

Thermal Non-Line-of-Sight Imaging

Tomohiro Maeda^{*1}, Yiqin Wang^{*2}, Ramesh Raskar¹, and Achuta Kadambi²

¹MIT Media Lab, MIT, Cambridge, MA 02139 USA

²Electrical Engineering Department, University of California, Los Angeles, CA 90095 USA

We propose a novel non-line-of-sight (NLOS) imaging framework with long-wave infrared (IR). At long-wave IR wavelengths, certain physical parameters are more favorable for high-fidelity reconstruction. In contrast to prior work in visible light NLOS, at long-wave IR wavelengths, the hidden heat source acts as a light source. This simplifies the problem to a single bounce problem. In addition, surface reflectance has a much stronger specular reflection in the long-wave IR spectrum than in the visible light spectrum. We reformulate a light transport model that leverages these favorable physical properties of long-wave IR. Specifically, we demonstrate 2D shape recovery and 3D localization of a hidden object. Furthermore, we demonstrate near real-time and robust NLOS pose estimation of a human figure, the first such demonstration, to our knowledge.

Index Terms—computational photography, non-line-of-sight imaging, infrared imaging

I. INTRODUCTION

Non-line-of-sight (NLOS) imaging has gained significant interest in recent years. The ability to look around corners has many applications, such as search-and-rescue operations, and autonomous vehicle navigation.

Previous methods in NLOS imaging use visible light (wavelengths at $0.4\text{-}0.7\mu\text{m}$). In contrast, this paper proposes a passive NLOS method that operates with long-wave IR (wavelengths at $8\text{-}14\mu\text{m}$). This change of wavelength inspires the design of a fundamentally new light transport formulation. Fig. 1 shows two advantages of our passive NLOS imaging with long-wave IR as compared to visible light. First, since the hidden object is a light source rather than a light reflector, one has to consider fewer light paths (e.g., a one-bounce, rather than a two-bounce scenario). Second, the surface reflectance of ordinary materials (from a rough metallic surface to colored acrylic) have a stronger specular reflection in the long-wave IR spectrum than in the visible light spectrum. This enables us to generalize passive NLOS imaging to new geometries, to scenes without occlusions or shadows.

Contributions: To our knowledge, this is the first “seeing around corners” paper at long-wave IR wavelengths. Our technical contributions include:

- A novel image formation model for NLOS with long-wave IR;
- A BRDF-based NLOS object reconstruction technique; and
- A robust NLOS human pose estimation with noisy reflection image.

Taken together, the technique has applications in passive, low-cost NLOS imaging under real-time operating constraints.

Limitations: Our NLOS reconstruction is limited to objects with a known uniform long-wave IR radiation (temperature) at a single depth (planar). We further assume isotropic long-wave IR radiation of the object. Such assumptions seem strong, but some common object of interest such as the human body has a known long-wave IR emission, and black-body radiation is

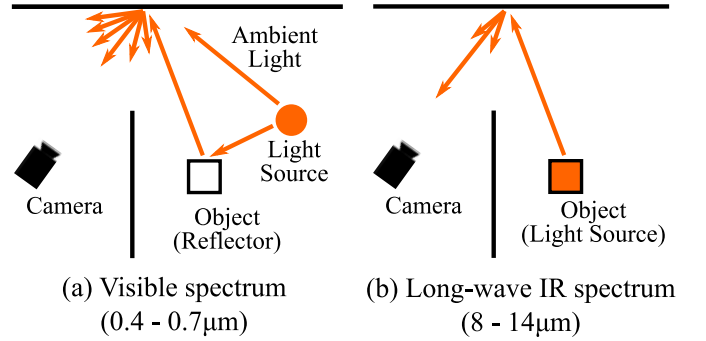


Fig. 1: **NLOS imaging with long-wave IR is fundamentally different from NLOS at the visible wavelengths.** (a) The visible light problem is a two-bounce problem, where the hidden object is a reflector. (b) The long-wave IR problem is a one-bounce problem since the hidden object is a light emitter. In addition to the bounce order, the two problems also differ in their BRDF characteristics.

isotropic [1]. While diffuse surface reflectance is the challenge for NLOS imaging in the visible spectrum, absorption at the surface is the challenge for Thermal NLOS. We demonstrate reconstructions on surfaces with low albedo, but the low SNR could limit the range of materials for our reconstruction and pose estimation algorithms to work.

II. RELATED WORKS

We briefly describe works that relate to this paper. In contrast to prior work in visible-light NLOS imaging, our approach leverages long-wave IR to see around corners. For our extension to succeed, we reformulate the light transport model and reconstruction algorithm for the long-wave IR spectrum. Table 1 compares the general limitations of the proposed approach with other approaches.

Long-wave IR (thermal) imaging: All heat sources radiate light at $8\text{-}14\mu\text{m}$ wavelengths (long-wave IR). A thermal camera captures the long-wave IR spectrum and has been traditionally used for temperature measurements. Many appli-

^{*}T. Maeda and Y. Wang contributed equally to this work.

TABLE I: Comparing the proposed methods with the related NLOS imaging methods.

	ToF	Coherence	RGB	Proposed
Illumination	Active	Active/Passive	Active/Passive	Passive
Cost	High ^a	Low	Low	Low
Ambient Light	Robust	Not Robust	Not Robust	Robust
Depth Resolution	Fine	None ^b	Coarse ^c	Coarse
2D Shape Resolution	Fine	Fine	Coarse	Fine

^a Amplitude modulated continuous wave ToF camera is affordable, but ToF sensors (streak cameras, SPAD) often used for NLOS imaging are costly.

^b Motion can be resolved in 3D [2], but reconstruction is limited in 2D scene.

^c Requires multiple occlusions for depth recovery.

cations can be enhanced at thermal wavelengths. The presence of a human body has a unique long-wave IR signature, which enables robust body pose estimation [3]. Subsequent work in human-computer interaction has exploited the unique long-wave IR reflectance for a novel human-computer interface [4]. A combination of long-wave IR and heat transport enables transient light transport capture at video framerates [5]. Recently, thermal cameras have been a great interest for sensor fusion techniques [6]–[8]. In contrast to prior art, this paper is a first attempt to exploit the beneficial properties of the long-wave IR reflectance, to see around corners.

NLOS with time-of-flight: There is a myriad of prior works that utilize time-of-flight information to see around corners. Perhaps, the first attempt originated with Kirmani et al. [9]. In a seminal demonstration, Velten et al. [10] showed high-quality 3D reconstruction results of a hidden object around a corner using elliptical back-projection. Extensions and reformulations of back-projection have spawned an exciting new area of study [11]–[17]. Recently proposed confocal optical setup formulates a linear inverse problem as a deconvolution problem, enabling fast and efficient reconstruction [18]. While a streak camera and a single photon avalanche diode are expensive, NLOS can be performed with amplitude modulated continuous wave (AMCW) ToF camera, which costs less by several orders of magnitude [19]–[21]. A transient renderer can be used to match NLOS scene with ToF measurements to recover the hidden scene [22]. Other applications of ToF camera in computational imaging such as de-scattering and light transport decomposition can be found in [23]–[27]. In contrast to the ToF methods, this paper exploits the passive reflectance at long-wave IR wavelengths.

NLOS with coherence: The coherence of light is useful for NLOS imaging. Memory effect can be used to recover the hidden scene by performing phase-retrieval algorithms on the auto-correlation of the captured image [28]. Speckle patterns produced by the interaction of active coherent illumination and a hidden scene can be used to track the small motion of

TABLE II: Comparing assumptions and capability of the proposed and the previous passive NLOS techniques.

