

# Scattering and Gathering for Spatially Varying Blurs

Nicholas Chimitt , Member, IEEE, Xingguang Zhang , Graduate Student Member, IEEE,  
Yiheng Chi , Student Member, IEEE, and Stanley H. Chan , Senior Member, IEEE

**Abstract**—A spatially varying blur kernel  $h(\mathbf{x}, \mathbf{u})$  is specified by an input coordinate  $\mathbf{u} \in \mathbb{R}^2$  and an output coordinate  $\mathbf{x} \in \mathbb{R}^2$ . For computational efficiency, we sometimes write  $h(\mathbf{x}, \mathbf{u})$  as a linear combination of spatially invariant basis functions. The associated pixelwise coefficients, however, can be indexed by either the input coordinate or the output coordinate. While appearing subtle, the two indexing schemes will lead to two different forms of convolutions known as *scattering* and *gathering*, respectively. We discuss the origin of the operations. We discuss conditions under which the two operations are identical. We show that scattering is more suitable for simulating how light propagates and gathering is more suitable for image filtering such as denoising.

**Index Terms**—Spatially varying blur, basis representation, scattering, gathering.

## I. INTRODUCTION

In two-dimensional space, the convolution between an input image  $J(\mathbf{x})$  and a shift-invariant kernel  $h(\mathbf{x})$  produces an output image  $I(\mathbf{x})$  via the well-known integral

$$I(\mathbf{x}) = \int_{\mathbb{R}^2} h(\mathbf{x} - \mathbf{u}) J(\mathbf{u}) d\mathbf{u}. \quad (1)$$

In this equation,  $\mathbf{x} \in \mathbb{R}^2$  is a two-dimensional coordinate in the output space and  $\mathbf{u} \in \mathbb{R}^2$  is a coordinate in the input space. This definition is ubiquitous in all shift-invariant systems.

If the kernel  $h$  is spatially *varying*, then it is no longer a function of the coordinate difference  $\mathbf{x} - \mathbf{u}$  but a function of two variables  $\mathbf{x}$  and  $\mathbf{u}$ . The resulting kernel  $h(\mathbf{x}, \mathbf{u})$  will give the input-output relationship via the integral

$$I(\mathbf{x}) = \int_{\mathbb{R}^2} h(\mathbf{x}, \mathbf{u}) J(\mathbf{u}) d\mathbf{u}, \quad (2)$$

also known as the superposition integral.

While spatially varying kernels are more difficult to analyze because they cannot be directly handled by Fourier transforms [1], they are common in image *formation* and image *processing*

Manuscript received 10 March 2023; revised 11 December 2023 and 19 February 2024; accepted 22 February 2024. Date of publication 11 March 2024; date of current version 22 March 2024. This work was supported in part by the Intelligence Advanced Research Projects Activity (IARPA) under Contract 2022-21102100004, and in part by the National Science Foundation under Grant CCRS-2030570 and Grant IIS-2133032. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Marco Felipe Duarte. (Corresponding author: Stanley H. Chan.)

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

Digital Object Identifier 10.1109/TSP.2024.3375638

which are twins in many situations, such as in the case of kernel estimation [2], [3], [4] or image restoration [5], [6], [7]. In image formation, spatially varying kernels are used to model how light propagates from the object plane to the image plane. These kernels are known as point spread functions (PSFs) which may be spatially varying due to various degradations in the medium or aberrations in the imaging system (such as spherical aberration). In image processing, spatially varying kernels are used to filter the input image for applications such as denoising or interpolation. The spatially varying nature in these situations come from examples such as non-local edge-aware filters where the shapes and orientations of the filters change depending on the image.

The theme of this paper is about the decomposition of the kernel in terms of basis functions. If  $h$  is *spatially invariant*, we may express it via the equation

$$h(\mathbf{x} - \mathbf{u}) = \sum_{m=1}^M a_m \varphi_m(\mathbf{x} - \mathbf{u}), \quad (3)$$

where  $\{\varphi_1, \varphi_2, \dots, \varphi_M\}$  are orthogonal basis functions. These functions could be as simple as the derivatives of Gaussians, or they can be learned from a dataset of kernels via principal component analysis (PCA). The scalars  $\{a_1, a_2, \dots, a_M\}$  are the basis coefficients, often constructed according the local image statistics or the underpinning physics.

In the case of *spatially varying* kernels, the subject of this paper, the basis representation in (3) needs to be modified so that it can take the two variables  $\mathbf{x}$  and  $\mathbf{u}$  into account. However, there is an ambiguity due to the existence of two options which we call *gathering* and *scattering*:

$$(\text{Gathering}) \quad h(\mathbf{x}, \mathbf{u}) = \sum_{m=1}^M a_{\mathbf{x},m} \varphi_m(\mathbf{x} - \mathbf{u}), \quad (4)$$

$$(\text{Scattering}) \quad h(\mathbf{x}, \mathbf{u}) = \sum_{m=1}^M a_{\mathbf{u},m} \varphi_m(\mathbf{x} - \mathbf{u}). \quad (5)$$

In both options, the spatially varying kernel  $h(\mathbf{x}, \mathbf{u})$  is written as a combination of *invariant* kernels  $\{\varphi_1, \varphi_2, \dots, \varphi_M\}$ . These  $\varphi_m$ 's are spatially invariant, so they can be written as  $\varphi_m(\mathbf{x} - \mathbf{u})$ . The difference between the two options lies in the coefficient  $a_{\mathbf{x},m}$  and  $a_{\mathbf{u},m}$ . Both possibilities have utility; this paper is aimed at describing their appropriateness.

*Remark:* Readers may wonder if we can define  $h(\mathbf{x}, \mathbf{u})$  using a global  $a_m$  instead of a pixelwise  $a_{\mathbf{x},m}$  or  $a_{\mathbf{u},m}$ . If we do so, for example by defining  $h(\mathbf{x}, \mathbf{u}) = \sum_{m=1}^M a_m \varphi_m(\mathbf{x} - \mathbf{u})$ , then

$h(\mathbf{x}, \mathbf{u})$  will be invariant because it is a linear combination of invariant basis functions. This will defeat the purpose studying a set of varying kernels.  $\square$

At first glance, the two choices above seem subtle to an extent that one may expect a minor difference in terms image quality. However, the two equations have two fundamentally different physical meanings. To give readers a preview of the main claims of the paper, we summarize as follows:

- **Gathering:**  $a_{\mathbf{x},m}$  is for image *processing* such as denoising filter, interpolation filter, etc.
- **Scattering:**  $a_{\mathbf{u},m}$  is for image *formation* such as modeling motion blur or atmospheric turbulence.

The study of time variant systems goes back to classical time-variant filter banks, wavelets, or Kalman filtering/state-space applications [8], [9], [10], [11], [12], [13]. The study of spatially varying kernels as a sum of invariant ones has been performed in works such as [14], [15], [16], [17], [18], [19] as applied models but often do not address the reason for the forward model chosen. The role of this work is to clearly highlight the difference between the two approximations from the side of modeling and describing where each one is more appropriately applied.

## II. THE OPERATIONS OF GATHERING AND SCATTERING

### A. Understanding Gathering

To present the operation of gathering, we can substitute the basis representation of (4) into (2), resulting in

$$\begin{aligned} I(\mathbf{x}) &\stackrel{\text{by (2)}}{=} \int_{\mathbb{R}^2} h(\mathbf{x}, \mathbf{u}) J(\mathbf{u}) d\mathbf{u} \\ &\stackrel{\text{by (4)}}{=} \int_{\mathbb{R}^2} \left( \sum_{m=1}^M a_{\mathbf{x},m} \varphi_m(\mathbf{x} - \mathbf{u}) \right) J(\mathbf{u}) d\mathbf{u} \\ &= \sum_{m=1}^M a_{\mathbf{x},m} \left( \int_{\mathbb{R}^2} \varphi_m(\mathbf{x} - \mathbf{u}) J(\mathbf{u}) d\mathbf{u} \right). \end{aligned}$$

Recognizing that the integral is a spatially invariant convolution, we can show that

$$I(\mathbf{x}) = \sum_{m=1}^M a_{\mathbf{x},m} \underbrace{(\varphi_m * J)(\mathbf{x})}_{\text{invariant blur}}. \quad (6)$$

linear combination of invariant blurs

Therefore, spatially varying convolution is replaced by a sum of  $M$  spatially *invariant* convolutions. Each term in the sum  $(\varphi_m * J)(\mathbf{x})$  is the convolution between the image  $J(\mathbf{u})$  and the basis functions  $\varphi_m(\mathbf{u})$ .

In terms of computation, we start with the image  $J(\mathbf{u})$ , precompute the filtered images  $(\varphi_m * J)(\mathbf{x})$ , then form a weighted sum of these images to construct the output  $I(\mathbf{x})$ . The weights are pixelwise, and so the resulting blur is spatially varying. Fig. 1 illustrates how this idea can be implemented. Because

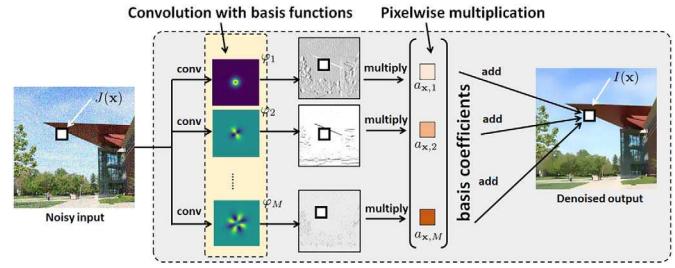


Fig. 1. **Gathering is for image processing.** In gathering, we construct a set of convolved images  $(\varphi_m * J)(\mathbf{x})$  and *gather* them via pixelwise multiplication. The process offers a significant saving in terms of computation.

