

Clutter rejection in passive non-line-of-sight imaging via blind multispectral unmixing

CONNOR HASHEMI,^{*} **D** RAFAEL AVELAR, AND JAMES LEGER **D**

Department of Electrical and Computer Engineering, University of Minnesota, 200 Union St. SE, Minneapolis, MN 55455, USA *hashe037@umn.edu

Abstract: Passive non-line-of-sight imaging methods that utilize scattered light to "look around corners" are often hindered by unwanted sources that overwhelm the weaker desired signal. Recent approaches to mitigate these "clutter" sources have exploited dependencies in the *spectral content*, or color, of the scattered light. A particularly successful method utilized blind source separation methods to isolate the desired imaging signal with minimal prior information. This current paper quantifies the efficacy of several preconditioning and unmixing algorithms when blind source separation methods are employed for passive multispectral non-line-of-sight imaging. Using an OLED television monitor as the source of both the desired signals and clutter, we conducted multiple controlled experiments to test these methods under a variety of scene conditions. We conclude that the preconditioner is a vital component as it greatly decreases the power and correlation of the clutter. Additionally, the choice of unmixing algorithm significantly impacts the reconstruction quality. By optimizing these two components, we find that effective image retrieval can be obtained even when the clutter signals are as much as 670 times stronger than the desired image.

© 2024 Optica Publishing Group under the terms of the Optica Open Access Publishing Agreement

1. Introduction

Passive non-line-of-sight (NLOS) imaging utilizes light scattered from rough surfaces to retrieve information hidden or occluded from the observer. Such imaging scenarios include objects located in a hidden hallway or around a street corner [1–12]. These imaging techniques have many applications, such as quickly locating survivors in a burning building during search-and-rescue or discerning hidden pedestrians and cars in autonomous driving. Several passive NLOS imaging methods have been proposed over the last several years, including occlusion-based methods [1–4], light field methods [5–8], and thermal imaging methods [9–12]. However, most realistic scenarios contain a large number of unwanted signals which negatively impact reconstruction of the desired scene. For example, the undesired ambient light in Ref. [2] was one thousand times stronger than the desired scattered light from the hidden human subjects. We refer to these undesired signals as "clutter" rather than "noise" since they have significant structure and therefore cannot be mitigated by stochastic denoising techniques.

There have been several attempts to attenuate strong clutter signals. These include incorporating prior information or models of the clutter [1,13,14], utilizing scene movement as a discriminant [2,15], applying slow spatially-varying assumptions [3,4], measuring the angle of the received light [5,7], and using non-visible wavelengths [9-12,16]. Although these clutter rejection methods have achieved some success, there is no single method that 1) requires minimal information about the clutter, 2) can image stationary objects, 3) can remove both fast and slow spatially-varying clutter, and 4) can perform image retrieval while using only visible light.

Recently, we have shown that the *spectral content* of a scene can be used to reject clutter while satisfying all the above criteria [17,18]. In particular, in Ref. [17] we incorporated aspects of blind source separation (BSS) to successfully attenuate clutter signals under several challenging conditions. For example, with just five visible spectral filters, we were able to reconstruct

2133

Optics EXPRESS

images that were 30 times more accurate in the presence of complex clutter that was 15 dB stronger than the desired imaging signal (results shown later in Fig. 4). Additionally, since BSS is "blind," the method required very few priors about the hidden scene. However, adapting BSS to multispectral passive NLOS imaging is not straightforward since passive NLOS signals typically originate from diffusely scattered light in the presence of background lighting that contains a large dynamic range. This scenario differs considerably from applications of BSS in hyperspectral unmixing [19,20]. To address these unique difficulties, Ref. [17] implemented a three-step BSS pipeline consisting of (1) applying a preconditioner to the raw measurements, (2) performing signal unmixing, and (3) discriminating between the clutter and desired signals. Although this work demonstrated the efficacy of applying BSS to the spectral content of scattered light, many unexplored questions remain concerning this process. Understanding the theory and best practices for these methods is essential for future improvements and application.

The current paper explores and quantifies the effectiveness of the first two steps of the three-step pipeline developed in Ref. [17]: preconditioning and unmixing. First, we introduce the general method of multispectral NLOS imaging and describe the benefit of incorporating multiple wavelength measurements. We then examine the BSS pipeline and consider multiple candidate preconditioners and algorithms. Next, we perform laboratory experiments using an OLED television monitor to generate a controlled hidden scene, allowing us to vary the number and strength of the clutter signals and the spectral complexity of the hidden scene with precision. These experiments are designed to quantitatively measure the performance of our various methods. Finally, we provide a general discussion and determine which methods and algorithms are most effective.

2. Multispectral passive NLOS imaging

2.1. Monochrome scattered radiance equation

The scattered radiance equation describes how light from a hidden scene on the left side of Fig. 1(a) is received by a camera on the right side of Fig. 1(a) after scattering off of a flat rough surface. Typically, this scattering is described by a 4-D light field which contains two spatial dimensions ξ , η on the scattering surface and two angular dimensions θ , ϕ eminating from the surface. For simplicity, this paper will just consider a 2-D slice of the 4-D light field which includes the horizontal position ξ and horizontal angle θ , but the extension to the entire 4-D light field is straightforward.

Ignoring wavelength variations λ for now, the scattered radiance equation across ξ and θ is given by

$$l^{\text{scat}}(\xi, \theta^{\text{scat}}) = \int_{\Theta^{\text{inc}}(\xi)} l^{\text{inc}}(\xi, \theta^{\text{inc}}) f(\theta^{\text{inc}}, \theta^{\text{scat}}) \cos \theta^{\text{inc}} d\theta^{\text{inc}},$$
(1)

where $l^{\text{scat}}(\xi, \theta^{\text{scat}})$ is the measured radiance scattered off the surface (*i.e.* the signal recorded by the camera) as a function of spatial scattering location ξ and scattering angle θ^{scat} , $l^{\text{inc}}(\xi, \theta^{\text{inc}})$ is the incident radiance from the hidden scene (*i.e.* the objects to be reconstructed) as a function of spatial scattering location ξ and incident angle θ^{inc} , $\Theta^{\text{inc}}(\xi)$ is the angular region of the hidden scene visible from surface location ξ , and $f(\theta^{\text{inc}}, \theta^{\text{scat}})$ describes the bi-directional reflectance distribution function (BRDF) containing the reflectance properties of the surface as a function of incoming and outgoing angles.