	Bouman-17 [30]	Saunders-19 [31]	Proposed
Assumptions	Occlusion geometry	Shape and size of occluding object, Known target depth	Known long-wave IR radiation, Single depth
Capabilities	1D tracking, Localization with 2 corners	2D RGB reconstruction	2D shape reconstruction, Localization

multiple objects [2] and recover the hidden scene [29]. While coherence approaches for NLOS imaging can be done with affordable standard cameras, these methods often suffer from a small field of view.

NLOS with standard cameras: Standard camera combined with active illumination can track the motion [32] or perform colored reconstruction [33] of objects around a corner. Occlusion geometries found in the corner scenes can be used to track [30] or reconstruct the hidden objects with passive sensing [31], [34], [35]. Data-driven approaches have shown that it is possible to learn to achieve NLOS imaging using a standard camera [36]. While our method is also a passive NLOS solution, our use of long-wave IR wavelengths requires a reformulation of the light transport model and reconstruction methodology. Though previous passive NLOS solutions require occluding geometries, ours does not. Table 2 compares the assumptions and capability of the proposed technique with previously proposed passive NLOS imaging techniques.

NLOS and surface reflectance: Ordinarily, NLOS imaging was studied for Lambertian surfaces. Kadambi et al. were the first to generalize the seeing around corners problem to exploit arbitrary, non-Lambertian BRDFs [21]. However, Kadambi et al. did not study the NLOS imaging problem without time-of-flight measurement. Although our method also exploits the reflectance function, we need to reformulate the forward and inverse problem for the new scenario of passive, long-wave IR. Sasaki et al. proposed BRDF-based reconstruction of a light field with passive sensing [37]. We mainly focus on 3D localization as 2D reconstruction is trivial with the narrow specular BRDF of the long-wave IR.

Pose estimation: In this paper, we show the results of a human pose estimation around the corner. For line-of-sight scenes, estimation of human pose is not a new problem. Model-based approaches [38], [39] use parameters such as joints orientations that match with articulated templates [39]. The pictorial structure methods [40]–[42] model human body with spring-like joints [43]. Data-driven approach for pose estimation has been successful in recent years [44]–[46]. Although this paper leverages existing pose estimation algorithms, we propose a heuristic combination of denoising and pose estimation that works well on the noisy NLOS measurements.

III. NLOS THERMAL IMAGE FORMATION MODEL

In this section, we describe a new image formation model for NLOS imaging with long-wave IR because the hidden

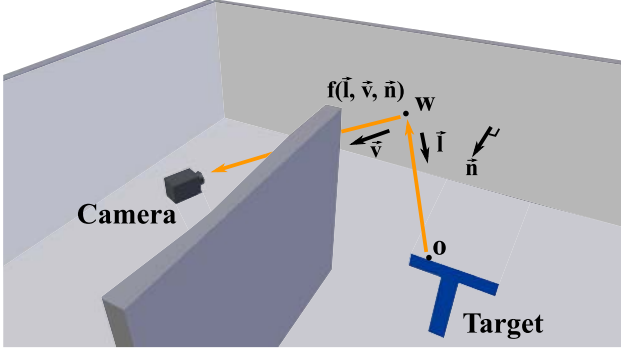


Fig. 2: **Corner setup.** The canonical scene setup consists of a camera looking at the wall at a corner. o and w denote point locations on a target object and wall, and unit vectors $\vec{l}, \vec{v}, \vec{n}$ denote the directions of incoming, outgoing light and surface normal of the wall. $f(\vec{l}, \vec{v}, \vec{n})$ is the BRDF of the wall.

object is a light emitter rather than a reflector and a wall surface has a specular dominant BRDF. We take a corner setup illustrated in Fig. 2 throughout this paper.

A. Emissivity and Albedo

When long-wave IR interacts with a wall, it is either absorbed, reflected or transmitted. Kirchhoff's law of thermal radiation states that emissivity ϵ of material is equivalent to absorptivity. When the transmission is negligible, the albedo of a surface ρ can be written as

$$\rho = 1 - \epsilon \quad (1)$$

B. Temperature and long-wave IR intensity

Stefan-Boltzmann law gives a relationship between the object temperature T_{obj} and long-wave IR radiance:

$$E = \epsilon_o \sigma T_{obj}^4, \quad (2)$$

where ϵ_o is the emissivity of the surface, and σ is the Stefan-Boltzmann constant.

A thermal camera captures the sum of long-wave IR emitted from a wall surface and reflected light from an object. When an object is not present, the camera observes reflected ambient radiation. From Stefan-Boltzmann law, the radiance of the corner wall without the hidden object can be written as

$$I_{background} = \epsilon_w \sigma T_w^4 + (1 - \epsilon_w) \sigma T_{amb}^4, \quad (3)$$

where T_w, T_{amb} denote the wall temperature and ambient temperature. ϵ_s is the emissivity of the corner wall. When an object at a temperature T_{obj} is present in the hidden scene, the radiance of the wall can be written as

$$I_{obj} = \epsilon_w \sigma T_w^4 + (1 - \epsilon_w) \epsilon_o \sigma T_{obj}^4. \quad (4)$$

Subtracting Eq. 4 from Eq. 3, we get the radiance of long-wave IR light due to the long-wave IR radiation of the object,

$$I = (1 - \epsilon_w) \sigma (\epsilon_o T_{obj}^4 - T_{amb}^4). \quad (5)$$

We treat an object as a long-wave IR source with its effective radiance after background subtraction of ambient temperature is

$$E_{obj} = \sigma (\epsilon_o T_{obj}^4 - T_{amb}^4). \quad (6)$$

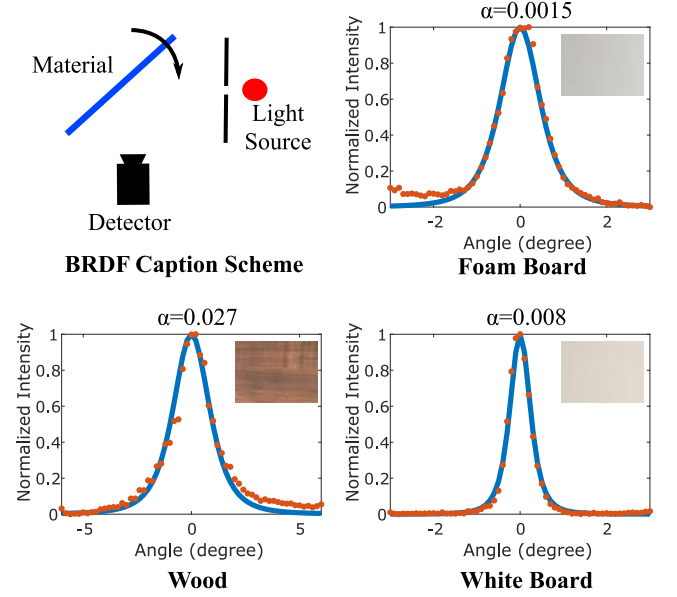


Fig. 3: **Fitting specular BRDF model to long-wave IR surface reflectance of common materials.** The top left diagram shows our BRDF caption scheme. A soldering iron was used as a light source, and a rotation stage was used to capture long-wave IR intensity over the varying angle between half vector $\vec{h}$ and surface normal $\vec{n}$. The plots show GGX specular BRDF models fitted to the measurements of a variety of materials. Fitting to other materials can be found in Fig. 9.

C. Specular BRDF Model

There are various analytic models to approximate the BRDF of surfaces [47]–[51]. Our key observation is that the BRDF of common materials in the long-wave IR spectrum can be approximated with a specular BRDF model as shown in Fig. 3. This observation agrees with the work of Bennett and Porteus, which showed that the strength of diffuse reflection decreases by a factor of $1/\lambda^4$, where λ is the wavelength of light [52]. As long-wave IR has a much longer wavelength than visible light, measured diffuse term of BRDF was negligible, and specular BRDF models approximate the BRDF of various materials well. While Tanaka et al. [53] suggests that diffuse reflection may occur by heat transfer from directed long-wave IR sources, we didn't observe heat transfer for sources with isotropic heat emission.