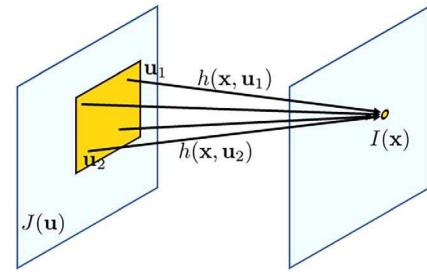


Fig. 2. The operation derived in (6) follows the intuitive argument that a neighboring input pixels are mapped to an output. Thus, each  $h(\mathbf{x}, \mathbf{u})$  is defined based on the output coordinate  $\mathbf{x}$ : At every  $\mathbf{x}$ , there is a set of basis functions that define  $h(\mathbf{x}, \mathbf{u})$ .

of the ordering of operations, this has also been referred to as convolution-product [20] in the more mathematical literature.

The illustration in Fig. 1 should remind readers of many classical signal processing techniques. For example, the steerable filters by Freeman and Adelson [21] employed the idea to detect image edges using a bank of Laplacian filters. The idea was also used in applied mathematics literature. For example, Nagy and O’Leary [22] used the idea to decompose the spatially varying kernels using non-overlapped blocks. Nagy and O’Leary’s formulation is a special case of our model by making the coefficients  $a_{\mathbf{x},m}$  as a binary mask. In [23], Popkin et al. used the idea to perform fast computation of spatially varying blurs with the basis functions computed via PCA.

To readers who prefer the perspective of deep learning, gathering as in (4) is what we call *convolution* in deep neural networks. To see why gathering is identical to the convolution (in deep-learning), we can refer to Fig. 2. Starting with a local neighborhood in the input space, we assign each pixel a weight which is specified by the kernel and map the sum to the output coordinate.

We summarize our findings here:

### What is **gathering (convolution-product)?**

- We apply spatially invarying kernels first, and then combine the results with weights.
- Consistent with convolution presented in Oppenheim and Wilsky [24], “flip, shift, and integrate”.
- Equivalent to the “convolution” in deep neural networks.

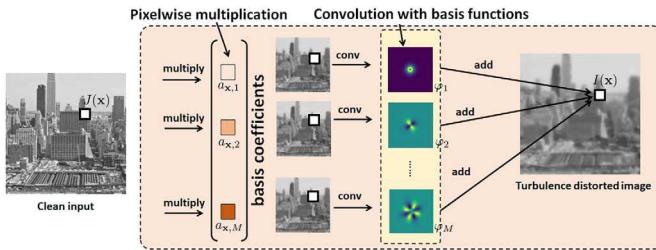


Fig. 3. **Scattering is for image formation.** Here we show a modified version of Mao et al.’s phase-to-space (P2S) transform [25] by switching the order of multiplication and convolution. The new formulation follows the derivation in (7). We argue that this formulation is the better (and the physically valid) approach when modeling how light propagates from the object plane to the image plane.

### B. Understanding Scattering

For scattering, we follow the same approach as we did for gathering. Substituting (5) into (2), we see that

$$\begin{aligned} I(\mathbf{x}) &\stackrel{\text{by (2)}}{=} \int_{\mathbb{R}^2} h(\mathbf{x}, \mathbf{u}) J(\mathbf{u}) d\mathbf{u} \\ &\stackrel{\text{by (5)}}{=} \int_{\mathbb{R}^2} \left( \sum_{m=1}^M a_{\mathbf{u},m} \varphi_m(\mathbf{x} - \mathbf{u}) \right) J(\mathbf{u}) d\mathbf{u} \\ &= \sum_{m=1}^M \left( \int_{\mathbb{R}^2} \varphi_m(\mathbf{x} - \mathbf{u}) \cdot (a_{\mathbf{u},m} J(\mathbf{u})) d\mathbf{u} \right). \end{aligned}$$

The last equation is trickier to understand. For a fixed index  $m$ , the product term  $a_{\mathbf{u},m} J(\mathbf{u})$  is a pixelwise multiplication of the coefficient map  $a_{\mathbf{u},m}$  and the input image  $J(\mathbf{u})$ . Therefore, if we express it using the elementwise multiplication  $\odot$ , we can write the equation as

$$I(\mathbf{x}) = \sum_{m=1}^M \underbrace{\left( \varphi_m * \underbrace{\left( a_m \odot J \right)}_{\text{pixelwise weighted input}} \right)}_{\text{linear combination of invariant blurs}}(\mathbf{x}). \quad (7)$$

Again thinking in terms of computation, we need to first construct the  $M$  coefficient maps  $\{a_{\mathbf{u},1}, \dots, a_{\mathbf{u},M}\}$  where each  $a_{\mathbf{u},m}$  is a value in an array of the same size as the image  $J$ . We multiply  $J$  with each coefficient map to obtain  $M$  ‘‘weighted images’’  $\tilde{J}_m(\mathbf{u}) = a_{\mathbf{u},m} J(\mathbf{u})$ . Then we convolve these  $M$  intermediate images through  $(\varphi_m * \tilde{J}_m)(\mathbf{u})$ . Finally, summing over the  $M$  images will give us the final result.

The process is summarized in Fig. 3. Comparing it with Fig. 1, we notice that the difference is order of multiplication and convolution. In Fig. 1, the pixelwise multiplication is done *after* the convolutions whereas in Fig. 3, it is done *before* the convolutions. Since the two processes involve the same number of convolutions and multiplications, the complexity of executing the computation is identical.

The role of scattering has been considered in a variety of mathematical studies, such as in Busby et al. [20] or Escande et al. [26] in which various bounds can be placed on the

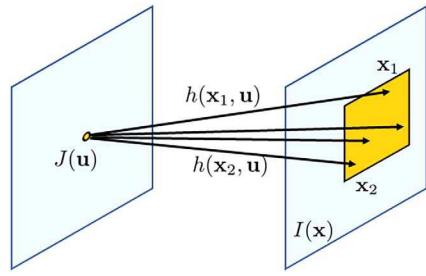


Fig. 4. The operation behind the scattering equation: a point source  $\delta(\mathbf{u})$  emits light which propagates to the object plane. The point spread function (aka the blur kernel  $h$ ) seen in the object plane is centered at  $\mathbf{u}$ . At each fixed  $\mathbf{u}$ , the pixel values  $h(\mathbf{x}, \mathbf{u})$  are determined by how the propagation medium affects the amplitude and phase at coordinate  $\mathbf{x}$ .

approximation of a product-convolution for a general operator (2). Alger et al. [27] demonstrated that product-convolution arises naturally in the case of locally-invariant kernels and proposes an adaptive scheme for the coefficients.

The situation described in (7) is the mirror of Fig. 2. Instead of having a local neighborhood whose pixels are mapped to a common output coordinate, the input pixel distributes its influence to a local neighborhood in the output space, as shown in Fig. 4. In deep learning, we call it the *transposed convolution*, which is typically used to upsample the features in tasks such as image super-resolution.

We summarize our findings here:

What is **scattering (product-convolution)?**

- We apply weighted averaging first, and then filter the weighted averages, and finally add.
- Consistent with how light propagates. See Goodman’s *Fourier Optics* [28] and our next section.
- Equivalent to the ‘‘transposed convolution’’ in deep neural networks.

### C. Conditions for Equivalence

After elaborating on the computations of gathering and scattering, we now explain the conditions under which the two are equivalent. As expected, the two are only equivalent in highly special cases.

Let us look at the equations more carefully through the lens of matrices and vectors. Let  $\mathbf{H} \in \mathbb{R}^{N \times N}$  be the matrix representation of the spatially varying blur kernel  $h(\mathbf{x}, \mathbf{u})$ . We assume that there is a set of *circulant* matrices  $\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_M$  representing the set of  $M$  spatially invariant basis functions  $\{\varphi_1, \varphi_2, \dots, \varphi_M\}$ .

We consider two sets of diagonal matrices. For every index  $m$ , we define

$$\mathbf{D}_m^{(g)} = \text{diag} \left\{ \begin{bmatrix} a_{\mathbf{x}_1, m} \\ \vdots \\ a_{\mathbf{x}_N, m} \end{bmatrix} \right\}, \quad \mathbf{D}_m^{(s)} = \text{diag} \left\{ \begin{bmatrix} a_{\mathbf{u}_1, m} \\ \vdots \\ a_{\mathbf{u}_N, m} \end{bmatrix} \right\}$$

Then, the gathering and scattering equations are

$$(\text{Gathering}) : \mathbf{H}^{(g)} = \sum_{m=1}^M \mathbf{D}_m^{(g)} \mathbf{H}_m \quad (8)$$

$$(\text{Scattering}) : \mathbf{H}^{(s)} = \sum_{m=1}^M \mathbf{H}_m \mathbf{D}_m^{(s)} \quad (9)$$

In other words, the difference lies in how we order the diagonal matrices and the spatially invariant convolution matrices. We now present a theorem that can be used to show gathering and scattering are mutually exclusive in a wide range of cases.