With V total scattering locations ξ , M total scattering angles θ^{scat} , and P total incident angles θ^{inc} , the scattered radiance can be described in discrete terms using a simple forward model:

$$\boldsymbol{l}^{\text{scat}} = \boldsymbol{F}\boldsymbol{l}^{\text{inc}},\tag{2}$$

where $I^{\text{scat}} \in \mathbb{R}^{(V \cdot M) \times 1}$ is the lexicographically-scanned scattered radiance vector, $I^{\text{inc}} \in \mathbb{R}^{(V \cdot P) \times 1}$ is the lexicographically-scanned incident radiance vector, and $F \in \mathbb{R}^{(V \cdot M) \times (V \cdot P)}$ is the forward



Fig. 1. (a) depicts our NLOS imaging setup with associated variables; (b) depicts a cluttered occlusion-based imaging example.

operator that describes the scattering process. In passive NLOS imaging it is typically assumed F is known or calculated beforehand.

2.2. Occlusion-based imaging and clutter

Since the BRDF of many common materials is approximately Lambertian in visible light (i.e. $f(\theta^{\text{inc}}, \theta^{\text{scat}}) \approx \rho/\pi$ where ρ is the reflectivity of the surface), it is often insufficient to rely solely on the BRDF to solve for the incident light l^{inc} . Occluders in the scene (such as the occluding edge of the wall in Fig. 1(a)) can cast "penumbras" or "shadows" which create a spatially-varying incident domain $\Theta(\xi)$ by revealing or occluding different parts of the hidden scene across different ξ locations. This improves the condition number of matrix F in Eq. (2) to the extent that it can often be inverted to solve for the hidden scene l^{inc} . One possible inversion is given by the equation

$$\boldsymbol{l}^{\text{inc}} = F^+ \boldsymbol{l}^{\text{scat}},\tag{3}$$

where + signifies the Moore-Penrose inverse. Utilizing occluders to image the hidden scene is called occlusion-based imaging and has been shown to be very successful across a variety of occluders and scattering surfaces [1-4].

Often, only a portion of the hidden scene casts information-rich shadows. This part of the hidden scene (shown in green in Fig. 1(a)) is called the computational field-of-view (CFOV) [3]. Any part of the hidden scene that lies outside the CFOV is not affected by the occluder and thus cannot be reconstructed by occlusion-based imaging. These objects and their associated signals are known as clutter. Since the forward model in Eq. (2) is linear, we can describe the measured scattered radiance as a collection of CFOV and clutter signals:

$$l^{\text{scat}} = l^{\text{cfov}} + l^{\text{clut}},\tag{4}$$

where $l^{cfov} \in \mathbb{R}^{(V \cdot M) \times 1}$ contains the scattered radiance derived from the CFOV region while $l^{clut} \in \mathbb{R}^{(V \cdot M) \times 1}$ contains the scattered radiance derived from the clutter (*i.e.* not from the CFOV region). If the clutter signals are not entirely orthogonal to the occlusion-based forward model,

or $F^+l^{\text{clut}} \neq 0$, then the clutter signals can negatively impact the reconstructions of the CFOV scene. This is particularly problematic in practical situations where the clutter can be orders of magnitude stronger than the CFOV signals. As mentioned in the introduction, *the main goal of this work* is to mitigate I^{clut} under the conditions of stationary scenes at visible wavelengths while making minimal assumptions about the clutter. A short depiction of cluttered occlusion-based imaging is shown in Fig. 1(b).

2.3. Multispectral Linear Mixture

Most passive NLOS imaging methods use the forward model shown in Eq. (2) which implicitly assumes that the light is composed of a single-wavelength. We hypothesize that incorporating wavelength information into Eq. (2) will aid in the clutter removal process. If we assume that 1) the scattering surface is spectrally white, 2) the scattering surface is rough, and 3) the wavelength range is limited to the visible domain, then the multispectral scattered light can be described as a linear mixture of K uniformly-colored objects in the hidden scene:

$$L^{\text{scat}} = \Lambda S^{\text{scat}},\tag{5}$$

where $L^{\text{scat}} \in \mathbb{R}^{N \times (V \cdot M)}$ is the multispectral scattered radiance matrix with N total wavelength measurements, $\Lambda \in \mathbb{R}^{N \times K}$ contains the spectral information for the K uniformly-colored scene objects, and $S^{\text{scat}} \in \mathbb{R}^{K \times (V \cdot M)}$ contains the lexicographically-scanned scattered radiance distribution vectors of each object that is independent of wavelength, where $S_k^{\text{scat}} = F I_k^{\text{inc}}$ and I_k^{inc} is the incident radiance from object k [17]. Note that the number of uniformly-colored objects K can be derived from tools such as k-means clustering of the scene's spectral content.

Equation (5) represents a *linear mixture*, where each row of L^{scat} measures a different combination of the rows from the "source matrix" S^{scat} , each weighted according to the "mixing matrix" Λ [21]. It is important to realize that the mixture in Eq. (5) contains *both* the clutter and CFOV radiance described in Eq. (4):

$$L^{\text{scat}} = \Lambda^{\text{cfov}} S^{\text{cfov}} + \Lambda^{\text{clut}} S^{\text{clut}}, \tag{6}$$

where Λ^{cfov} , S^{cfov} are the CFOV components while Λ^{clut} , S^{clut} are the clutter components. In comparison to the monochrome signal described by Eq. (4), the multispectral mixture in Eq. (6) contains some redundancy which can be used to "unmix" $\Lambda^{cfov}S^{cfov}$ from $\Lambda^{clut}S^{clut}$. This is the essential idea behind using multiple wavelengths for clutter rejection in NLOS imaging. Figure 2 depicts the multispectral mixture; more details about the derivation can be found in the appendix and in Ref. [17].