While several specular BRDF models reasonably fit the measured BRDF, we found that microfacet-based Cook-Torrance model [48] with the GGX distribution [49] best fits the observations (see Appendix A). Hence, we write BRDF of the wall as a function of unit vectors that represent directions of the surface normal of a wall $\vec{n}$, incoming light $\vec{l}$, and outgoing light $\vec{v}$.

$$f(\vec{l}, \vec{v}, \vec{n}) = \frac{D(\vec{h}, \vec{n}) F(\vec{v}, \vec{h}) G(\vec{l}, \vec{v}, \vec{h}, \vec{n})}{4(\vec{l} \cdot \vec{n})(\vec{v} \cdot \vec{n})}, \quad (7)$$

where $\vec{h} = (\vec{l} + \vec{v}) / \|\vec{l} + \vec{v}\|$ is a half vector. $D(\vec{h}, \vec{n})$, $G(\vec{l}, \vec{v}, \vec{h}, \vec{n})$ and $F(\vec{v}, \vec{h})$ denote surface normal distribution function, the shadowing factor, and Fresnel function (see

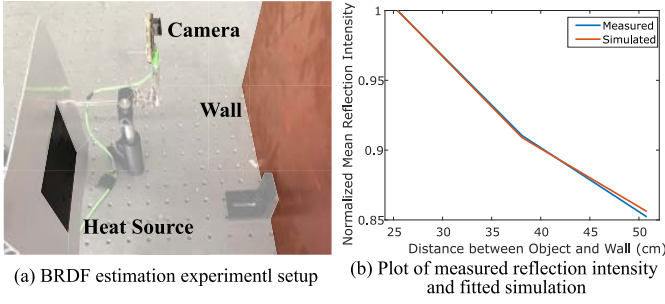


Fig. 4: **Alternative approach for BRDF parameter estimation.** (a) Experimental setup for measurements. The same scene is simulated in physical-based render using the Cook-Torrance model with GGX distribution, parameterized by α . (b) When α is estimated such that the simulated measurements match with the experimental measurement.

Appendix B for details). This BRDF model is parameterized by a single parameter α , which represents ‘roughness’ of the surface.

D. Light Transport Model

A light transport model describes how long-wave IR light emitted from the object reaches the camera. Let us consider a planar object with a temperature of T_{obj} . This object acts as an area light source with radiance E_{obj} described in Eq. 6. The irradiance of a wall from a single point of the object can be written as

$$L(\mathbf{o}, \mathbf{w}) = \frac{E_{obj}}{\pi \|\mathbf{o} - \mathbf{w}\|^2}, \quad (8)$$

where $\mathbf{o}, \mathbf{w}$ denote positions of points on the object and the wall. The flux on a camera pixel corresponding to a wall patch $\mathbf{W}_i$ can be written as an integral of flux over points on $\mathbf{W}_i$ and outgoing directions of photons that reach the camera $\vec{V}$.

$$I(\mathbf{o}, \mathbf{W}_i) = \int_{\mathbf{w} \in \mathbf{W}_i} \int_{\vec{v} \in \vec{V}} (1 - \epsilon_w) L(\mathbf{o}, \mathbf{w}) f(\vec{l}, \vec{v}, \vec{n}) d\vec{v} d\mathbf{w}, \quad (9)$$

where $1 - \epsilon$ is equivalent to the albedo ρ . Finally, the observed intensity from the hidden object at the camera pixel seeing a wall patch $\mathbf{W}_i$ is written as

$$I(\mathbf{W}_i) = \int_{\mathbf{o} \in \mathbf{O}} I(\mathbf{o}, \mathbf{W}_i) d\mathbf{o}, \quad (10)$$

where the integral is over all the points on the object.

IV. NLOS SCENE RECONSTRUCTION

Our reconstruction method efficiently recovers the 2D shape of the object from a narrow specular BRDF in the long-wave IR spectrum. When specular reflection has some width, the intensity of the object reflection drops as the object moves further away from the wall. Using the the known long-wave IR emission of the object, our algorithm recovers the distance between the wall and the object.

We assume that the geometry of the wall is known with respect to the camera. Now, we describe our reconstruction algorithm step by step.

A. BRDF Estimation

BRDF of surfaces can be recovered by measuring the angle-varying intensity of the surface reflection as illustrated in Fig. 3. This can also be accomplished by moving the light source or the detector instead of moving the material directly.

Alternative approach for BRDF fitting: Full capture of BRDF usually requires a precisely controlled environment. Instead of capturing full BRDF, the estimation of roughness parameter α of the GGX distribution can be performed in a simpler manner as summarized in Fig. 4.

We take k measurements with a known long-wave IR source. We simulate the same scene in a physically-based renderer to estimate α that is consistent with the measurement. We denote the average intensity of reflection of the object as $\mu = [\mu_1, \dots, \mu_k]$. Simulated measurements $\hat{\mu} = [\hat{\mu}_1(\alpha, \epsilon), \dots, \hat{\mu}_k(\alpha, \epsilon)]$ are parameterized by α and albedo ρ . Because the albedo is a constant outside of integrals in Eq. 9 and 10, the ratio of elements in μ and $\hat{\mu}$ is parameterized only by α . Hence, α can be estimated as

$$\hat{\alpha} = \arg \min_{\alpha} \left\| \frac{\mu}{\mu_1} - \frac{\hat{\mu}}{\hat{\mu}_1} \right\|^2. \quad (11)$$

$\hat{\mu}$ is continuous with respect to α , so gradient descent can be used, where the gradient is numerically approximated through the simulation. Once α is estimated, the albedo of the wall can be estimated as

$$\hat{\rho} = \rho_s \frac{\mu_0 E_s}{\hat{\mu}_0 E_{obj}}, \quad (12)$$

where ρ_s, E_s are the simulated albedo and radiance.

B. Forward Model Construction

Light transport matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ represents the mapping between the radiance of the hidden object $\mathbf{x} \in \mathbb{R}^n$ to the measurement $\mathbf{y} \in \mathbb{R}^m$:

$$\mathbf{y} = \mathbf{A}\mathbf{x}, \quad (13)$$

Once the parameters of the BRDF model are estimated, light transport model $\hat{\mathbf{A}}$ can be constructed using Eq. 9 and 10. The integrals in these equations are hard to evaluate analytically. We use Monte Carlo path tracing simulation to approximate the integrals. We discretize the hidden scene and recover $\hat{\mathbf{A}}$ by moving a small area light source at discrete voxel locations in the simulation.

C. Reconstruction

In the long-wave IR spectrum, recovery of the 2D shape is trivial as surface reflectance has a dominant specular reflection. Hence, we reconstruct the 2D shape of the object first, then estimate the location of the object using uniform, known temperature. Fig. 5 illustrates an overview of our reconstruction algorithm. Our technique assumes uniform temperature and single depth of the hidden object but does not require prior knowledge on the 2D shape or the size of the object.