*Theorem 1:* Let  $\mathbf{H} \in \mathbb{R}^{N \times N}$  be a circulant matrix formed by  $[h_0, \dots, h_{N-1}]$  with at least two consecutive non-zero elements,  $h_m \neq 0$  and  $h_{m+1} \neq 0$ . Let  $\mathbf{A} = \text{diag}[a_0, \dots, a_{N-1}]$  and  $\mathbf{B} = \text{diag}[b_0, \dots, b_{N-1}]$  be two diagonal matrices. Then,  $\mathbf{AH} = \mathbf{HB}$  if and only if  $\mathbf{A} = \mathbf{B} = \lambda \mathbf{I}$  for some constant  $\lambda$  where  $\mathbf{I}$  is the identity matrix.

*Proof:* For the forward direction, we begin by writing out the  $(i, j)$ th element of each matrix product

$$\begin{aligned} [\mathbf{AH}]_{ij} &= a_i h_{ij} = a_i h_{(j-i) \bmod N}, \\ [\mathbf{HB}]_{ij} &= b_j h_{ij} = b_j h_{(j-i) \bmod N}. \end{aligned}$$

Suppose  $h_m \neq 0$ , then for any pair  $(i, j)$  in the set  $\mathcal{M} = \{(i^*, j^*) | (j^* - i^*) \bmod N = m\}$ , the following holds true:  $a_i h_m = b_j h_m \implies a_i = b_j$ . The set  $\mathcal{M}$  can be written in terms of  $m$  as  $\mathcal{M} = \{(0, m), (1, (m+1) \bmod N), \dots, (N-1, (m+N-1) \bmod N)\}$ , with a similar set  $\mathcal{M}'$  for  $m+1$  corresponding to  $h_{m+1} \neq 0$ . Through  $\mathcal{M}$  and  $\mathcal{M}'$  we have  $a_k = b_{(m+k) \bmod N}$  and  $a_l = b_{(m+1+l) \bmod N}$ .

Observing that  $a_0 = b_m \bmod N = b_{(m+1) \bmod N}$  and in general  $a_k = b_{(m+k) \bmod N} = b_{(m+k+1) \bmod N}$ , we set a sequence  $p_k = (m+k) \bmod N$  for  $k = 0, \dots, N-1$ . Since it can be observed  $p_k$  does not repeat within a cycle of  $N$ , we can establish  $b_{p_0} = \dots = b_{p_{N-1}} \implies b_0 = \dots = b_{N-1} \implies a_0 = \dots = a_{N-1}$ , and finally that all  $a_i$  and  $b_j$  are equal.

For the reverse direction,  $\mathbf{A} = \mathbf{B} = \lambda \mathbf{I}$  can be used to show  $\mathbf{AH} = \lambda \mathbf{IH} = \lambda \mathbf{H} = \mathbf{H} \lambda \mathbf{I} = \mathbf{HB}$ .  $\square$

The result of the previous theorem implies that if we have a convolution matrix  $\mathbf{H}_m$  (which is circulant by definition) that satisfies the consecutive non-zero property, for the scattering and gathering operations to be equivalent, we need

$$\mathbf{D}_m^{(g)} \mathbf{H}_m = \mathbf{H}_m \mathbf{D}_m^{(s)}, \quad \text{for all } m.$$

*Theorem 1* asserts that we need  $\mathbf{D}_m^{(g)} = \mathbf{D}_m^{(s)} = \lambda \mathbf{I}$ . But if  $\mathbf{D}_m^{(g)} = \mathbf{D}_m^{(s)} = \lambda \mathbf{I}$ , then the underlying blur must be spatially invariant.

Another consequence of Theorem 1 is that

$$\sum_{m=1}^M \mathbf{D}_m^{(g)} \mathbf{H}_m \neq \sum_{m=1}^M \mathbf{H}_m \mathbf{D}_m^{(s)}. \quad (10)$$

Therefore, the gathering and scattering equations (4) and (5) are *mutually exclusive* under the conditions stated in Theorem 1. If we say that  $h(\mathbf{x}, \mathbf{u})$  can be exactly represented by the gathering equation, then there will be an approximation error when representing  $h(\mathbf{x}, \mathbf{u})$  using the scattering equation, and vice versa.

Under what **conditions** would scattering = gathering?

- When the underlying blur is spatially invariant or does not satisfy the consecutive non-zero property.
- Scattering and gathering are mutually exclusive. We cannot simultaneously have (4) and (5) for a spatially varying blur. If one is the correct representation, the other will have approximation error.

#### D. Normalization

When performing a convolution, it is often necessary to ensure that the image intensity is not amplified or attenuated due to an improper normalization. For example, in a spatially invariant blur, we almost always require that

$$\int_{\mathbb{R}^2} h(\mathbf{u}) d\mathbf{u} = 1,$$

assuming that  $h(\mathbf{u}) \geq 0$  for all  $\mathbf{u}$ . Otherwise, if the integral is less than unity, the resulting (convolved) image will appear to be dimmer. Translated to matrices and vectors, this is equivalent to  $\mathbf{H}\mathbf{1} = \mathbf{1}$  for an all-one vector  $\mathbf{1}$ , assuming that  $\mathbf{H}$  is circulant.

Suppose that we have a sequence of spatially invariant blurs  $\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_M$  satisfying the property that  $\mathbf{H}_m \mathbf{1} = \mathbf{1}$  for all  $m$ . We want the diagonal matrices  $\mathbf{D}_1^x, \mathbf{D}_2^x, \dots, \mathbf{D}_M^x$  to be defined in such a way that the gathering equation (8) will give us

$$\mathbf{1} \stackrel{(\text{we want})}{=} \mathbf{H}^x \mathbf{1} = \left( \sum_{m=1}^M \mathbf{D}_m^{(g)} \mathbf{H}_m \right) \mathbf{1} = \sum_{m=1}^M \mathbf{D}_m^{(g)} \mathbf{1}.$$

Therefore, as long as we can ensure that the sum of the  $M$  diagonal matrices  $\{\mathbf{D}_m^{(g)} | m = 1, \dots, M\}$  is a vector of all one's, we are guaranteed to have  $\mathbf{H}^x$  to have unit rows. Converting this into the basis representation, it is equivalent to asking

$$\sum_{m=1}^M a_{\mathbf{x}, m} = 1, \quad \text{for all } m, \quad (11)$$

which is reasonably easy to satisfy. For implementation, if  $\mathbf{H}^x$  does not have rows sum to the unity such that  $\mathbf{H}^x \mathbf{1} \neq \mathbf{1}$ , the simplest approach is to define a diagonal matrix  $\mathbf{D}$  such that  $\mathbf{D}^{-1} \mathbf{H}^x \mathbf{1} = \mathbf{1}$ . From the derivations above, it is clear that the diagonal matrix should have the elements

$$\mathbf{D} = \text{diag} \left\{ \sum_{m=1}^M \mathbf{D}_m^{(g)} \mathbf{H}_m \mathbf{1} \right\}. \quad (12)$$

Therefore, the overall operation applied to an image is

$$\widehat{\mathbf{I}} = \text{diag} \left\{ \sum_{m=1}^M \mathbf{D}_m^{(g)} \mathbf{H}_m \mathbf{1} \right\}^{-1} \left( \sum_{m=1}^M \mathbf{D}_m^{(g)} \mathbf{H}_m \mathbf{J} \right), \quad (13)$$

where  $\mathbf{J} \in \mathbb{R}^N$  is the vector representation of the input, and  $\widehat{\mathbf{I}} \in \mathbb{R}^N$  is the output.

The normalization of the scattering equation is more complicated because the diagonal matrices do not commute with

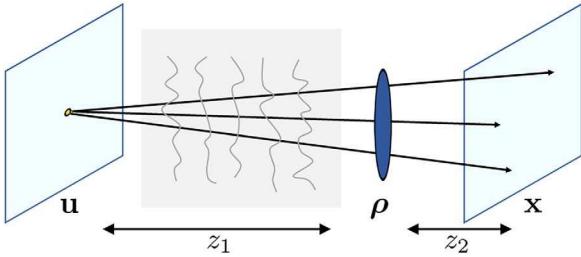


Fig. 5. The coordinate system of a typical optical system, in the presence of a random medium.

the spatially invariant blurs, and so we are not able to simplify the equation:

$$\mathbf{1} \stackrel{\text{(we want)}}{=} \mathbf{H}^u \mathbf{1} = \sum_{m=1}^M \mathbf{H}_m \left( \mathbf{D}_m^{(s)} \mathbf{1} \right).$$

The only work around solution is to define a diagonal matrix  $\mathbf{D}$  such that  $\mathbf{D}^{-1} \mathbf{H}^u \mathbf{1} = \mathbf{1}$ . This would require that

$$\mathbf{D} = \text{diag} \left\{ \sum_{m=1}^M \mathbf{H}_m \mathbf{D}_m^{(s)} \mathbf{1} \right\}. \quad (14)$$

In other words, when computing the scattering equation and if normalization is required, then we need to compute the set of convolutions *twice*: once for the input, and once for the normalization constant,

$$\widehat{\mathbf{I}} = \text{diag} \left\{ \sum_{m=1}^M \mathbf{H}_m \mathbf{D}_m^{(g)} \mathbf{1} \right\}^{-1} \left( \sum_{m=1}^M \mathbf{H}_m \mathbf{D}_m^{(g)} \mathbf{J} \right). \quad (15)$$

To summarize the normalization:

If **normalization** is needed, then

- For gathering: Either ensure  $\sum_{m=1}^M a_{\mathbf{x},m} = 1$  for all  $\mathbf{x}$ , or perform the calculation in (13).
- For scattering: Perform the calculation in (15). The cost is twice the number of convolutions than gathering.

