–832.
- [25] L. Zhu, F. Zhou, B. Liu, and O. Göksel, "HDRfeat: A feature-rich network for high dynamic range image reconstruction," 2022, *arXiv:2211.04238*.
- [26] Y. Chi, X. Zhang, and S. H. Chan, "HDR imaging with spatially varying signal-to-noise ratios," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2023, pp. 5724–5734.
- [27] R. Pourreza-Shahri and N. Kehtarnavaz, "Exposure bracketing via automatic exposure selection," in *Proc. IEEE Int. Conf. Image Process.*, 2015, pp. 320–323.
- [28] N. Barakat, A. N. Hone, and T. E. Darcie, "Minimal-bracketing sets for high-dynamic-range image capture," *IEEE Trans. Image Process.*, vol. 17, no. 10, pp. 1864–1875, Oct. 2008.
- [29] B. Guthier, S. Kopf, and W. Effelsberg, "Optimal shutter speed sequences for real-time HDR video," in *Proc. IEEE Int. Conf. Imag. Syst. Techn. Proc.*, 2012, pp. 303–308.
- [30] K. F. Huang and J. C. Chiang, "Intelligent exposure determination for high quality HDR image generation," in *Proc. IEEE Int. Conf. Image Process.*, 2013, pp. 3201–3205.
- [31] K. Hirakawa and P. J. Wolfe, "Optimal exposure control for high dynamic range imaging," in *Proc. IEEE Int. Conf. Image Process.*, 2010, pp. 3137–3140.
- [32] S. W. Hasinoff, F. Durand, and W. T. Freeman, "Noise-optimal capture for high dynamic range photography," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2010, pp. 553–560.
- [33] Z. Wang, J. Zhang, M. Lin, J. Wang, P. Luo, and J. Ren, "Learning a reinforced agent for flexible exposure bracketing selection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 1820–1828.
- [34] Y. Chi, A. Gnanasambandam, V. Koltun, and S. H. Chan, "Dynamic low-light imaging with quanta image sensors," in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 122–138.
- [35] C. Li, X. Qu, A. Gnanasambandam, O. A. Elgendy, J. Ma, and S. H. Chan, "Photon-limited object detection using non-local feature matching and knowledge distillation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2021, pp. 3976–3987.
- [36] A. Gnanasambandam and S. H. Chan, "HDR imaging with quanta image sensors: Theoretical limits and optimal reconstruction," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 1571–1585, 2020.
- [37] S. H. Chan, "What does a one-bit quanta image sensor offer?," *IEEE Trans. Comput. Imag.*, vol. 8, pp. 770–783, 2022.
- [38] S. H. Chan, "On the insensitivity of bit density to read noise in one-bit quanta image sensors," *IEEE Sensors J.*, vol. 23, no. 4, pp. 3666–3674, Feb. 2023.
- [39] A. Gnanasambandam and S. H. Chan, "Exposure-referred signal-to-noise ratio for digital image sensors," *IEEE Trans. Comput. Imag.*, vol. 8, pp. 561–575, 2022.
- [40] S. H. Chan, X. Wang, and O. A. Elgendy, "Plug-and-play ADMM for image restoration: Fixed-point convergence and applications," *IEEE Trans. Comput. Imag.*, vol. 3, no. 1, pp. 84–98, Mar. 2017.
- [41] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, and M.-H. Yang, "Restormer: Efficient transformer for high-resolution image restoration," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2022, pp. 5728–5739.
- [42] E. Pérez-Pellitero, S. Catley-Chandar, A. Leonardis, and R. Timofte, "NTIRE 2021 challenge on high dynamic range imaging: Dataset, methods and results," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, 2021, pp. 691–700.
- [43] P. Korshunov, H. Nemoto, A. Skodras, and T. Ebrahimi, "Crowdsourcing-based evaluation of privacy in HDR images," *Proc. SPIE*, vol. 9138, 2014, Art. no. 913802.
- [44] N. K. Kalantari and R. Ramamoorthi, "Deep high dynamic range imaging of dynamic scenes," *ACM Trans. Graph.*, vol. 36, no. 4, pp. 1–12, 2017.
- [45] G. Yu, G. Sapiro, and S. Mallat, "Solving inverse problems with piecewise linear estimators: From Gaussian mixture models to structured sparsity," *IEEE Trans. Image Process.*, vol. 21, no. 5, pp. 2481–2499, May 2012.
- [46] C. Bao, J. F. Cai, and H. Ji, "Fast sparsity-based orthogonal dictionary learning for image restoration," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2013, pp. 3384–3391.
- [47] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian Denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.
- [48] S. W. Zamir et al., "Multi-stage progressive image restoration," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 14821–14831.
- [49] V. Monga, Y. Li, and Y. C. Eldar, "Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing," *IEEE Signal Process. Mag.*, vol. 38, no. 2, pp. 18–44, Mar. 2021.
- [50] D. Gilton, G. Ongie, and R. Willett, "Deep equilibrium architectures for inverse problems in imaging," *IEEE Trans. Comput. Imag.*, vol. 7, pp. 1123–1133, 2021.
- [51] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