2D shape recovery: Let’s consider a simple case where we know the distance between the hidden planar object and

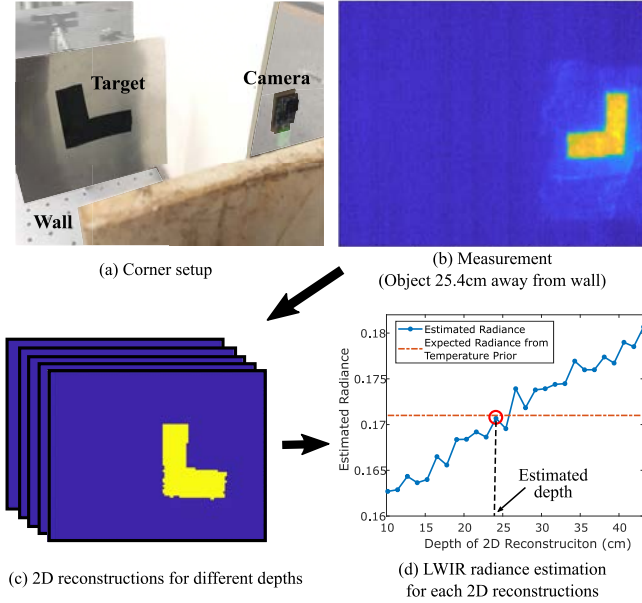


Fig. 5: **Reconstruction algorithm for 2D shape recovery and 3D localization.** (a) Corner setup. From (b) measurement, (c) we perform 2D shape reconstruction at different depths. (d) Given the fitted BRDF model, long-wave IR radiance of an object is estimated for each depth. We estimate depth by finding a depth whose estimated radiance is the closest to expected radiance from the temperature prior.

the wall. Since BRDF of a wall only has a narrow specular component, the 2D shape of the object is visible. Hence, we approximate the forward model with ideal mirror reflection $\mathbf{M} \approx \hat{\mathbf{A}} \in \mathbb{R}^{m \times n}$ for an efficient shape recovery. Under the uniform temperature assumption, we can write $x = E_{obj} \hat{\mathbf{x}}$, where $\hat{\mathbf{x}} \in \{0, 1\}$ is an estimated 2D shape that spans over a fixed depth. The object shape can be recovered as

$$\hat{\mathbf{x}}_i \approx \mathbb{I} \left\{ \sum_{j=1}^m \mathbb{I} \{ \mathbf{M}_{ji} \geq \gamma_1 \} \mathbb{I} \{ \mathbf{y}_j \geq \gamma_2 \} \geq 1 \right\}, \quad (14)$$

, where $\mathbb{I} \{ \cdot \} \in \{0, 1\}$ is an indicator. $\hat{\mathbf{x}}_i = 1$ indicates the object is present, and 0 indicates otherwise.

3D localization: We perform 2D shape reconstruction for multiple depths $\mathbf{d} = [d_1, d_2, \dots, d_l]$. For each 2D reconstruction $[\hat{\mathbf{x}}^{(1)}, \hat{\mathbf{x}}^{(2)}, \dots, \hat{\mathbf{x}}^{(l)}]$, we estimate long-wave IR radiance of the object that fits with the measurement.

$$\hat{E}_{obj}^{(i)} = \arg \min_E \left\| \mathbf{y} - E \hat{\mathbf{A}}^{(i)} \hat{\mathbf{x}}^{(i)} \right\|^2, \quad (15)$$

where $\hat{\mathbf{A}}^{(i)}$ is a light transport model for voxels at fixed distance from the wall d_i . In practice, the discretized model in simulation does not match perfectly with the measurement (e.g., mismatch of the discretized reconstruction and continuous world). We overcome this problem by taking an average of object reflection intensity values from $\mathbf{y}$ and $E \hat{\mathbf{A}}^{(i)} \hat{\mathbf{x}}^{(i)}$, and minimize the difference of these two quantities. Then, we estimate the 3D location of the object as follows:

$$\hat{d} = \arg \min_{d_i \in \mathbf{d}} \left\| E_{obj} - \hat{E}_{obj}^{(i)} \right\|^2. \quad (16)$$

The 3D localization capability is limited by emissivity and roughness of the wall, size of the object, and temperature of the object. See Section VI-A for further discussions.

V. NLOS 2D POSE ESTIMATION

The 2D shape of the hidden object is visible from the reflection because of the dominant specular reflection in the long-wave IR spectrum. We exploit this fact to NLOS 2D pose estimation. In this section, we extend line-of-sight pose estimation algorithms to NLOS scenes. We show that in the long-wave IR, even a simple preprocessing of thermal images can enable a robust pose estimation around corners.

A. 2D Pose Estimation

We follow Yang-Ramanan model for 2D pose estimation, which estimates human pose using a mixture of non-oriented pictorial structures. Yang-Ramanan 2D pose estimation outputs small patches within images that are likely to contain a specific class of body parts such as head and torso. The key ideas of this approach are as follows. First, local features of the feature such as the histogram of oriented gradients (HOG) [54]. Second, the spatial connection between the patches. Third, the consistency of local features in connected patches. We refer readers to [55] for further details of the algorithm.

While Yang-Ramanan model works well for a line-of-sight body in RGB images, it fails on noisy thermal reflection images of a body in NLOS. This is because HOG of NLOS thermal images is much less structured than typical line-of-sight RGB images as shown in Fig. 7 (a).

B. Processing Thermal Image

Typical NLOS body images on a wall with high emissivity look ‘blurry’ (Fig. 10 (b)). Based on our analysis on BRDF discussed in Section III-C, we conclude that the ‘blurriness’ is due to low SNR caused by the low albedo of the surface. From this observation, we apply denoising to thermal images to improve NLOS pose estimation.

We denoise thermal images with two methods. First, we use a median filter, which removes the additive noise and preserves the edges. Second, we further denoise images with total variation (TV) regularization:

$$\mathbf{I}_{TV} = \arg \min_{\hat{\mathbf{I}}} \left\| \mathbf{I}_{med} - \hat{\mathbf{I}} \right\|^2 + \lambda \left\| \nabla \hat{\mathbf{I}} \right\|_1, \quad (17)$$

where $\mathbf{I}_{med}$ is an image processed with median filtering. The second term encourages smaller total variation of the resulting image. See Section VI-B for analysis on how denoising improves HOG features of NLOS body reflections, and thus, improves pose estimation accuracy in thermal images.

VI. ANALYSIS

A. Analysis on Depth Recovery

The BRDF-based object localization method exploits the fact that the intensity of the reflection drops as the object moves further away from the wall. The emissivity ϵ , the roughness of a wall α , and the size of the object affect the

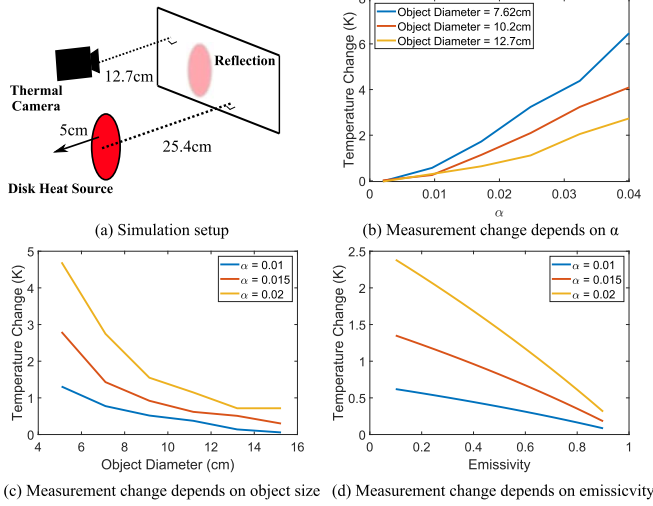


Fig. 6: **Depth recovery sensitivity depends on roughness, size of the object, and emissivity of the wall.** (a) Simulation setup for thermal measurement of the object reflections. (b-d) Plots of temperature change when the object is moved from 25.4cm to 30.4cm away from the wall.

depth recoverability. The direct evaluation of depth resolution is challenging due to the complexity of the GGX BRDF model. However, we provide the intuition of what affects the depth recoverability by analyzing how temperature measurement of the reflection changes when the hidden object moves away from the corner wall. We simulated the thermal images using Mitsuba renderer, and plotted how the temperature of the reflection image changes as the object moves further away from the wall (Fig. 6). The simulated ambient and object temperatures were 37, 10°C, and the object was moved from 25.4cm to 30.4cm away from the wall. The horizontal distance between the object and camera is 10.2cm.