### III. ORIGIN OF THE SCATTERING EQUATION

In this section, we explain the origin of the scattering equation from a physics point of view. To make our discussions concrete, the running example we use is a point source propagating a random medium as shown in Fig. 5.

We follow the coordinate system defined in Fig. 5. The object plane uses the coordinate  $\mathbf{u} \in \mathbb{R}^2$ . We can think of it as the *input* coordinate. As light propagates through the random medium, it reaches at the aperture of a lens. The coordinate on the lens is denoted by  $\rho \in \mathbb{R}^2$ . The image plane uses coordinate  $\mathbf{x} \in \mathbb{R}^2$ , which is also the *output* coordinate. Deriving the PSF equation from the Rayleigh-Sommerfeld integral will be too lengthy for our paper. Therefore, we skip the derivation and refer the readers to [28, Ch. 4 & 5] or [29]. Our approach is to highlight the four components of Fig. 5.

**Source to Aperture.** The propagation of a point source from the object plane to the aperture, in the absence of the random medium, is characterized by the free-space propagation. The electromagnetic field defined upon the aperture is given by [28, Eq 5-25]

$$U(\mathbf{u}, \rho) = \frac{1}{j\lambda z_1} \exp \left\{ j \frac{k}{2z_1} |\mathbf{u} - \rho|^2 \right\}, \quad (16)$$

where  $k = 2\pi/\lambda$  is the wave number, and  $z_1$  is the distance from the source to the aperture. The notation  $|\mathbf{u} - \rho|$  denotes the Euclidean distance between the two coordinates  $\mathbf{u}$  and  $\rho$ . This equation describes a parabolic wave propagating outward from  $\mathbf{u}$ . The farther apart  $\mathbf{u}$  and  $\rho$  is, the weaker the field  $U(\mathbf{u}, \rho)$  will become.

**Aperture and lens.** Right at the lens, the incident field will be imparted by the pupil function of the lens and its phase response. For a lens with a focal length of  $f$ , the field at the exit of the aperture is [28, Eq 5-26]

$$U'(\mathbf{u}, \rho) = U(\mathbf{u}, \rho) P(\rho) \exp \left\{ j \frac{k}{2f} |\rho|^2 \right\}, \quad (17)$$

where  $P(\rho)$  is the pupil function, typically chosen to be a circular indicator function.

**Aperture to image.** When the incident field exits the lens, it propagates via Fresnel diffraction to the image plane. Referring to [28, Eq 5-27], we can show that

$$\underbrace{U''(\mathbf{x}, \mathbf{u})}_{=h(\mathbf{x}, \mathbf{u})} = \frac{1}{j\lambda z_2} \int_{\mathbb{R}^2} U'(\mathbf{u}, \rho) \exp \left\{ j \frac{k}{2z_2} |\mathbf{x} - \rho|^2 \right\} d\rho. \quad (18)$$

Notice that the final electromagnetic field  $U''(\mathbf{x}, \mathbf{u})$  arriving at the image plane is originated from a point source. As such,  $U''(\mathbf{x}, \mathbf{u})$  is the point spread function  $h(\mathbf{x}, \mathbf{u})$ .

The PSF  $h(\mathbf{x}, \mathbf{u})$  can be expressed (with some approximation) as [28, Eq 5-36]:

$$h(\mathbf{x}, \mathbf{u}) = \underbrace{\kappa \int_{\mathbb{R}^2} P(\rho) \exp \left\{ -j \frac{k}{z_2} (\mathbf{x} - S\mathbf{u})^T \rho \right\} d\rho}_{=h(\mathbf{x}-\mathbf{u}) \text{ if } S=1}, \quad (19)$$

where  $S = -z_2/z_1$  is the magnification factor and  $\kappa = 1/(\lambda^2 z_1 z_2)$ . If, for simplicity, we assume  $z_1 = -z_2$  so that  $S = 1$ , then  $h(\mathbf{x}, \mathbf{u})$  is completely characterized by the coordinate difference  $\mathbf{x} - \mathbf{u}$ . This will give us  $h(\mathbf{x}, \mathbf{u}) = h(\mathbf{x} - \mathbf{u})$ , and so  $h(\mathbf{x}, \mathbf{u})$  represents a spatially *invariant* kernel.

**Random medium.** The fourth element we need to discuss, which is also the source of the problem, is the random medium. The random medium introduces a random amplitude and phase distortion as

$$R_{\mathbf{u}}(\rho) = \underbrace{A_{\mathbf{u}}(\rho)}_{\text{amplitude}} \times \underbrace{\exp \{ -j \phi_{\mathbf{u}}(\rho) \}}_{\text{phase}}. \quad (20)$$

Notice that in this definition, the distortion has a coordinate pair  $(\mathbf{u}, \rho)$ . The position  $\mathbf{x}$  is not present and has no impact at this

stage. Since the impact of the random medium takes place from the source-to-aperture, we append (20) to (16) to find

$$U(\mathbf{u}, \boldsymbol{\rho}) = \frac{A_{\mathbf{u}}(\boldsymbol{\rho})}{j\lambda z_1} \exp\{-j\phi_{\mathbf{u}}(\boldsymbol{\rho})\} \exp\left\{j\frac{k}{2z_1}|\mathbf{u} - \boldsymbol{\rho}|^2\right\}.$$

Consequently, the point spread function is

$$h(\mathbf{x}, \mathbf{u}) = \kappa \int_{\mathbb{R}^2} g_{\mathbf{u}}(\boldsymbol{\rho}) \exp\left\{-j\frac{k}{z_2}(\mathbf{x} - \mathbf{u})^T \boldsymbol{\rho}\right\} d\boldsymbol{\rho} \quad (21)$$

where again we assumed  $S = 1$  and defined

$$g_{\mathbf{u}}(\boldsymbol{\rho}) \stackrel{\text{def}}{=} A_{\mathbf{u}}(\boldsymbol{\rho}) \exp\{-j\phi_{\mathbf{u}}(\boldsymbol{\rho})\} P(\boldsymbol{\rho}).$$

Therefore, (21) can be seen as the Fourier transform of the random medium and the pupil function via

$$h(\mathbf{x}, \mathbf{u}) = \text{Fourier}\left\{g_{\mathbf{u}}(\boldsymbol{\rho})\right\} \Big|_{\frac{\mathbf{x}-\mathbf{u}}{\lambda z_2}}, \quad (22)$$

where we specify that the transform is evaluated at the coordinate  $(\mathbf{x} - \mathbf{u})/(\lambda z_2)$ . This indicates that the PSF  $h(\mathbf{x}, \mathbf{u})$  is, in part, written as a function of  $\mathbf{x} - \mathbf{u}$ . However, since  $g_{\mathbf{u}}(\boldsymbol{\rho})$  is indexed by  $\mathbf{u}$ , the resulting PSF  $h(\mathbf{x}, \mathbf{u})$  should inherit the index  $\mathbf{u}$ . This will give us

$$\begin{aligned} h(\mathbf{x}, \mathbf{u}) &= \text{some function of } (\mathbf{x} - \mathbf{u}), \text{ index by } \mathbf{u}, \\ &= \sum_{m=1}^M a_{\mathbf{u}, m} \varphi_m(\mathbf{x} - \mathbf{u}), \end{aligned} \quad (23)$$

where in the last step we use the linear combination of basis functions as the model for such  $h(\mathbf{x}, \mathbf{u})$ . The basis function  $\varphi_m$  here captures the spatial invariance, whereas the coefficients  $a_{\mathbf{u}, m}$  capture the spatially varying indices.

This derivation explains why in optics simulation, such as imaging through random media, one must follow the *scattering* equation if we choose to represent the PSF using a set of spatially invariant basis functions. This is due to the *source* location determining the response of the system. This extends to other optical distortions where the source location parameterizes the system error.

#### IV. DO THESE ACTUALLY MATTER?

The question to ask now is: given the gathering equation and the scattering equation, does it really matter if we choose the “wrong” one? The goal of this section is to answer this question through a few examples.

##### A. Scattering Works for Optical Simulation

Our first example considers the problem of simulating the resulting image given an incoherent light source. The light source  $J(\mathbf{u})$  we consider here consists of two delta functions:

$$J(\mathbf{u}) = \delta(\mathbf{u} + \Delta) + \delta(\mathbf{u} - \Delta),$$

where  $\Delta$  is a small displacement. For convenience, we define a plane with two halves; relative to the separation,  $\delta(\mathbf{u} + \Delta)$  is on the left and  $\delta(\mathbf{u} - \Delta)$  is on the right.

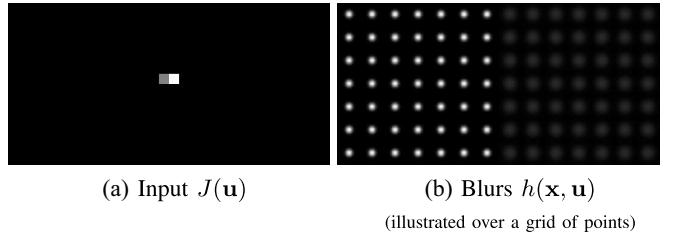


Fig. 6. Visualization of an example with (a)  $J(\mathbf{u})$  and (b) a grid of spatially varying blur kernels.