- [52] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 586–595.
- [53] A. Chakrabarti, "Learning sensor multiplexing design through back-propagation," in *Proc. 30th Int. Conf. Neural Inf. Process. Syst.*, 2016, pp. 3089–3097.
- [54] D. Zwillinger and S. Kokoska, *CRC Standard Probability and Statistics Tables and Formulae*. Boca Raton, FL, USA: Chapman & Hall, 2000.



Xiangyu Qu (Student Member, IEEE) received the B.S. degree in electrical engineering and the Ph.D. degree in electronic and computer engineering from Purdue University, West Lafayette, IN, USA, in 2017 and 2023, respectively. In 2016, he joined the i2Lab as a Team Member for the Signal Processing Cup 2016, and continued doing research on computational photography during his senior year. His research interests include photon limited imaging, computer vision, and machine learning. He was the recipient of the Purdue University Ross Fellowship.



Yiheng Chi (Graduate Student Member, IEEE) received the B.S. degree in 2018 in computer engineering from Purdue University, West Lafayette, IN, USA, where he is currently working toward the Ph.D. degree. His research interests include photon limited imaging, imaging on mobile devices, image and video restoration, denoising, computer vision, and machine learning. He is also a regular reviewer for international conferences and journals, including CVPR, ICIP, IEEE OPEN JOURNAL OF SIGNAL PROCESSING, IEEE TRANSACTIONS ON IMAGE PROCESSING, and

Optics Express.



Stanley H. Chan (Senior Member, IEEE) received the B.Eng. degree in electrical engineering from the University of Hong Kong, Hong Kong, in 2007, the M.A. degree in mathematics, and the Ph.D. degree in electrical engineering from University of California San Diego, La Jolla, CA, USA, in 2009 and 2011, respectively. From 2012 to 2014, he was a Post-doctoral Research Fellow with Harvard University, Cambridge, MA, USA. In 2014, he joined Purdue University, West Lafayette, IN, USA, where he is currently an Elmore Associate Professor of electronic and computer engineering and an Associate Professor of statistics. He has authored a popular undergraduate textbook *Introduction to Probability for Data Science*, Michigan Publishing 2021, and coauthored the book *Computational Imaging through Atmospheric Turbulence*, Now Publisher 2023. His research interests include single-photon imaging, imaging through atmospheric turbulence, and computational photography. He was the recipient of the 2022 IEEE Signal Processing Society Best Paper Award. He is also a Senior Area Editor of the IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING. From 2016 to 2018, he was an Associate Editor for the *OSA Optics Express*. From 2015 to 2020, he was an Elected Member of the IEEE Signal Processing Society Technical Committee in Computational Imaging.