Fig. 2. Depiction of a multispectral linear mixture. (a) measurement; (b) linear mixture in Eq. (5); (c) separation of CFOV and clutter objects in Eq. (6).

3. Preconditioners and unmixing algorithms

3.1. Blind Source Separation (BSS) Models

Blind source separation models represent a class of techniques that can be used to "unmix" the linear mixture in Eq. (5) into an estimated mixing matrix $\tilde{\Lambda}$ and source matrix \tilde{S}^{scat} [21,22]. It assumes minimal prior information about the mixture - hence the term "blind." As mentioned in the introduction, there are many challenges in adapting BSS to the multispectral passive NLOS imaging mixture in Eq. (5) and Eq. (6). To make the problem more manageable, we focus on extracting only S^{cfov} in Eq. (6) since it contains the most useful information about the CFOV objects such as their shape and location. Previously in Ref. [17] we utilized a 3-step approach to extract S^{cfov} given by 1) preconditioning, 2) unmixing, and 3) clutter rejection. First, the measurements undergo linear "preconditioning" which converts the measurements into a new domain that is numerically easier to unmix. Second, an unmixing algorithm estimates the predicted source matrix \tilde{S}^{scat} with prior knowledge of the number of uniformly-colored objects K (note that this can also be estimated "blindly"). Third, some simple criteria is used to determine if the scattering distribution of object k belongs to S^{cfov} or S^{clut} . Finally, S^{cfov} is reconstructed while S^{clut} is discarded. The BSS pipeline is depicted in Fig. 3 while Fig. 4 depicts an experiment performed in Ref. [17] with five visible spectral filters and clutter light bulbs that were 15 dB stronger than the CFOV objects. Note that the CFOV objects (i.e. book and can) are not of uniform color yet can still be reconstructed via BSS.



Fig. 3. Depiction of the BSS model described in Section 3.1. This paper focuses on parts a-c which include aspects of the preconditioning transforms and unmixing algorithms.

In the BSS model there are many valid choices for the preconditioners (step 1) and unmixing algorithms (step 2) that are unexplored. There also are other criteria that can be considered for the clutter rejection (step 3), such as rejecting objects based on spectral content or prior knowledge of clutter structure. However, this paper will not focus on the clutter rejection step since it is more straightforward and depends on the specific application.

3.2. Preconditioning

The first step of the BSS pipeline, preconditioning, defines a linear transform $\mathcal{P}\{\cdot\}$ that converts the scattered radiance measurements L^{scat} into "preconditioned" measurements $\mathcal{P}\{L^{\text{scat}}\}$ that are numerically easier to unmix. While the performance metric of $\mathcal{P}\{\cdot\}$ depends on the particular



Fig. 4. Results from the supplementary material of Ref. [17] using a setup similar to Fig. 1 and Fig. 3 with real hidden objects and visible spectral filters. Clutter sources are three light bulbs: two "soft-white" and one "daylight" temperature profiles. Note the reconstructions have vertical-uniformity due to the 1-D wall edge occluder.

unmixing algorithm used in step 2, in general we expect that decreasing the correlation between the scattering distributions in S^{scat} and decreasing the power of the clutter objects will improve the reconstruction results. This paper explores several different preconditioners $\mathcal{P}\{\cdot\}$ for the multispectral scattered light.

3.2.1. Spatial differential

$$\mathcal{P}_{\text{diff}}\{L^{\text{scat}}\} = D_{\xi}L^{\text{scat}},\tag{7}$$

where $D_{\xi} \in \mathbb{R}^{(V \cdot M) \times (V \cdot M)}$ performs the differential across ξ . Many occlusion-based imaging methods utilize $\mathcal{P}_{\text{diff}}$ since the penumbras from CFOV objects have sharp gradients while unobstructed ambient light has a smooth profile [2–4]. As a result, $\mathcal{P}_{\text{diff}}$ minimizes the power of slow spatially-varying clutter objects while maximizing the power of the CFOV objects.

3.2.2. Truncated spatial differential

$$\mathcal{P}_{\text{diff}}^{(Q)}\{L^{\text{scat}}\} = \sum_{i=1}^{Q} \frac{1}{\sigma_i} D_{\xi} L^{\text{scat}} \boldsymbol{u}_i \boldsymbol{u}_i^T,$$
(8)

where σ_i is the *i*th singular value and $u_i \in \mathbb{R}^{(V \cdot M) \times 1}$ is the *i*th right-singular vector of the matrix describing the discretized BRDF. Previously used in Ref. [5], $\mathcal{P}_{diff}^{(Q)}$ is unique to light field NLOS imaging since it performs best with non-Lambertian BRDFs in which θ^{scat} carries significant information. Compared to \mathcal{P}_{diff} , $\mathcal{P}_{diff}^{(Q)}$ is expected to better handle fast spatially-varying clutter sources which have significant parallax. Note that we use Q = 3 to obtain the best results with the scattering surface (white satin paint) used later in the experiments.

3.2.3. Least-squares (LS) reconstruction

$$\mathcal{P}_{\text{recon}}^{(Q)}\{L^{\text{scat}}\} = F^+ \left(\mathcal{P}_{\text{diff}}^{(Q)}\{L^{\text{scat}}\} \right),\tag{9}$$

where + is the Moore-Penrose inverse and *F* is the forward operator previously defined in Eq. (2). This reconstruction is very similar to the least squares inversion described in Eq. (3) with the addition of the previously mentioned preconditioner $\mathcal{P}_{diff}^{(Q)}\{L^{scat}\}$ in Eq. (8). It is also used in occlusion-based light field NLOS imaging in Ref. [5]. Compared to $\mathcal{P}_{diff}^{(Q)}\{L^{scat}\}$, $\mathcal{P}_{recon}^{(Q)}\{L^{scat}\}$ further minimizes the strength of clutter objects whose scattering distributions are near orthogonal to *F*. However, it is important to note that this preconditioner prevents the typical implementation of the third BSS step since it causes the residuals of the resulting \tilde{S}^{scat} to always be zero.