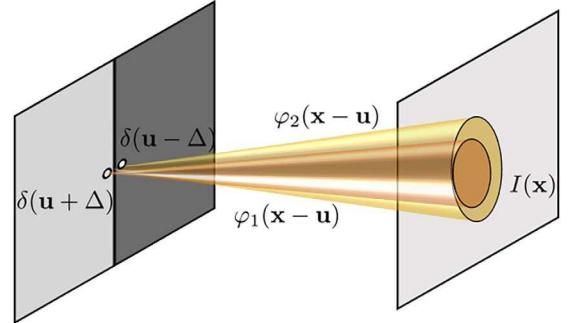


Fig. 7. Thought experiment with two points on the object plane, diffracting through two different metasurfaces. The resulting image should, in principle, be one superimposed diffraction pattern.

Imagine that in front of the light source, we put two transparent sheets with different phase profiles (which can be engineered using a meta-material). This will give us a spatially varying blur kernel  $h(\mathbf{x}, \mathbf{u})$ . By design of the transparent sheet, if light is emitted on the left hand side the blur has a smaller radius; if the light is emitted on the right hand side, then the blur has a larger radius. Thus, we write

$$h(\mathbf{x}, \mathbf{u}) = \begin{cases} \frac{1}{2\pi\sigma_1^2} e^{-\frac{\|\mathbf{x}-\mathbf{u}\|^2}{2\sigma_1^2}} \stackrel{\text{def}}{=} \varphi_1(\mathbf{x} - \mathbf{u}), & \mathbf{u} \in \text{left}, \\ \frac{1}{2\pi\sigma_2^2} e^{-\frac{\|\mathbf{x}-\mathbf{u}\|^2}{2\sigma_2^2}} \stackrel{\text{def}}{=} \varphi_2(\mathbf{x} - \mathbf{u}), & \mathbf{u} \in \text{right}, \end{cases} \quad (24)$$

where  $\sigma_1 < \sigma_2$ . Fig. 6(b) illustrates these spatially varying blur kernels. For visualization purposes, we show only the PSFs at a grid of points. In reality, the PSFs are defined continuously over  $\mathbf{u}$ .

Before we do any calculus, we can perform a thought experiment. Fig. 7 illustrates a hypothetical experimental setup. On the object plane there are two points emitting light through a meta surface with two different phase profiles. As the light propagates outward from the source through diffraction, the electromagnetic fields superimpose over each other. When the light reaches the aperture, the two diffraction patterns overlap. Therefore, the resulting image  $I(\mathbf{x})$ , without any calculation, should be one big diffraction pattern. It is impossible to obtain a sharp cutoff and two diffraction patterns.

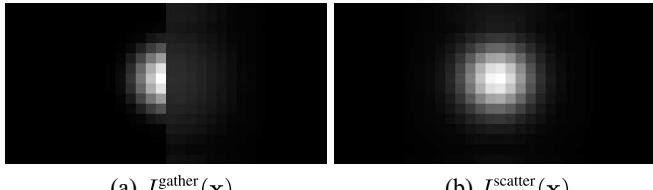
(a)  $I^{\text{gather}}(\mathbf{x})$  (b)  $I^{\text{scatter}}(\mathbf{x})$ 

Fig. 8. Comparison between gathering and scattering for the setup in Fig. 7. Notice that for this optical experiment, we should expect the resulting image to contain one big diffraction pattern. However, only the scattering equation demonstrates this.

With this in mind, we can write the PSF in terms of scattering and gathering as

$$h^{\text{gather}}(\mathbf{x}, \mathbf{u}) = \underbrace{\mathbb{I}\{\mathbf{x} \in \text{left}\}}_{=a_{\mathbf{x},1}} \times \varphi_1(\mathbf{x} - \mathbf{u}) + \underbrace{\mathbb{I}\{\mathbf{x} \in \text{right}\}}_{=a_{\mathbf{x},2}} \times \varphi_2(\mathbf{x} - \mathbf{u}), \quad (25)$$

$$h^{\text{scatter}}(\mathbf{x}, \mathbf{u}) = \underbrace{\mathbb{I}\{\mathbf{u} \in \text{left}\}}_{=a_{\mathbf{u},1}} \times \varphi_1(\mathbf{x} - \mathbf{u}) + \underbrace{\mathbb{I}\{\mathbf{u} \in \text{right}\}}_{=a_{\mathbf{u},2}} \times \varphi_2(\mathbf{x} - \mathbf{u}), \quad (26)$$

where we replaced  $a_{\mathbf{x},m}$  by  $a_{\mathbf{u},m}$  and  $\mathbb{I}\{\cdot\}$  is the indicator function. By comparing the gathering equation (25) and the scattering equation (26) with the original spatially varying  $h(\mathbf{x}, \mathbf{u})$  in (24), it is clear that only the scattering equation will match with the original  $h(\mathbf{x}, \mathbf{u})$  because they are both indexed by  $\mathbf{u}$ . However, to confirm that this is indeed the case, it would be helpful to look at the resulting image, as illustrated in Fig. 8.

We can also mathematically describe the images through use of the kernel  $h(\mathbf{x}, \mathbf{u})$  in either case in the superposition integral. The gathering kernel results in

$$I^{\text{gather}}(\mathbf{x}) = \int (a_{\mathbf{x},1}\varphi_1(\mathbf{x} - \mathbf{u}) + a_{\mathbf{x},2}\varphi_2(\mathbf{x} - \mathbf{u})) J(\mathbf{u}) d\mathbf{u}.$$

$$= \begin{cases} \varphi_1(\mathbf{x} + \Delta) + \varphi_1(\mathbf{x} - \Delta), & \mathbf{x} \in \text{left}, \\ \varphi_2(\mathbf{x} + \Delta) + \varphi_2(\mathbf{x} - \Delta), & \mathbf{x} \in \text{right}. \end{cases} \quad (27)$$

If we draw  $I^{\text{gather}}(\mathbf{x})$ , we will obtain the figure shown in Fig. 8(a). For scattering, we can carry out the same derivation and show that

$$I^{\text{scatter}}(\mathbf{x}) = \int_{-\infty}^{\infty} (a_{\mathbf{u},1}\varphi_1(\mathbf{x} - \mathbf{u}) + a_{\mathbf{u},2}\varphi_2(\mathbf{x} - \mathbf{u})) J(\mathbf{u}) d\mathbf{u}$$

$$= \varphi_1(\mathbf{x} + \Delta) + \varphi_2(\mathbf{x} - \Delta). \quad (28)$$

Similar to gathering, if we draw the resulting image, we will obtain the figure shown in Fig. 8(b). This is consistent with what we expect from Fig. 7 and the theoretical derivation in Section III.

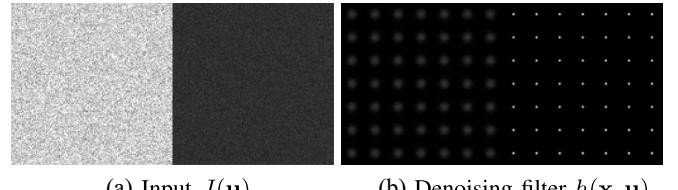
(a) Input  $J(\mathbf{u})$  (b) Denoising filter  $h(\mathbf{x}, \mathbf{u})$  (illustrated over a grid of points)

Fig. 9. Thought experiment of two noisy regions in an image. To denoise this image, ideally we would want to apply to different filters with a sharp boundary at the transition.

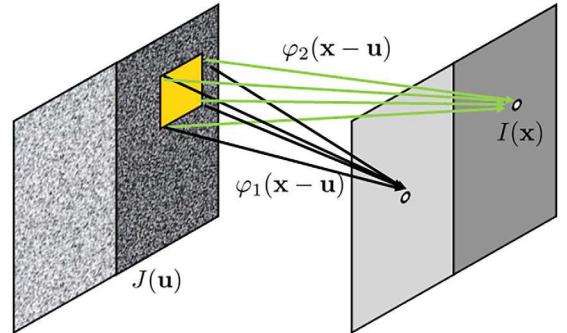


Fig. 10. Thought experiment with two noisy half-planes. As we perform the denoising step, we would hope that the sharp boundary is preserved.

### B. Gathering Works for Image Filtering

In the second example, we consider the problem of *image filtering*. The scenario is that we are given a noisy image  $J(\mathbf{u})$  that contains two regions:

$$J(\mathbf{u}) = \begin{cases} \theta_1 + \text{Gauss}(0, \sigma_1^2), & \mathbf{u} \in \text{left}, \\ \theta_2 + \text{Gauss}(0, \sigma_2^2), & \mathbf{u} \in \text{right}, \end{cases} \quad (29)$$

with two signal levels  $\theta_1$  and  $\theta_2$ , and two noise standard deviations  $\sigma_1$  and  $\sigma_2$  such that  $\sigma_1 > \sigma_2$ . In this equation,  $W_1(\mathbf{u}) \sim \text{Gauss}(0, \sigma_1^2)$  and  $W_2(\mathbf{u}) \sim \text{Gauss}(0, \sigma_2^2)$  denote Gaussian noise. For illustration, we show in Fig. 9(a) the case where  $\theta_1 = 0.8$ ,  $\theta_2 = 0.2$ ,  $\sigma_1 = 0.1$  and  $\sigma_2 = 0.02$ .