Roughness: The roughness of the wall α affects the depth resolution. If a wall is an ideal mirror, reflection intensity of an area source does not change when the hidden objects move to another depth. This is because the surface area of the object corresponding to a camera pixel grows by the square of the distance, while irradiance per area of the object decays by the square of the distance. Indeed, imperfect specular reflection allows reflection intensity to change as the hidden object changes its depth as illustrated in Fig. 6 (b).

Object size: When the hidden object is infinitely large, the integral in Eq. 9 and 10 becomes the same for any depth with the infinitely large object, which makes depth estimation impossible. When object size becomes too large, localization is challenging (Fig. 6(c)). This is because as the $(\vec{h} \cdot \vec{n})$ becomes larger, BRDF becomes negligibly small. However, our method is scalable to larger, practical scenarios because BRDF is a function of angles, and scaling both object size and scene geometry preserves $\vec{l}, \vec{v}$ and $\vec{n}$ (Appendix C).

Emissivity: When a wall has a high emissivity (low albedo), intensity change due to the hidden object could be

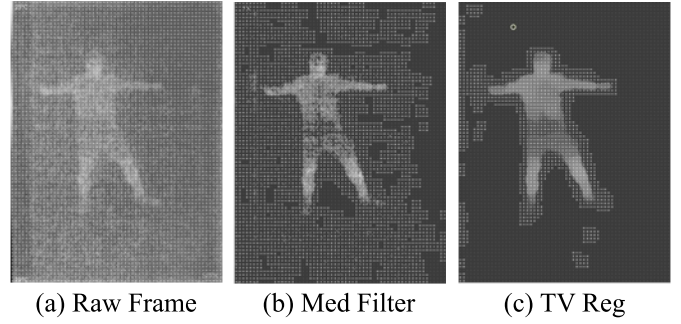


Fig. 7: **Comparing HOG features in the raw frames and the denoised frames.** (a) The raw thermal frame does not have a clear gradient structure due to a high noise level. (b) Median filtering mitigates noisy gradients. (c) TV regularization removes most of the gradient where a body is not present.

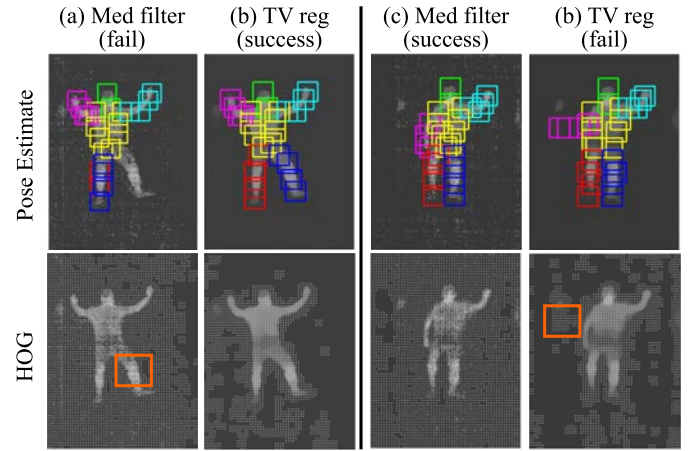


Fig. 8: **Analysis on failure cases.** (a,b) Median filtering did not remove high-frequency noise. Leg part surrounded by an orange box has only small gradients. (c,d) TV denoising introduced artificial gradients in the region surrounded by an orange box, which tricked the pose estimation algorithm.

below the noise level of the sensor. Some materials with high albedo in the visible spectrum could have low albedo in the long-wave IR spectrum. Such problems can be solved by averaging many frames to improve SNR in practice. Ambient and object temperature also limits the SNR. When the difference between these two temperatures is small, the long-wave IR radiation due to the presence of the hidden object becomes small.

Model mismatch: One source of model mismatch is from the discretization of the hidden scene while the physical world is continuous. For example, a small change in voxel location affects simulated intensity around the edge in the reflection image of the hidden object. We use averaged intensity for depth recovery to overcome this mismatch. Other sources of model mismatch come from the accuracy of approximation with analytic BRDF model and accuracy of integral approximation with the Monte Carlo method.

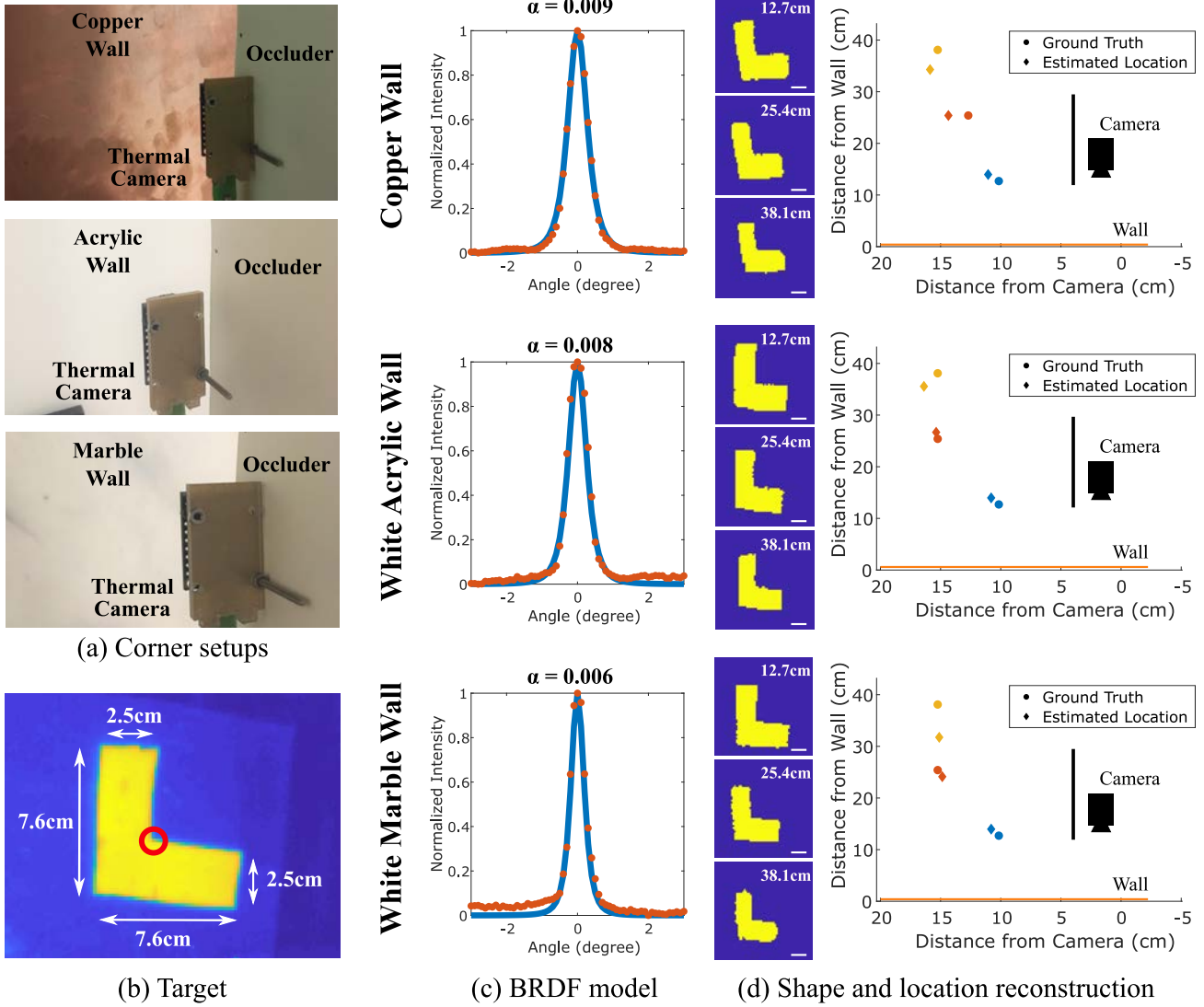


Fig. 9: **NLOS scene reconstruction results.** (a) Corner setups. In the visible spectrum, the object is not visible on the wall. (b) Thermal image of the target. (c) BRDF is fitted with GGX specular BRDF model. (d) Localization of the target (red circle shown in (b)). Reconstruction does not require prior knowledge on the shape or size of the hidden object.