3.2.4. "Optimized Preconditioning"

$$\mathcal{P}_{\text{opt}}\{L^{\text{scat}}\} = \hat{P}L^{\text{scat}} \text{ where } \hat{P} = \begin{vmatrix} I \\ 0 \end{vmatrix} \begin{bmatrix} F & S^{\text{clut}} \end{bmatrix}^+,$$
 (10)

where S^{clut} is the clutter objects' scattering distributions. "Optimized Preconditioning" was previously developed in Ref. [14] to mitigate clutter objects by placing their expected scattering distributions S^{clut} in the nullspace of the preconditioner. Under ideal circumstances where the clutter and forward model are nearly orthogonal, this would completely remove the effects of the clutter. However, this method is highly susceptible to noise and model mismatch. It is important to understand that, unlike the rest of the preconditioners, optimized preconditioning is *not* blind since it requires a measurement or approximation of S^{clut} ahead of time. However, we still include it in this paper to show the best non-blind result.

3.3. Unmixing algorithms

The second step of the BSS pipeline, the unmixing algorithm, separates the preconditioned measurements $\mathcal{P}\{L^{\text{scat}}\}$ into the estimated mixing matrix $\tilde{\Lambda}$ and estimated source matrix \tilde{S}^{scat} in Eq. (5). Since BSS operates under "blind" assumptions, each unmixing method employs a different metric to optimize the unmixing. There are a variety of BSS unmixing techniques, and we test the following popular algorithms.

3.3.1. Principal component analysis (PCA)

PCA is a popular dimensionality-reduction method that produces an ordered orthogonal basis. The first basis element, or principal component $w_1 \in \mathbb{R}^{(V \cdot M) \times 1}$, minimizes the projection error of the scattered radiance observations L^{scat} :

$$\min_{\mathbf{w}_1} ||L^{\text{scat}} - L^{\text{scat}} \mathbf{w}_1 \mathbf{w}_1^T ||_F^T \text{ where } \mathbf{w}_1^T \mathbf{w}_1 = 1.$$
(11)

Equation (11) is repeated in a similar fashion to find the subsequent principal components. For BSS, the first *K* principal components comprise the estimated scattering distributions \tilde{S}^{scat} . While PCA can be thought of as a "naïve" BSS solution since it is not specifically designed for unmixing linear mixtures, it has shown recent success in finding small signals in large biases for passive NLOS imaging [12]. In this paper, we implement PCA using singular value decomposition (SVD) in MATLAB.

3.3.2. Non-Negative Matrix Factorization (NMF)

NMF is a popular BSS technique that assumes both \tilde{S}^{scat} and $\tilde{\Lambda}$ are non-negative and minimizes the Euclidean distance:

$$\min_{\tilde{S}^{\text{scat}},\tilde{\Lambda}} ||L^{\text{scat}} - \tilde{\Lambda}\tilde{S}^{\text{scat}}||_{F}^{2} \text{ where } \tilde{\Lambda} \ge 0, \ \tilde{S}^{\text{scat}} \ge 0.$$
(12)

The biggest disadvantage of NMF is that the evolving solution can become trapped in a local minimum, making the ideal solution difficult to find. To increase reliability, we have use a cascading or multilayer NMF similar to Refs. [20,23] with three layers. We also repeated the algorithm 100 times with different random initialization states and kept the result with the smallest residual error.

3.3.3. Second-order blind identification (SOBI)

Independent component analysis (ICA) algorithms assume that the sources in S^{scat} are "independent" from each other, where each ICA method measures the degree of independence differently

[24]. SOBI approximates independence between sources s^{scat} in \tilde{S}^{scat} by using second-order statistics, namely that the sources are mutually uncorrelated:

$$\operatorname{Cov}(\boldsymbol{s}_i^{\text{scat}}, \boldsymbol{s}_i^{\text{scat}}) = 0 \text{ where } i \neq j.$$
(13)

To solve Eq. (13), SOBI performs joint-diagonalization with many different sample lags [25].

3.3.4. Joint approximate diagonalization of Eigen-matrices (JADE)

JADE [26] is an ICA algorithm that uses higher-order statistics to measure independence. It assumes that the fourth-order cross-cumulants between sources in \tilde{S}^{scat} are all equal to zero:

$$\operatorname{Cum}(\boldsymbol{s}_{i}^{\operatorname{scat}}, \boldsymbol{s}_{i}^{\operatorname{scat}}, \boldsymbol{s}_{k}^{\operatorname{scat}}, \boldsymbol{s}_{l}^{\operatorname{scat}}) = 0 \text{ where } i \neq j \neq k \neq l.$$
(14)

JADE has been shown to produce meaningful results in many fields. It is solved by jointdiagonalization in a similar manner to SOBI. However, because it uses higher-order statistics, it tends to be more robust than the SOBI algorithm.

4. Experiments using a television monitor as a controlled source

In this section we describe an experimental setup that uses a calibrated OLED television monitor to simulate multispectral light from a variety of object and clutter sources. Unlike the demonstration shown in Fig. 4 that used real objects, we generate synthetic light fields from the television monitor to more accurately control the hidden scene parameters and test the limitations our methods under a variety of conditions.