To denoise this image, we consider the simplest approach assuming that we *knew* the partition of the two regions. Suppose that we want to denoise the left side. Since we know that the noise is stronger, we shall apply a stronger filter. As illustrated in Fig. 10, for this filter to be effective along the boundary, we should apply a mask *after* the filtering is done.

The spatially varying filter we propose here takes the form

$$h(\mathbf{x}, \mathbf{u}) = \begin{cases} \frac{1}{\sqrt{2\pi s_1^2}} e^{-\frac{\|\mathbf{x}-\mathbf{u}\|^2}{2s_1^2}} & \stackrel{\text{def}}{=} \varphi_1(\mathbf{x} - \mathbf{u}), \quad \mathbf{x} \in \text{left}, \\ \frac{1}{\sqrt{2\pi s_2^2}} e^{-\frac{\|\mathbf{x}-\mathbf{u}\|^2}{2s_2^2}} & \stackrel{\text{def}}{=} \varphi_2(\mathbf{x} - \mathbf{u}), \quad \mathbf{x} \in \text{right}, \end{cases}$$

where we assume that  $s_1 > s_2$ . We are careful about the index in this equation, remarking that the conditions are applied to  $\mathbf{x}$  instead of  $\mathbf{u}$ . We will illustrate what will happen if the conditions are applied to  $\mathbf{u}$ .

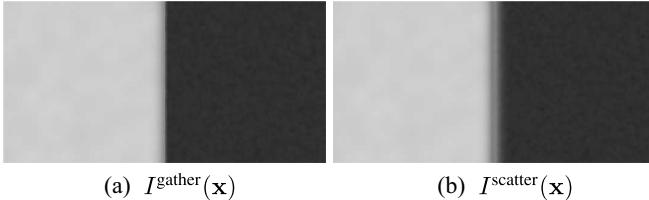


Fig. 11. Comparison between gathering and scattering for the setup in Fig. 9. Notice that for this denoising experiment, the better method should produce a sharp transition along the boundary.

The gathering and the scattering equation for this example are identical to those in (25) and (26). Most importantly, the coefficients  $a_{\mathbf{x},m}$  and  $a_{\mathbf{u},m}$  are binary masks indicating whether the pixel  $\mathbf{x}$  (or  $\mathbf{u}$ ) is on the left / right hand side.

The resulting images  $I^{\text{gather}}(\mathbf{x})$  and  $I^{\text{scatter}}(\mathbf{x})$  follow a similar derivation as in (27) and (28). It can be shown that the two possible approximations reduce to

$$I^{\text{gather}}(\mathbf{x}) = \begin{cases} (J * \varphi_1)(\mathbf{x}), & \mathbf{x} < 0, \\ (J * \varphi_2)(\mathbf{x}), & \mathbf{x} \geq 0, \end{cases} \quad (30)$$

$$I^{\text{scatter}}(\mathbf{x}) = [(J * \mathbb{I}\{\mathbf{x} \in \text{left}\}) * \varphi_1](\mathbf{x}) + [(J * \mathbb{I}\{\mathbf{x} \in \text{right}\}) * \varphi_2](\mathbf{x}). \quad (31)$$

This shows that gathering contains a sharp transition in its output while scattering is a sum of the two convolutions over partitions. We visualize the results in Fig. 11. While both of them can offer denoising to some extent, the gathering approach handles the boundary much better because the masking is performed *after* the filtering. If the masking is performed before the filtering, then (31) tells us that we are summing two convolutions of the same image. Therefore, the edge is blurred.

### C. Decompositions for Pixel-Wise PSFs

While the two previous examples illustrate the cases of binary masks, we now use a toy model of imaging through turbulence to demonstrate a case of a PSF per-pixel, requiring a more general approach than partitioning. Our toy model is Gaussian kernels with random shifts:

$$h_{\mathbf{u}}(\mathbf{x}) \propto \exp \left\{ -\frac{(\mathbf{x} - \mathbf{t}(\mathbf{u}))^T (\mathbf{x} - \mathbf{t}(\mathbf{u}))}{b(\mathbf{u})^2} \right\}. \quad (32)$$

Here  $\mathbf{t}(\mathbf{u})$  is a random vector field comprised of two independent fields  $X$  and  $Y$  as  $\mathbf{t}(\mathbf{u}) = [X(\mathbf{u}), Y(\mathbf{u})]^T$ . Furthermore  $b$  is a scalar random field which parameterizes the blur. While the exact details of the fields are not critical, we show the result of a grid of point sources distorted by our toy turbulence model in Fig. 12. For this example we primarily stick to mild blur and stronger geometric distortions, demonstrating some utility for motion-blur modeling.

Relying on per-pixel partitioning would require an entire convolution per-pixel and is therefore very computationally expensive. We instead represent each kernel as a sum of bases obtained via PCA. If an accurate PCA representation requires  $K \ll M$  basis functions, where  $M$  is the number of pixels in an image, this corresponds to a meaningful reduction in

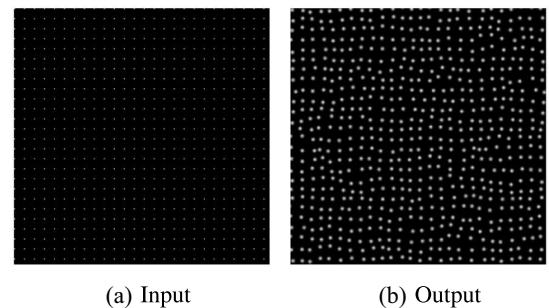


Fig. 12. Visualizing the effect of our toy turbulence model for point sources. We note the primary distortion given our chosen parameters is geometric distortion.

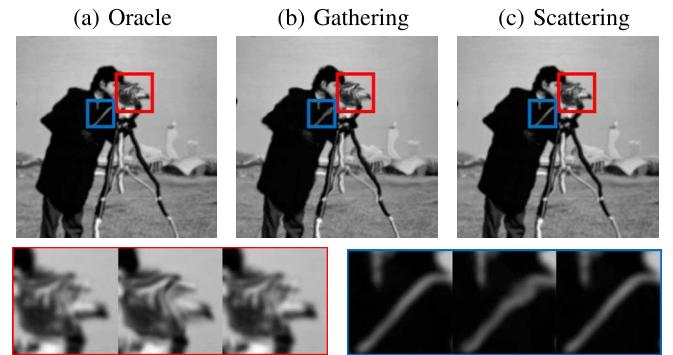


Fig. 13. A visual comparison of the oracle case and scattering and gathering for our toy turbulence model. We can see the oracle case matches scattering, but has some visual difference with gathering.

complexity. To achieve this, given that we have access to all the PSFs, we can take a PCA over the entire dataset:

$$\text{PCA}(\{h_{\mathbf{u}}\}_{\mathbf{u} \in \mathbb{R}^2}) \rightarrow \{\varphi_m, a_{\mathbf{u},m}\}_{m=1,2,\dots,M}. \quad (33)$$

By this, we can encode the information of each PSF into the invariant bases and spatially varying coefficients.

Continuing with our toy turbulence example, we can (tediously) simulate the “oracle” by using a partition per-pixel. We regard this as the correct result. We then apply the scattering and gathering approximations, facilitated by PCA, and see how they compare to the oracle. The comparison is shown in Fig. 13 where we can see that scattering (product-convolution) matches the oracle while gathering (convolution-product) has some errors, particularly near boundaries.

## V. DISCUSSIONS

As we write this paper, two underpinning questions are constantly asked: (1) what utility does it bring? (2) How does it affect how we solve an inverse problem? In this section, we briefly share our findings.

### A. Which Decomposition Should I Choose?

In the examples presented, we have considered both partition and modal interpretations. We consider the following as guidelines regarding which to use:

- 1) **Partition-like decomposition:** Suitable for a small amount of invariant kernels across the image.

2) **Modal decomposition:** Appropriate for systems which require a PSF per-pixel.

The partitioning approach can represent the scene with no error. To illustrate this, consider the superposition integral (2). Since we can partition this any way we like, we can partition it in a way such that each partition corresponds to an invariant convolution. For partitions  $\{P_m\}_{m=1,2,\dots,M}$  we can write

$$\int_{\mathbb{R}^2} h(\mathbf{x}, \mathbf{u}) J(\mathbf{u}) d\mathbf{u} = \sum_{m=1}^M \int_{P_m} h_m(\mathbf{x} - \mathbf{u}) J(\mathbf{u}) d\mathbf{u} \quad (34)$$

where  $h_m(\mathbf{x} - \mathbf{u})$  is the kernel associated with  $P_m$ . We utilized partition-like decompositions in Sections IV-A and IV-B.

Although we can consider a per-pixel partition (as done for the oracle of Section IV-C), it is often computationally infeasible. In these cases it is often preferable to represent each kernel as a sum of bases, such as a modal representation (e.g. PCA). Having previously discussed the computational benefit of a PCA decomposition, we introduce another interesting question, where can error arise in partition-like or modal decompositions? We highlight two errors which can arise in either case:

- 1) **Using a sub-optimal basis:** If one uses a poor basis, there may be errors representing each  $h(\mathbf{x}, \mathbf{u})$  faithfully.
- 2) **Using the wrong model:** If one uses the wrong model the result will be wrong. The degree of mismatch is related to the manner in which the kernel changes.