B. Analysis of Pose Estimation Improvement

Analytic evaluation of the pose estimation accuracy is challenging as a body has many degrees of freedom. We investigate how denoising improves HOG features of images, which is a base for various pose estimation algorithms.

Fig. 7 visualizes HOG features on raw and processed frames. Median filtering mitigates noisy gradients, and TV regularization removes gradients where a body is not present. Fig. 8 provides visual analysis of failure cases for median filtering and TV denoising. (a), (b) show that TV denoising could remove noise that median filter could not remove. (c), (d) show the case where TV denoising introduces artifacts, which results in pose estimation error.

A t-shirt and middle-length trousers were used for the hidden person’s clothing in this paper. Different types of clothing may affect the pose estimation capability. For example, if there is no heat transfer between human skin to the surface

of the clothes, the temperature of the clothes would be the same as the ambient temperature, which does not appear in the reflection.

VII. EXPERIMENTAL VALIDATIONS

We describe the simulation and experiment setups we used for validation of our techniques.

A. Simulation

A physically-based Mitsuba renderer [56], was used to approximate the light transport model. We simulated a wall with a Cook-Torrance model with GGX microsurface distribution. Mitsuba’s implementation of this model follows [49]. Though the Mitsuba renderer does not render thermal images, Eq. 2 allows us to convert a thermal image to a long-wave IR light intensity image.

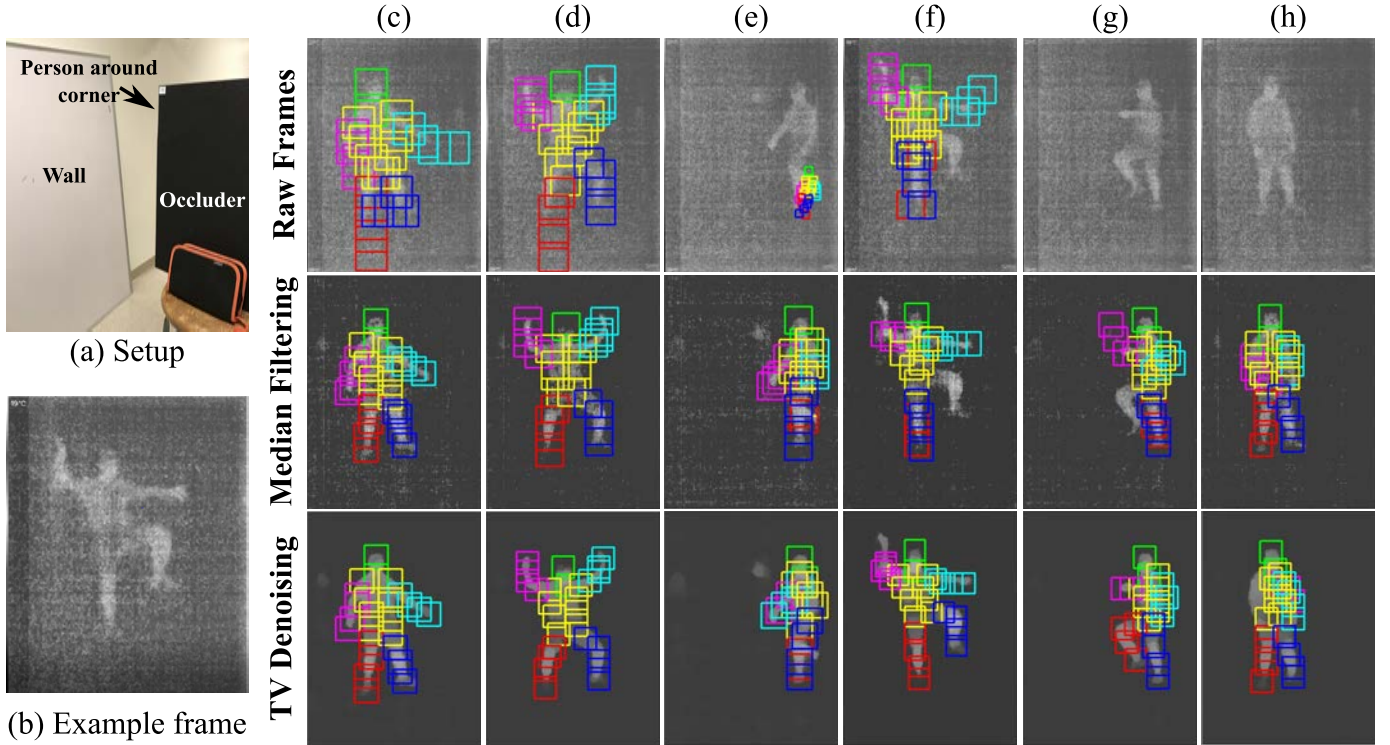


Fig. 10: **NLOS pose estimation results.** We show that processing raw frames of thermal images can improve the pose estimation algorithm. (a) shows NLOS imaging setup. (b) is an example of raw output from the camera. (c), (d), (e) Appropriate processing of raw thermal video frames results in better posture estimation. (f) (g) TV regularization could correct pose estimation on median filtering, while in (h), it could introduce artifacts that result in pose estimation error. (g) and (h) include human detection through score thresholding. Raw frames did not get high enough score for body detection.

B. Experiments

In this section, we describe the experimental setups for NLOS scenes.

BRDF caption: To estimate α accurately, we fitted the GGX specular BRDF model to intensity measurement captured by moving the wall material as illustrated in Fig. 3. We used PureThermal 2 module with FLIR Lepton 3.5 thermal sensor as a detector, and soldering iron as a point long-wave IR source. For the parameter fitting showed in Fig. 4, we placed 7.6 by 7.6cm heat source with a known temperature 25.4, 38.1, and 50.8cm away from the wall while the camera was fixed at 12.4cm away from the wall. The same scene was simulated in Mitsuba, and albedo ρ was estimated using Eq. 12.

NLOS reconstruction: Fig. 9 (a) illustrates the experimental corner setups with copper, acrylic and marble walls. To create a planar target, an aluminum plate was attached on an IR heater to distribute the heat uniformly, and a electric tape with emissivity close to 1 was placed for long-wave IR radiation. Heated aluminum with low emissivity emits negligible long-wave IR. The L-shaped planar Target temperature was set at between 50 – 60°C and the ambient temperature was around 32°C. We averaged 30 thermal images for each measurement. Fig. 9 (d) summarizes experimental setup geometry. The camera was 12.7cm away from the wall, and objects were placed at 25.4, 38.1, and 50.8cm away from the wall. The background image was captured before placing the heat source

for background subtraction of ambient temperature described in Eq. 6. For reconstruction of L shaped target, we chose the threshold γ_1 and γ_2 in Eq. 14 to be $\max\{y\}/1.5$ and 0.07 for all the reconstruction. The resolution of 2D shape reconstruction is 80 by 120 with a pixel size of 2.54 by 2.54mm. The depth candidates are sampled between 10.16 and 43.18cm (4-17inch) with the interval of 1.27cm (0.5inch).