4.1. Experimental setup

Figure 5 depicts the experimental setup. The left side of Fig. 5(a) shows an LG C2 OLED monitor displaying multiple uniformly-colored hidden objects. Two baffles are placed near the OLED monitor to create more complex object scattering distributions and to restrict the CFOV region. The imaging occluder is an opaque steel sheet covered with black felt and is meant to simulate the edge of a standard hallway. The right sides of Fig. 5(a) and Fig. 5(b) show a camera (with lens) that records images of the scattering surface from a variety of angles from 0 to 70 degrees in 1 degree increments (M = 71). The field-of-view of the camera is 10 cm. Note that the iris of the camera lens is reduced in size so that the scattering angle θ^{scat} can be highly resolved for each spatial location ξ on the scattering wall and rotation angle of the camera. The camera images contain 2048x2048 pixels. However, after correcting for distortion and performing spatial averaging, only 271 spatial ξ positions remain (V = 271). The signal-to-noise ratio of the camera is approximately 30 dB. The scattering surface is a white satin-painted surface which has a mix of specular and diffusive BRDF components. Its specular peak can be roughly approximated by a Gaussian with a full-width half-maximum of 20°.



Fig. 5. Depiction of setup used in experiments. Figure modified from Ref. [17].

To generate a multispectral hidden scene, the monochrome intensity of the TV screen was adjusted to correspond to each of the predefined spectral components of the scene, and data was

taken for each spectral component separately. This mimics multispectral measurements using monochrome screen intensities without the need for spectral filters. Figures 5(c) and 5(d) show the two different scene conditions. Figure 5(c) shows two clutter objects and one CFOV object (K = 3) with five total spectral filters (N = 5) while Fig. 5(d) shows only one clutter object and up to five CFOV objects ($2 \le K \le 6$) with six total spectral filters (N = 6). Note that the depicted colors of the objects represent the simulated spectral distributions.

4.2. Results and discussion

4.2.1. Preconditioner metrics

Although it is difficult to quantify the "effectiveness" of a preconditioner, it is reasonable to consider a reduction in the correlation among scattering distributions in S^{scat} and a reduction in the total power of the clutter sources $\Lambda^{\text{clut}}S^{\text{clut}}$ as effective metrics. Figure 6 depicts and compares the preconditioners applied to the screen objects displayed in Fig. 5(c). Without preconditioning (*i.e.* using just light field measurements L^{scat}) the scattered radiance from both clutter objects is considerably brighter than the CFOV object (350 and 10 times brighter) and they have a correlation to the CFOV object of 0.37 and 0.38.



Fig. 6. The effect of the BSS preconditioners in Section 3.2 on the objects in Fig. 5(c). (a) preconditioning depictions. Each column contains a different preconditioner applied to the scattered radiance from each scene object. Note the units change based on the preconditioner; (b) corresponding power ratio and correlation of each clutter element to the CFOV object (lower is better for both metrics).

As depicted in Fig. 6(b), each preconditioner greatly reduces the correlation and power metrics compared to using L^{scat} directly. While each preconditioner minimizes the correlations between the scattering distributions, the "optimized preconditioner" performs the best by reducing

the power ratio of the clutter objects relative to the CFOV objects to 0.7% and 31% of the non-preconditioned (light field) values. However, it is important to remember that "optimized preconditioning" requires prior knowledge of the clutter scattering distributions while the other methods do not. For the "blind" (*i.e.* limited prior knowledge) preconditioners, the "truncated differential" offers the best results by reducing the power ratio of the clutter sources to 17% and 65% of the non-preconditioned values.

4.2.2. Single object reconstructions with varying clutter strength

Figure 7 compares the reconstruction accuracy of the single CFOV object setup in Fig. 5(c) with a varying signal-to-clutter ratio (SCR) and several preconditioner and unmixing algorithms. The "baseline" curve is the spectrally-agnostic "LS reconstruction" in Eq. (9) which is commonly used in light field NLOS imaging [5]. Figure 8 compares the best results.



Fig. 7. Reconstruction accuracy as a function of signal-to-clutter ratio (SCR) for various preconditioner and unmixing algorithms. Screen setup is as shown in Fig. 5(c).



Fig. 8. Best reconstruction error results as a function of SCR for (a) each preconditioner and (b) each unmixing algorithm across all "blind" preconditioners.

In Fig. 8(a) the "LS recon" and "truncated differential" preconditioners performed the best in the single object experiments. This is also evident in Fig. 7 where the errors in (c) and (d) are generally smaller than the rest of the preconditioners while "light field" (*i.e.* no preconditioning) and "spatial differential" perform the poorest. In general, the best performing preconditioners in Fig. 6 were also the best performing preconditioners in the single object reconstructions in

Fig. 7 and Fig. 8. However, the one exception is "optimized preconditioning," which performs the best across the preconditioner metrics yet is only the third-best performer in the single object reconstruction in Fig. 8(a). This is a result of the preconditioner's high susceptibility to noise in the system due to the ill-conditioned inversion in Eq. (10).

Figure 8(b) summarizes the best performances of the unmixing algorithms across all "blind" preconditioners (*i.e.* all except "optimized preconditioning"). Overall, the JADE algorithm performs the best across the majority of the SCRs, followed by SOBI, PCA, and NMF. The performance of NMF and PCA is a bit unexpected. Intuitively, NMF (non-negative matrix factorization) would be the ideal candidate for unmixing multispectral light fields since both Λ and S^{scat} are non-negative (because negative radiance is not physical). The biggest reason NMF's performance as evident in Fig. 6 and Fig. 7. Since most of the preconditioners offers weaker performance as evident in Fig. 6 and Fig. 7. Since most of the preconditioners are not guaranteed to have non-negative constraints, this removes the main advantage of utilizing NMF. PCA, while not specifically designed for linear unmixing, is able to achieve surprisingly impressive results (see Fig. 8(b)) while being one of the least expensive algorithms. One reason for PCA's success is that it can extract weak signals across large biases. This is particularly well-suited to passive NLOS imaging which typically has a large dynamic range across the CFOV and clutter objects. This also can explain why the performance of PCA improves as the SCR becomes smaller in Fig. 8(b).