First, we illustrate how a sub-optimal basis can impact our representation. For non-negative functions  $J(\mathbf{u})$  and  $h(\mathbf{x}, \mathbf{u})$ , we can compute the relative approximation error for a scattering representation as follows:

$$\begin{aligned} \mathcal{L} &= \left| \int_{\mathbb{R}^2} J(\mathbf{u}) \left( h(\mathbf{x}, \mathbf{u}) - \sum_m a_{\mathbf{u},m} \varphi_m(\mathbf{u} - \mathbf{x}) \right) d\mathbf{u} \right| \\ &\leq \int_{\mathbb{R}^2} J(\mathbf{u}) \left| h(\mathbf{x}, \mathbf{u}) - \sum_m a_{\mathbf{u},m} \varphi_m(\mathbf{u} - \mathbf{x}) \right| d\mathbf{u}, \end{aligned} \quad (35)$$

with a similar result for a gathering approximation. Inspecting the difference term in (35), we note this is just the error in basis representation. Thus, the upper bound of the error is related to the basis representation error.

For a partition-like decomposition, this error arises from choosing the wrong function  $\varphi_m$  for partition  $P_m$ . This may stem from approximating a slowly varying kernel as an invariant one over a partition. For a modal decomposition, each kernel must be captured by the basis decomposition. We note that if all kernels come from the same distribution, a bound can likely be placed on the decomposition and one can say with confidence that the approximation holds.

The other noted listed possible source of error is incorrectly choosing scattering or gathering. Although we have already discussed their mutual exclusivity, consider their equality:

$$\sum_{m=1}^M a_{\mathbf{x},m} (\varphi_m * J)(\mathbf{x}) \stackrel{?}{=} \sum_{m=1}^M \left( \varphi_m * (a_m \odot J) \right)(\mathbf{x}). \quad (36)$$

Are there any conditions which scattering and gathering are approximately equal? To illustrate a possibility, we provide a version of the example from Section IV-A in Fig. 14, where

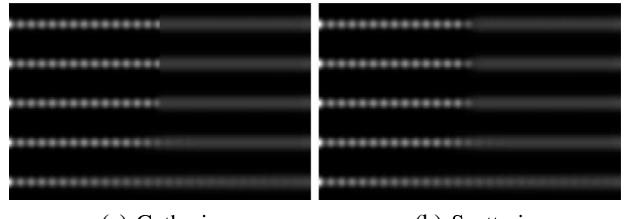


Fig. 14. Visualizing when gathering can approximate scattering. Each row increasingly considers a more slowly changing PSF. Although we do not show it here, the error between the two outputs shown monotonically decreases.

TABLE I  
COMPARISON ON THE TURBULENCE-TEXT DATASET

Model	Simulator in [31]	New Simulator
	CRNN / DAN / ASTER	CRNN / DAN / ASTER
TSRGAN	57.75 / 71.45 / 73.10	60.30 / 73.90 / 74.40
ESTRNN	84.50 / 96.25 / 95.45	87.10 / 97.80 / 96.95
TMT	79.25 / 85.20 / 88.00	80.90 / 87.25 / 88.55

the kernel becomes increasingly invariant as we move toward the bottom of the five rows. The top row corresponds to a sharp transition (equivalent to a partition), while each increasing row uses a more gradual transition.

As the coefficients become nearly invariant over the support of the kernels, the error in choosing the wrong model lessens. Mathematically, as the coefficients approach approximate invariance (relative to kernel support size) we can justify factoring out the coefficients  $a_m$  on the RHS of (36). Thus, the two become increasingly close, with them being exact when we achieve effectively invariant convolution.

The choice in a partition or modal decomposition and the source of errors is highly contextual, depending on various characteristics of the problem. Furthermore, there is nothing restricting us from blending approximations (partition and modal) and scattering and gathering depending on the problem. One such example is adaptively choosing the weights in a blended partition and modal fashion [27]. Thus, researchers have some flexibility in choosing the proper approximation for their problem.

### B. Utility: Simulating Atmospheric Turbulence

Scattering can be applied to the simulation of atmospheric turbulence. The latest turbulence simulators, based on phase-over-aperture [30], phase-to-space (P2S) transform [25], and dense-field P2S (DF-P2S) [31], are all using the *gathering* equation, having some mismatch with nature.

Table I shows a comparison between the DF-P2S simulator [31], and a new simulator implemented using the scattering equation.<sup>1</sup> Our testing dataset is based on the text recognition dataset released in the UG2+ challenge [32]. We tested three

<sup>1</sup>The new simulator contains a few other modifications including expanding the parameter space, and expanding the kernel supports. However, the biggest change is the adoption of the scattering equation.

image reconstruction models: TSRWGAN [33], ESTRNN [34], and TMT [35]. We report the recognition accuracy in terms of CRNN / DAN / ASTER from the restored images.

As we can see in this table, the new simulator indeed offers a sizable amount of improvement over the previous simulator. Since the change from the gathering to the scattering equation attributes to a substantial portion of the simulator update, the utility of our study is evident.

### C. We Have an Inverse Problem, Which Model to Use?

Solving an inverse often requires an optimization. For a spatially varying blur problem, the typical formulation is

$$\hat{\mathbf{J}} = \underset{\mathbf{J}}{\operatorname{argmin}} \|\mathbf{I} - \mathbf{H}\mathbf{J}\|^2 + \lambda R(\mathbf{J}), \quad (37)$$

for some regularization functions  $R(\mathbf{J})$ . Since we know that the scattering formulation applies to image formation, we can instead write this as

$$\hat{\mathbf{J}} = \underset{\mathbf{J}}{\operatorname{argmin}} \left\| \mathbf{I} - \left( \sum_{m=1}^M \mathbf{H}_m \mathbf{D}_m^{(s)} \right) \mathbf{J} \right\|^2 + \lambda R(\mathbf{J}). \quad (38)$$

This problem is extremely difficult to solve. Even if  $\mathbf{D}_m^u$  is binary, we cannot solve for individual regions and combine them. Furthermore, the blurs  $\mathbf{H}_m$  are summed which further complicates the problem. Although we accurately describe the forward model, the inverse problem becomes challenging.

On the other hand, the gathering equation will give us

$$\hat{\mathbf{J}} = \underset{\mathbf{J}}{\operatorname{argmin}} \left\| \mathbf{I} - \left( \sum_{m=1}^M \mathbf{D}_m^{(g)} \mathbf{H}_m \right) \mathbf{J} \right\|^2 + \lambda R(\mathbf{J}). \quad (39)$$

If  $\mathbf{D}_m^{(g)}$  is binary (as in Nagy and O’Leary [22]), we can partition the image into  $M$  smaller regions and solve them individually or through overlapping partitions such as in [36] or by recurrent neural networks with spatially varying weights as in [37]. The parallelism offered by the model is computationally appealing, and it has been proven useful. The caveat is that there is a model mismatch, and thus is a proxy to the original problem. However, if one can justify an assumption of approximate spatial invariance, one can utilize the fact that gathering may approximate scattering. In a general sense for arbitrary  $\mathbf{D}_m^{(g)}$ , any solution we obtain, regardless if we can prove convergence of the optimization algorithm, will not resemble the true solution to (37).

Our advice to practitioners, when solving an inverse problem related to the spatially varying blur, is to have the sense of awareness for this kind of mismatch. For deep neural networks, the mismatch is often less of an issue when the capacity of the network is large enough. However, the training data needs to capture enough of the physics in order to generalize well.

## VI. CONCLUSION

The approximations of gathering and scattering for a spatially varying blur were discussed in this work. They are mutually exclusive, in the sense that if one is the exact representation

of the original blur, the other one can only be an approximation. They become identical if the underlying blur kernel is spatially invariant. In summary, we recognize the following key points:

**Gathering has an origin of image filtering.** It is an effective approach to speed up the spatially varying blur via a small set of invariant blurs. The approach is to filter the image first, and then combine the filtered images through a pixelwise mask. Gathering offers better edge awareness in tasks such as image denoising.

**Scattering is originated from light propagation physics.** It is an accurate description according to the scalar diffraction theory in Fourier optics. The approach is to weigh the images and then perform filtering afterward. Scattering is the model for describing how light propagates through a random medium.

## ACKNOWLEDGMENT

The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of IARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation therein.

## REFERENCES

- [1] S. H. Chan, “Constructing a sparse convolution matrix for shift varying image restoration problems,” in *Proc. IEEE Int. Conf. Image Process.*, 2010, pp. 3601–3604.
- [2] A. Chakrabarti, T. Zickler, and W. T. Freeman, “Analyzing spatially-varying blur,” in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2010, pp. 2512–2519.
- [3] X. Zhu, S. Cohen, S. Schiller, and P. Milanfar, “Estimating spatially varying defocus blur from a single image,” *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4879–4891, Dec. 2013.
- [4] Y. Huang, É. Chouzenoux, and V. Elvira, “Probabilistic modeling and inference for sequential space-varying blur identification,” *IEEE Trans. Comput. Imag.*, vol. 7, pp. 531–546, 2021.
- [5] M. M. Sondhi, “Image restoration: The removal of spatially invariant degradations,” *Proc. IEEE*, vol. 60, no. 7, pp. 842–853, Jul. 1972.
- [6] M. Suin, K. Purohit, and A. N. Rajagopalan, “Degradation aware approach to image restoration using knowledge distillation,” *IEEE J. Sel. Topics. Signal Process.*, vol. 15, no. 2, pp. 162–173, Feb. 2021.
- [7] D. O’Connor and L. Vandenberghe, “Total variation image deblurring with space-varying kernel,” *Comput. Optim. Appl.*, vol. 67, no. 3, pp. 521–541, Jul. 2017.
- [8] B. Saleh and N. Subotic, “Time-variant filtering of signals in the mixed time frequency domain,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 6, pp. 1479–1485, Dec. 1985.
- [9] I. Sodagar, K. Nayebi, and T. P. Barnwell, “Time-varying filter banks and wavelets,” *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 2983–2996, Nov. 1994.
- [10] I. Sodagar, K. Nayebi, T. P. Barnwell, and M. J. T. Smith, “Time-varying analysis-synthesis systems based on filter banks and post filtering,” *IEEE Trans. Signal Process.*, vol. 43, no. 11, pp. 2512–2524, Nov. 1995.
- [11] F. Wang and V. Balakrishnan, “Robust Kalman filters for linear time-varying systems with stochastic parametric uncertainties,” *IEEE Trans. Signal Process.*, vol. 50, no. 4, pp. 803–813, Apr. 2002.
- [12] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, “A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking,” *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.
- [13] G. Wang, “The most general time-varying filter bank and time-varying lapped transforms,” *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3775–3789, Oct. 2006.