Pose estimation: Fig. 10 (a) shows the experimental setup, where a thermal camera looked at a whiteboard with a viewing angle of roughly 45 degrees. The camera was roughly 1m away from the wall, and the hidden person wearing a t-shirt and middle-length trousers was 0.4m away from the wall. The ambient temperature was around 20°C. We used a Seek Compact Pro thermal camera with 320 by 240 pixels and 70mK thermal sensitivity. The thermal camera outputs frames with the normalized contrast. We used a median filter with a kernel size of 3, and TV regularization with $\lambda = 0.2$ for denoising. The implementation of pose estimation followed [57]. Denoising and pose estimation took about 1.8 seconds on a 160 by 120 downsampled thermal image on an unoptimized laptop computer. This result shows the potential for real-time NLOS pose estimation.

C. Results

Estimated roughness α and albedo ρ for copper, acrylic, and marble surfaces are 0.009, 0.008, 0.006, and 0.84, 0.038,

0.094. Fig. 9 (c) shows the fitted model and intensity measurements. Fig. 9 (d) summarizes 2D shape reconstruction and localization.

Fig. 10 illustrates some frames of video that compares body pose estimation performance. Table III shows the accuracy of pose estimation evaluated over 210 video frames. NLOS body detection was performed by score thresholding. See the supplemental material for the resulting video. We showed that denoising thermal images improves the accuracy of a pose estimation algorithm on NLOS thermal images. Surprisingly, TV denoising did not improve pose estimation accuracy from median filtered images other than left leg detection. As shown in Fig 8, this may be due to artifacts that the TV regularizer introduces, while there are frames where TV denoising improves pose-estimation accuracy.

TABLE III: Comparing accuracy of pose estimation for raw thermal frames and denoised thermal frames.

	Head	Torso	R Arm	L Arm	R Leg	L Leg
Raw Frames	0.70	0.72	0.56	0.49	0.46	0.42
Med Filter	0.98	0.98	0.83	0.69	0.90	0.82
TV Reg	0.98	0.94	0.77	0.68	0.86	0.86

VIII. DISCUSSION AND CONCLUSION

In summary, we have proposed and demonstrated an early attempt at NLOS scene reconstruction and human pose estimation in the long-wave IR spectrum. Going forward, we believe that NLOS at long wave IR is an exciting prospect for two reasons:

- The problem is simplified to one-bounce of photons, as the object is a source of light; and
- Many surfaces are specular at long-wave IR wavelengths.

Object as a source of light (one-bounce problem): Recall that in the visible spectrum, the hidden objects act as reflectors, making NLOS imaging a two-bounce problem. However, in this paper, the hidden object is a source of long-wave IR radiation. This simplification allows NLOS imaging to be modeled as a one-bounce problem. Moreover, if the temperature of the object is known, the emitter radiance is easily estimated from Stefan-Boltzmann’s law. We exploit this unique property of long-wave IR for a novel NLOS reconstruction method that can generalize to both 2D and 3D cases.

Narrow specular BRDF: From rough metallic surface to ‘diffuse’ whiteboard, various materials have dominant specular surface reflection in the long-wave IR spectrum. This allows us to recover the 2D shape of the hidden object easily while it is challenging in the visible spectrum. With imperfect specular reflection, the intensity of reflection changes as the distance between the wall and the hidden object changes. We exploited this for 3D localization of the hidden object.

Limitations and future Work: Our reconstruction algorithm is limited to a planer object with a uniform, known long-wave emission. For simplicity, we exploited that the blurring due to the reflection of the wall is trivial due to the narrow specular reflection to recover the 2D shape of the hidden

object. However, there is still a small blurring due to the reflection, and deblurring algorithms can be used to recover the finer details of the object with a non-uniform temperature profile. We observed that some materials exhibit specular reflection, but low albedo, as seen in our pose estimation results. While we showed that simple denoising makes the pose estimation algorithm robust, denoisers for thermal images can be designed to improve the accuracy.

Outlook and conclusion: We hope this paper inspires new algorithms and light transport models across different wavelengths. Going forward, one can imagine a sensor fusion approach that combines the variation in BRDF across different wavelengths to improve NLOS imaging performance. We believe that long-wave IR could be the method of choice to see around corners for many applications. Our results are powered by the simplicity of a one-bounce problem and a strong specular surface reflectance. Future work can build upon such favorable physics to further realize the vision of a camera that can see around corners.

ACKNOWLEDGMENT

We thank the anonymous reviewers for their detailed comments. The work of all MIT affiliated authors was supported by the Media Lab consortium funding. The work of UCLA-affiliated authors is partially supported by an NSF Research Initiation Award (IIS 1849941). R. Raskar is partially supported by the DARPA REVEAL program (N00014-18-1-2894).

REFERENCES

- [1] G. B. Rybicki and A. P. Lightman, *Radiative Processes in Astrophysics*. New York, NY: Wiley, 1985.
- [2] B. M. Smith, M. O’Toole, and M. Gupta, “Tracking multiple objects outside the line of sight using speckle imaging,” in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.
- [3] S. Iwasawa, K. Ebihara, J. Ohya, and S. Morishima, “Real-time estimation of human body posture from monocular thermal images,” in *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR ’97)*, ser. CVPR ’97. Washington, DC, USA: IEEE Computer Society, 1997, pp. 15–.
- [4] A. Sahami Shirazi, Y. Abdelrahman, N. Henze, S. Schneegass, M. Khalilbeigi, and A. Schmidt, “Exploiting thermal reflection for interactive systems,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI ’14. New York, NY, USA: ACM, 2014, pp. 3483–3492.
- [5] K. Tanaka, N. Ikeya, T. Takatani, H. Kubo, T. Funatomi, and Y. Mukaigawa, “Time-resolved light transport decomposition for thermal photometric stereo,” in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.
- [6] W. Treible, P. Saponaro, S. Sorensen, A. Kolagunda, M. O’Neal, B. Phelan, K. Sherbondy, and C. Kambhamettu, “Cats: A color and thermal stereo benchmark,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017, pp. 134–142.
- [7] Y. Cao, B. Xu, Z. Ye, J. Yang, Y. Cao, C.-L. Tisse, and X. Li, “Depth and thermal sensor fusion to enhance 3d thermographic reconstruction,” *Opt. Express*, vol. 26, no. 7, pp. 8179–8193, Apr 2018.
- [8] A. Torabi, G. Massé, and G.-A. Bilodeau, “An iterative integrated framework for thermal-visible image registration, sensor fusion, and people tracking for video surveillance applications,” *Comput. Vis. Image Underst.*, vol. 116, no. 2, pp. 210–221, Feb. 2012.
- [9] A. Kirmani, T. Hutchison, J. Davis, and R. Raskar, “Looking around the corner using transient imaging,” in *2009 IEEE 12th International Conference on Computer Vision*, Sept 2009, pp. 159–166.