4.2.3. Spectral overlap in single object reconstructions

While the SCR is one important metric to gauge the difficulty of BSS, another important metric is the spectral overlap between the scene objects. If this overlap is large, it is much more difficult to utilize the differences in the spectral content to separate the objects in the scene. Figure 9 compares the performance of the unmixing algorithms at a fixed SCR of 1:40 while the average spectral overlap between the CFOV object and clutter objects in Fig. 5(c) is varied. In this experiment the "truncated differential" preconditioner was used since it had the best overall results in Fig. 8(a). As expected, each algorithm performs best when the spectral overlap is small. However, the JADE algorithm clearly is more robust than the others for most degrees of spectral overlap. Note that, while we tried to keep the SCR constant, it decreased as the overlap increased which is why the baseline performance in blue varies slightly.



Fig. 9. Effects of spectral overlap between CFOV object and clutter objects. (a) "spectral overlap" definition; (b) unmixing algorithm performance.

4.3. Multiple object reconstructions

As the number of hidden CFOV objects increases, the number of total hidden objects K also increases and the linear mixture in Eq. (5) becomes more complex. While this should not affect the efficacy of the preconditioners, it will degrade the performance of the unmixing algorithms. Figures 10 and 11 depict the unmixing results with just a single clutter source (clutter 1 in Fig. 5(a)) and multiple hidden CFOV objects as shown in the screen setup in Fig. 5(d). The

Research Article

Optics EXPRESS

"truncated differential" preconditioner is used and the SCR is set to 1:50 with an average spectral overlap of 0.90. The "baseline" is again the "LS reconstruction" from Eq. (9).



Fig. 10. Reconstructions of multiple CFOV objects in Fig. 5(d) experiment.



Fig. 11. Structural similarity (SSIM) index of reconstructions in Fig. 10 compared to "Clutterless Recon."

As is evident in the reconstructions displayed in Fig. 10, the reconstructed color of the objects is often inaccurate. Note that this color inaccuracy was also noted in our previous work in Ref. [17] and is clearly visible in the results in Fig. 4. This inaccuracy is most likely due to the fact that the BSS model in Section 3.1 is designed to extract the estimated scattering distributions \tilde{S}^{cfov} since they provide the most important information about the hidden scene. As a result, the majority of the methods (*i.e.* the preconditioners and unmixing algorithms) refine the estimate of \tilde{S}^{cfov} . There is no such refinement or priors applied to the estimated spectral content $\tilde{\Lambda}^{cfov}$, resulting in the poor color fidelity displayed in Fig. 10. Furthermore, with multiple CFOV objects, the reconstructed scattering distributions in \tilde{S}^{cfov} overlap and merge with each other. For example, the estimated BSS component for object one actually contains elements of objects two and three. While neither the poor color fidelity nor the distribution overlap affect the single-object reconstructions shown previously, they greatly affect the multi-object reconstructions since the multiple BSS components interfere with each other when combining together in the final reconstruction.

The reconstruction accuracy utilized in previous experiments proved to be insufficient to gauge the reconstruction quality due to the large color inaccuracies in the multi-object experiments. As an alternative metric, we calculated the structural similarity (SSIM) index of each reconstruction compared to the clutterless reconstruction, where a larger SSIM index corresponds to a reconstruction whose structure more accurately matches the clutterless version. The SSIM curves in Fig. 11 roughly correspond to the visual appearance of the reconstructions in Fig. 10, with PCA performing the best and JADE performing the worse with a large number of hidden objects. This

is a surprising result, since JADE performed the best in the previous single-object experiments (also seen as a high SSIM value for one object in the curve of Fig. 11).

5. Conclusion

In this paper we have explored and quantified the efficacy of several preconditioners and unmixing algorithms in blind source separation (BSS) of multispectral non-line-of-sight (NLOS) imagery to reject clutter and improve the reconstruction performance in low-signal scenarios. The use of a television monitor to generate the source light fields allowed us to quantitatively evaluate the performance of single-object and multiple-object reconstructions as a function of several parameters. As evident in our results, preconditioning is a vital step to mitigate many of the difficulties in unmixing NLOS imagery and greatly improves the clutter rejection performance. For blind (*i.e.* no or few priors) applications, the best performing preconditioner was the "truncated differential." In non-blind applications, the "optimized preconditioning" performed better as long as it was paired with a strong denoiser or regularizer to mitigate its susceptibility to noise. From our tests of unmixing algorithms, the independent component analysis (ICA) methods represented by SOBI and JADE performed best and had the most consistent results. In particular, JADE performed better with fewer hidden objects whereas SOBI performed better with more. However, principal component analysis (PCA), which is mainly used for dimensionality-reduction, also performed surprisingly well with multiple hidden objects and could be a less-expensive solution. Finally, recovering the true colors of the hidden objects and preventing overlap between different scattering distributions remain a challenge. These issues require more robust solutions when applying BSS methods to multiple-object reconstructions and more spectrally-complex scenes.

While there are many future research routes, there currently are no methods to refine the estimated spectral content of the hidden scene. One way to improve spectral recovery would be to design the unmixing algorithms based on the unique problems of multispectral NLOS imaging. For example, traditional ICA methods which are designed for general applications could be modified to exploit the high dynamic range and relatively simple scattering distributions present in passive NLOS imaging. In addition, since many clutter sources have a similar spectral content (*e.g.* incandescent light bulbs with similar color temperatures), there are opportunities to utilize spectral priors to identify clutter and improve spectral reconstructions, perhaps in a deep-learning environment similar to Ref. [27].

Appendix: multispectral linear mixture derivation

This section reviews the derivation of the multispectral linear mixture in Eq. (5) that was previously developed in Ref. [17].