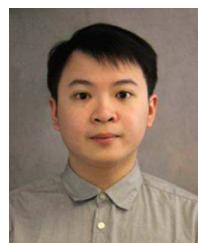
- [14] T. Lauer, "Deconvolution with a spatially variant PSF," in *Astronomical Data Analysis II*, vol. 4847, J.-L. Starck and F. D. Murtagh, Eds., Bellingham, WA, USA: SPIE, 2002, pp. 167–173.
- [15] R. C. Flicker and F. J. Rigaut, "Anisoplanatic deconvolution of adaptive optics images," *J. Opt. Soc. Am. A*, vol. 22, no. 3, pp. 504–513, Mar. 2005.
- [16] M. Hirsch, S. Sra, B. Schölkopf, and S. Harmeling, "Efficient filter flow for space-variant multiframe blind deconvolution," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2010, pp. 607–614.
- [17] M. Arivogvindan, J. Shaevitz, J. McGowan, J. W. Sedat, and D. A. Agard, "A parallel product-convolution approach for representing depth varying point spread functions in 3D widefield microscopy based on principal component analysis," *Opt. Express*, vol. 18, no. 7, pp. 6461–6476, Mar. 2010.
- [18] T. W. Mitchel, B. Brown, D. Koller, T. Weyrich, S. Rusinkiewicz, and M. Kazhdan, "Efficient spatially adaptive convolution and correlation," 2020, *arXiv:2006.13788*.
- [19] K. Yanny, K. Monakhova, R. W. Shuai, and L. Waller, "Deep learning for fast spatially varying deconvolution," *Optica*, vol. 9, no. 1, pp. 96–99, Jan. 2022.
- [20] R. C. Busby and H. A. Smith, "Product-convolution operators and mixed-norm spaces," *Trans. Amer. Math. Soc.*, vol. 263, no. 2, pp. 309–341, 1981.
- [21] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 9, pp. 891–906, Sep. 1991.
- [22] J. G. Nagy and D. P. O'Leary, "Restoring images degraded by spatially variant blur," *SIAM J. Sci. Comput.*, vol. 19, no. 4, pp. 1063–1082, 1998.
- [23] T. Popkin, A. Cavallaro, and D. Hands, "Accurate and efficient method for smoothly space-variant Gaussian blurring," *IEEE Trans. Image Process.*, vol. 19, no. 5, pp. 1362–1370, May 2010.
- [24] A. V. Oppenheim, A. S. Willsky, and S. H. Nawab, *Signals & Systems*, 2nd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 1996.
- [25] Z. Mao, N. Chmitt, and S. H. Chan, "Accelerating atmospheric turbulence simulation via learned phase-to-space transform," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 14759–14768.
- [26] P. Escande and P. Weiss, "Approximation of integral operators using product-convolution expansions," *J. Math. Imaging Vis.*, vol. 58, no. 3, pp. 333–348, Jul. 2017.
- [27] N. Alger, V. Rao, A. Myers, T. Bui-Thanh, and O. Ghattas, "Scalable matrix-free adaptive product-convolution approximation for locally translation-invariant operators," *SIAM J. Sci. Comput.*, vol. 41, no. 4, pp. A2296–A2328, 2019.
- [28] J. W. Goodman, *Introduction to Fourier Optics*, 3rd ed. Englewood, CO, USA: Roberts and Company, 2005.
- [29] S. H. Chan and N. Chmitt, "Computational imaging through atmospheric turbulence," *Found. Trends® Comput. Graph. Vis.*, vol. 15, no. 4, pp. 253–508, 2023.
- [30] N. Chmitt and S. H. Chan, "Simulating anisoplanatic turbulence by sampling intermodal and spatially correlated Zernike coefficients," *Opt. Eng.*, vol. 59, no. 8, pp. 1–26, 2020.
- [31] N. Chmitt, X. Zhang, Z. Mao, and S. H. Chan, "Real-time dense field phase-to-space simulation of imaging through atmospheric turbulence," *IEEE Trans. Comput. Imag.*, vol. 8, pp. 1159–1169, 2022.
- [32] "Bridging the gap between computational photography and visual recognition: 5th UG2+ prize challenge," in *Proc. Comput. Vis. Pattern Recognit. Conf. (CVPR)*, 2022, p. 1. Accessed: Jan. 3, 2023. [Online]. Available: [http://cvpr2022.ug2challenge.org/dataset22\\_t3.html](http://cvpr2022.ug2challenge.org/dataset22_t3.html)
- [33] D. Jin et al., "Neutralizing the impact of atmospheric turbulence on complex scene imaging via deep learning," *Nature Mach. Intell.*, vol. 3, pp. 876–884, Oct. 2021.
- [34] Z. Zhong, Y. Gao, Y. Zheng, B. Zheng, and I. Sato, "Real-world video deblurring: A benchmark dataset and an efficient recurrent neural network," *Int. J. Comput. Vis.*, pp. 1–18, Jan. 2023.
- [35] X. Zhang, Z. Mao, N. Chmitt, and S. H. Chan, "Imaging through the atmosphere using turbulence mitigation transformer," 2022, *arXiv:2207.06465*.
- [36] M. Fornasier, A. Langer, and C.-B. Schönlieb, "A convergent overlapping domain decomposition method for total variation minimization," *Numer. Math.*, vol. 116, no. 4, pp. 645–685, Oct. 2010.
- [37] W. Ren et al., "Deblurring dynamic scenes via spatially varying recurrent neural networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 8, pp. 3974–3987, Aug. 2022.



**Nicholas Chmitt** (Member, IEEE) received the B.Eng. degree in electrical and computer engineering from the Purdue University Northwest, Westville, IN, USA, in 2017, and the Ph.D. degree in electrical engineering from the Purdue University, West Lafayette, IN, USA, in 2023. He is currently a Postdoctoral Researcher. His research works have been published at top computer vision conferences such as CVPR/ICCV, and journals such as IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING and SPIE Optical Engineering. He is a Co-Author of the book titled *Computational Imaging through Atmospheric Turbulence* (Now Publisher, 2023). His research interests include computational imaging and imaging through atmospheric turbulence.



**Xingguang Zhang** (Graduate Student Member, IEEE) received the B.E. degree in optoelectronic information science and engineering from Zhejiang University, in 2018, and the M.S. degree in electrical and computer engineering from Purdue University, in 2021. He joined the i2Lab in 2021 as a Ph.D. Student. His research interests include turbulence modeling and mitigation, generative models, domain adaptation, high dynamic range imaging, and machine learning for video restoration. His research works have been published at top computer vision conferences, such as CVPR/ICCV, and journals, such as IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING. He co-organized the UG2+ workshop at the CVPR 2023 and 2024, he also serves as a Reviewer for international conferences and journals, including ICPR, ICME, ECCV, and IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE.



**Yiheng Chi** (Student Member, IEEE) received the B.S. degree in computer engineering from the Purdue University, West Lafayette, IN, USA, in 2018. 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, IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I: REGULAR PAPERS, Signal Processing: Image Communication, and Optics Express.



**Stanley H. Chan** (Senior Member, IEEE) received the B.Eng. degree in electrical engineering from the University of Hong Kong, in 2007, the M.A. degree in mathematics, and the Ph.D. degree in electrical engineering from the University of California, San Diego, in 2009 and 2011, respectively. From 2012 to 2014, he was a Postdoctoral Research Fellow with the Harvard University. He joined the Purdue University, West Lafayette, IN, USA, in 2014, where he is currently an Elmore Associate Professor of electrical and computer engineering. He is a Co-Author of the book titled *Computational Imaging through Atmospheric Turbulence* (Now Publisher, 2023). His research interests include single-photon imaging, imaging through atmospheric turbulence, and computational photography. He is a recipient of the IEEE Signal Processing Society Best Paper Award 2022 and the IEEE International Conference on Image Processing Best Paper Award 2016. At Purdue, he is a recipient of multiple teaching awards. He is the author of a popular undergraduate textbook titled *Introduction to Probability for Data Science* (Michigan Publishing, 2021). He is currently serving as a Senior Area Editor of IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING (AE 2018–2022, SAE 2022–now), and a Senior Area Editor of IEEE Open Journal on Signal Processing (2022–now). He is an Advising Member of the IEEE Signal Processing Society Technical Committee in Computational Imaging. He was an Elected Member of the CI-TC in 2015–2020 and an Elected Member of the IVMSP-TC in 2023–2025.