The addition of a spectral dimension to the monochrome scattered radiance equation in Eq. (1) results in a multispectral scattered radiance equation:

$$l^{\text{scat}}(\xi, \theta^{\text{scat}}, \lambda) = \int_{\Theta^{\text{inc}}(\xi, \lambda)} l^{\text{inc}}(\xi, \theta^{\text{inc}}, \lambda) f(\theta^{\text{inc}}, \theta^{\text{scat}}, \lambda) \cos \theta^{\text{inc}} d\theta^{\text{inc}},$$
(15)

where $\Theta^{\text{inc}}(\xi, \lambda)$ is the angular domain of the hidden scene as a function of spatial position ξ and wavelength λ , $l^{\text{inc}}(\xi, \theta^{\text{inc}}, \lambda)$ is the multispectral incident radiance, and $f(\theta^{\text{inc}}, \theta^{\text{scat}}, \lambda)$ is the multispectral BRDF.

To simplify Eq. (15), we make the assumption that λ is confined to the visible domain and therefore the occluders are opaque across all wavelengths. This allows us to simplify $\Theta^{\text{inc}}(\xi, \lambda)$ as

$$\Theta^{\rm inc}(\xi,\lambda) = \Theta^{\rm inc}(\xi). \tag{16}$$

To simplify the multispectral BRDF in the visible domain, we use a dichromatic reflection model [28,29] for $f(\theta^{\text{inc}}, \theta^{\text{scat}}, \lambda)$ which separates the BRDF into surface scattering and subsurface

scattering components:

---+

$$f(\theta^{\text{inc}}, \theta^{\text{scat}}, \lambda) = f^{\text{spec}}(\theta^{\text{inc}}, \theta^{\text{scat}}) + \alpha(\lambda) f^{\text{diff}}(\theta^{\text{inc}}, \theta^{\text{scat}}), \tag{17}$$

where f^{spec} is the surface scattering (*i.e.* specular reflection), f^{diff} is the subsurface scattering (*i.e.* diffuse reflection), and $\alpha(\lambda)$ is the diffuse albedo of the material ($\alpha(\lambda) \leq 1$).

The final assumption is that the hidden scene can be represented by a collection of *K* total uniformly-colored objects:

$$l^{\rm inc}(\xi,\theta^{\rm inc},\lambda) = \sum_{k=1}^{K} l_k^{\rm inc}(\xi,\theta^{\rm inc},\lambda) = \sum_{k=1}^{K} l_k s_k^{\rm inc}(\xi,\theta^{\rm inc})\gamma_k(\lambda),\tag{18}$$

where $l_k^{\text{inc}}(\xi, \theta^{\text{inc}}, \lambda)$ is the multispectral light from a uniformly-colored object k, l_k is a constant that represents the brightness of object k, $s_k^{\text{inc}}(\xi, \theta^{\text{inc}})$ is the normalized incident distribution of object k, and $\gamma_k(\lambda)$ is the normalized spectrum of object k. Note that this uniform color assumption can be applied to multi-colored objects by assigning each color in the multicolored object to a different value k. However, for scenes containing objects with a large number of colors, the total number K could become sufficiently large as to degrade the performance of the BSS algorithm.

With the substitutions made in Eqs. (16), (17), and (18), the multispectral scattered radiance can be described as

$$l^{\text{scal}}(\xi, \theta^{\text{scal}}, \lambda) = \sum_{k=1}^{K} \left[l_k \gamma_k(\lambda) \int_{\Theta^{\text{inc}}(\xi)} s_k^{\text{inc}}(\xi, \theta^{\text{inc}}) \left[f^{\text{spec}}(\theta^{\text{inc}}, \theta^{\text{scal}}) + \alpha(\lambda) f^{\text{diff}}(\theta^{\text{inc}}, \theta^{\text{scal}}) \right] \cos \theta^{\text{inc}} d\theta^{\text{inc}} \right].$$
(19)

To simplify Eq. (19), we convert the incident distributions $s_k^{\text{inc}}(\xi, \theta^{\text{inc}})$ to specular $s_k^{\text{spec}}(\xi, \theta^{\text{scat}})$ and diffuse $s_k^{\text{diff}}(\xi, \theta^{\text{scat}})$ scattering distributions by

$$s_{k}^{\text{spec}}(\xi, \theta^{\text{scat}}) = \int_{\Theta^{\text{inc}}(\xi)} s_{k}^{\text{inc}}(\xi, \theta^{\text{inc}}) f^{\text{spec}}(\theta^{\text{inc}}, \theta^{\text{scat}}) \cos \theta^{\text{inc}} d\theta^{\text{inc}}$$

$$s_{k}^{\text{diff}}(\xi, \theta^{\text{scat}}) = \int_{\Theta^{\text{inc}}(\xi)} s_{k}^{\text{inc}}(\xi, \theta^{\text{inc}}) f^{\text{diff}}(\theta^{\text{inc}}, \theta^{\text{scat}}) \cos \theta^{\text{inc}} d\theta^{\text{inc}}.$$
(20)

Substituting Eq. (20) into Eq. (19) leads to the final simplified equation:

$$l^{\text{scat}}(\xi, \theta^{\text{scat}}, \lambda) = \sum_{k=1}^{K} l_k \gamma_k(\lambda) \left[s_k^{\text{spec}}(\xi, \theta^{\text{scat}}) + \alpha(\lambda) s_k^{\text{diff}}(\xi, \theta^{\text{scat}}) \right].$$
(21)

If the scattering surface is spectrally white ($\alpha(\lambda) = 1$), then a discretized version of Eq. (21) simplifies to the multispectral mixture used in Eq. (5). If it is *not* spectrally white, it can still be expressed as a linear mixture with four matrix components $\Lambda^{\text{spec}}S^{\text{spec}} + \Lambda^{\text{diff}}S^{\text{diff}}$ instead of the two components ΛS^{scat} used in Eq. (5).

Disclosures. The authors declare no conflicts of interest.

Data availability. Majority of the data underlying the results presented in this paper are available at [30] and the rest may be obtained from the authors upon reasonable request.

References

 A. Torralba and W. T. Freeman, "Accidental pinhole and pinspeck cameras : revealing the scene outside the picture," in *IEEE Conference on Computer Vision and Pattern Recognition*, (2012), pp. 374–381.

Vol. 32, No. 2/15 Jan 2024 / Optics Express 2146

Research Article

Optics EXPRESS

- K. L. Bouman, V. Ye, A. B. Yedidia, et al., "Turning Corners into Cameras: Principles and Methods," in Proceedings of the IEEE International Conference on Computer Vision, (2017), pp. 2270–2278.
- C. Saunders, J. Murray-Bruce, and V. K. Goyal, "Computational periscopy with an ordinary digital camera," Nature 565(7740), 472–475 (2019).
- S. W. Seidel, Y. Ma, J. Murray-Bruce, et al., "Corner Occluder Computational Periscopy: Estimating a Hidden Scene from a Single Photograph," 2019 IEEE International Conference on Computational Photography (ICCP) pp. 1–9 (2019).
- D. Lin, C. Hashemi, and J. Leger, "Passive non-line-of-sight imaging using plenoptic information," J. Opt. Soc. Am. A 37(4), 540 (2020).
- T. Sasaki and J. R. Leger, "Light-field reconstruction from scattered light using plenoptic data," in *Unconventional* and Indirect Imaging, Image Reconstruction, and Wavefront Sensing 2018, vol. 10772 (International Society for Optics and Photonics, 2018), p. 1077203.
- T. Sasaki and J. Leger, "Light field reconstruction from scattered light using plenoptic data," J. Opt. Soc. Am. A 37(4), 653 (2020).
- T. Sasaki and J. Leger, "Object depth estimation from scattered light using plenoptic data," J. Opt. Soc. Am. A 38(2), 211–228 (2021).
- T. Maeda, Y. Wang, R. Raskar, et al., "Thermal Non-Line-of-Sight Imaging," in 2019 IEEE International Conference on Computational Photography (ICCP), (IEEE, 2019), pp. 1–11.
- M. Kaga, T. Kushida, T. Takatani, *et al.*, "Thermal non-line-of-sight imaging from specular and diffuse reflections," IPSJ Trans. on Comput. Vis. Appl. 11(1), 8–9 (2019).
- T. Sasaki, C. Hashemi, and J. Leger, "Passive 3D location estimation of non-line-of-sight objects from a scattered infrared light field," Opt. Express 29(26), 43642–43661 (2021).
- C. Hashemi, T. Sasaki, and J. Leger, "Parallax-Driven Denoising of Passive Non-Line-of-Sight Thermal Imagery," 2023 IEEE International Conference on Computational Photography (ICCP) pp. 1–12 (2023).
- M. Baradad, V. Ye, A. B. Yedidia, et al., "Inferring Light Fields from Shadows," in *IEEE Conference on Computer Vision and Pattern Recognition*, (IEEE, 2018), pp. 6267–6275.
- C. Saunders and V. K. Goyal, "Fast Computational Periscopy in Challenging Ambient Light Conditions through Optimized Preconditioning," 2021 IEEE International Conference on Computational Photography, ICCP 2021 pp. 1–9 (2021).
- A. B. Yedidia, M. Baradad, C. Thrampoulidis, et al., "Using Unknown Occluders to Recover Hidden Scenes," in Proceedings of the IEEE Computer Vision and Pattern Recognition Conference, (IEEE, 2019), pp. 12231–12239.
- T. Sasaki, E. N. Grossman, and J. R. Leger, "Estimation of the 3D spatial location of non-line-of-sight objects using passive THz plenoptic measurements," Opt. Express 30(23), 41911–41921 (2022).
- C. Hashemi, R. Avelar, and J. Leger, "Isolating Signals in Passive Non-Line-of-Sight Imaging using Spectral Content," IEEE Trans. Pattern Anal. Mach. Intell. PP, 1–12 (2023).
- C. Hashemi and J. Leger, "Exploiting the Visible Spectrum to Look Around Corners," in *Computational Optical Sensing and Imaging*, (Optica Publishing Group, 2020), pp. CTh5C-3.
- W.-K. Ma, J. M. Bioucas-dias, T.-h. Chan, *et al.*, "A signal processing perspective on hyperspectral unmixing: insights from remote sensing," IEEE Signal Process. Mag. 31(1), 67–81 (2013).
- R. Rajabi and H. Ghassemian, "Spectral unmixing of hyperspectral imagery using multilayer NMF," IEEE Geosci. Remote Sensing Lett. 12(1), 38–42 (2015).
- 21. J. F. Cardoso, "Blind signal separation: Statistical principles," Proc. IEEE 86(10), 2009–2025 (1998).
- P. Comon and C. Jutten, Handbook of Blind Source Separation: Independent component analysis and applications (2010).
- A. Cichocki, R. Zdunek, and S. I. Amari, "Hierarchical ALS algorithms for nonnegative matrix and 3D tensor factorization," Lect. Notes Comput. Sci. (including subseries Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 4666 LNCS, 169–176 (2007).
- 24. J. Shlens, "A Tutorial on Independent Component Analysis," arXiv, arXiv:1404.2986 (2014).
- A. Belouchrani, K. Abed-Meraim, J. F. Cardoso, *et al.*, "A blind source separation technique using second-order statistics," IEEE Trans. Signal Process. 45(2), 434–444 (1997).
- 26. J. F. Cardoso and A. Souloumiac, "Blind beamforming for non-Gaussian signals," (1993).
- 27. F. Bao, X. Wang, S. H. Sureshbabu, et al., "Heat-assisted detection and ranging," Nature 619(7971), 743–748 (2023).
- 28. S. A. Shafer, "Using color to separate reflection components," Color Res. Appl. **10**(4), 210–218 (1985).
- 29. H. C. Lee, E. J. Breneman, and C. P. Schulte, "Modeling Light Reflection for Computer Color Vision," IEEE Trans. Pattern Anal. Machine Intell. **12**(4), 402–409 (1990).
- C. Hashemi, R. Avelar, and J. R. Leger, "Multispectral-NLOS-Imaging," GitHub (2023) [accessed 5 December 2023], https://github.com/Hashe037/Multispectral-NLOS-Imaging.