- [10] A. Velten, T. Willwacher, O. Gupta, A. Veeraraghavan, M. G. Bawendi, and R. Raskar, "Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging," *Nature communications*, vol. 3, p. 745, 2012.
- [11] V. Arellano, D. Gutierrez, and A. Jarabo, "Fast back-projection for non-line of sight reconstruction," *Opt. Express*, vol. 25, no. 10, pp. 11 574–11 583, May 2017.
- [12] O. Gupta, T. Willwacher, A. Velten, A. Veeraraghavan, and R. Raskar, "Reconstruction of hidden 3d shapes using diffuse reflections," *Opt. Express*, vol. 20, no. 17, pp. 19 096–19 108, Aug 2012.
- [13] F. Heide, M. O'Toole, K. Zhang, D. B. Lindell, S. Diamond, and G. Wetzstein, "Robust non-line-of-sight imaging with single photon detectors," *CoRR*, vol. abs/1711.07134, 2017.
- [14] A. K. Pediredla, M. Buttafava, A. Tosi, O. Cossairt, and A. Veeraraghavan, "Reconstructing rooms using photon echoes: A plane based model and reconstruction algorithm for looking around the corner," in *2017 IEEE International Conference on Computational Photography (ICCP)*, May 2017, pp. 1–12.
- [15] C.-Y. Tsai, K. Kutulakos, S. G. Narasimhan, and A. Sankaranarayanan, "The geometry of first-returning photons for non-line-of-sight imaging," 07 2017, pp. 2336–2344.
- [16] M. L. Manna, F. Kine, E. Breitbach, J. Jackson, T. Sultan, and A. Velten, "Error backprojection algorithms for non-line-of-sight imaging," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–1, 2018.
- [17] X. Liu, I. Guillén, M. L. Manna, J. H. Nam, S. A. Reza, T. H. Le, D. Gutierrez, A. Jarabo, and A. Velten, "Virtual wave optics for non-line-of-sight imaging," *CoRR*, vol. abs/1810.07535, 2018.
- [18] M. O'Toole, D. B. Lindell, and G. Wetzstein, "Confocal Non-Line-of-Sight Imaging Based on the Light-Cone Transform," *Nature*, 2018.
- [19] F. Heide, L. Xiao, W. Heidrich, and M. B. Hullin, "Diffuse mirrors: 3d reconstruction from diffuse indirect illumination using inexpensive time-of-flight sensors," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, June 2014, pp. 3222–3229.
- [20] A. Kadambi, R. Whyte, A. Bhandari, L. Streeter, C. Barsi, A. Dorrington, and R. Raskar, "Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles," *ACM Transactions on Graphics (TOG)*, vol. 32, no. 6, p. 167, 2013.
- [21] A. Kadambi, H. Zhao, B. Shi, and R. Raskar, "Occluded imaging with time-of-flight sensors," *ACM Trans. Graph.*, vol. 35, no. 2.
- [22] J. Iseringhausen and M. B. Hullin, "Non-line-of-sight reconstruction using efficient transient rendering," *CoRR*, vol. abs/1809.08044, 2018.
- [23] S. Jayasuriya, A. K. Pediredla, S. Sivaramakrishnan, A. C. Molnar, and A. Veeraraghavan, "Depth fields: Extending light field techniques to time-of-flight imaging," *2015 International Conference on 3D Vision*, pp. 1–9, 2015.
- [24] K. Tanaka, Y. Mukaigawa, H. Kubo, Y. Matsushita, and Y. Yagi, "Recovering transparent shape from time-of-flight distortion," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4387–4395, 2016.
- [25] A. Kadambi, J. Schiel, and R. Raskar, "Macroscopic interferometry: Rethinking depth estimation with frequency-domain time-of-flight," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016, pp. 893–902.
- [26] C. Tsai, A. Veeraraghavan, and A. C. Sankaranarayanan, "Shape and reflectance from two-bounce light transients," in *2016 IEEE International Conference on Computational Photography (ICCP)*, May 2016, pp. 1–10.
- [27] G. Satat, M. Tancik, and R. Raskar, "Towards photography through realistic fog," in *2018 IEEE International Conference on Computational Photography (ICCP)*, May 2018, pp. 1–10.
- [28] O. Katz, P. Heidmann, M. Fink, and S. Gigan, "Non-invasive real-time imaging through scattering layers and around corners via speckle correlations," *Nature Photonics*.
- [29] A. Viswanath, P. Rangarajan, D. MacFarlane, and M. P. Christensen, "Indirect imaging using correlative," in *Imaging and Applied Optics 2018 (3D, AO, AIO, COSI, DH, IS, LACSEA, LS&C, MATH, pcAOP)*. Optical Society of America, 2018, p. CM2E.3.
- [30] K. L. Bouman, V. Ye, A. B. Yedidia, F. Durand, G. W. Wornell, A. Torralba, and W. T. Freeman, "Turning corners into cameras: Principles and methods," in *2017 IEEE International Conference on Computer Vision (ICCV)*, Oct 2017, pp. 2289–2297.
- [31] C. Saunders, J. Murray-Bruce, and V. Goyal, "Computational periscopy with an ordinary digital camera," *Nature*, vol. 565, p. 472, 01 2019.
- [32] J. Klein, C. Peters, J. Martín, M. Laurenzis, and M. B. Hullin, "Tracking objects outside the line of sight using 2d intensity images," *Scientific Reports*, vol. 6, no. 32491, Aug. 2016.
- [33] W. Chen, S. Daneau, F. Mannan, and F. Heide, "Steady-state non-line-of-sight imaging," *CoRR*, vol. abs/1811.09910, 2018.
- [34] C. Thrampoulidis, G. Shulkind, F. Xu, W. T. Freeman, J. H. Shapiro, A. Torralba, F. N. C. Wong, and G. W. Wornell, "Exploiting occlusion in non-line-of-sight active imaging," *IEEE Transactions on Computational Imaging*, vol. 4, no. 3, pp. 419–431, Sept 2018.
- [35] M. Baradad, V. Ye, A. B. Yedidia, F. Durand, W. T. Freeman, G. W. Wornell, and A. Torralba, "Inferring light fields from shadows," in *CVPR*, 2018.
- [36] M. Tancik, T. Swedish, G. Satat, and R. Raskar, "Data-driven non-line-of-sight imaging with a traditional camera," in *Imaging and Applied Optics 2018 (3D, AO, AIO, COSI, DH, IS, LACSEA, LS&C, MATH, pcAOP)*. Optical Society of America, 2018, p. IW2B.6.
- [37] T. Sasaki and J. Leger, "Light-field reconstruction from scattered light using plenoptic data," in *Unconventional and Indirect Imaging, Image Reconstruction, and Wavefront Sensing 2018*, J. Dolne and P. Bones, Eds., vol. 10772. United States: SPIE, 1 2018.
- [38] J. O'Rourke and N. I. Badler, "Model-based image analysis of human motion using constraint propagation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 6, pp. 522–536, 1980.
- [39] D. Hogg, "Model-based vision: a program to see a walking person," *Image and Vision computing*, vol. 1, no. 1, pp. 5–20, 1983.
- [40] M. Andriluka, S. Roth, and B. Schiele, "Pictorial structures revisited: People detection and articulated pose estimation," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, 2009, pp. 1014–1021.
- [41] P. F. Felzenszwalb and D. P. Huttenlocher, "Pictorial structures for object recognition," *International journal of computer vision*, vol. 61, no. 1, pp. 55–79, 2005.
- [42] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2d pose estimation using part affinity fields," *arXiv preprint arXiv:1611.08050*, 2016.
- [43] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient matching of pictorial structures," in *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, vol. 2. IEEE, 2000, pp. 66–73.
- [44] A. Toshev and C. Szegedy, "DeepPose: Human pose estimation via deep neural networks," *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Jun 2014.
- [45] A. Newell, K. Yang, and J. Deng, "Stacked hourglass networks for human pose estimation," in *Computer Vision – ECCV 2016*, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. Cham: Springer International Publishing, 2016, pp. 483–499.
- [46] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2d pose estimation using part affinity fields," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1302–1310, 2017.
- [47] J. F. Blinn, "Models of light reflection for computer synthesized pictures," *SIGGRAPH Comput. Graph.*, vol. 11, no. 2, pp. 192–198, Jul. 1977.
- [48] R. L. Cook and K. E. Torrance, "A reflectance model for computer graphics," *ACM Trans. Graph.*, vol. 1, no. 1, pp. 7–24, Jan. 1982.
- [49] B. Walter, S. R. Marschner, H. Li, and K. E. Torrance, "Microfacet models for refraction through rough surfaces," in *Proceedings of the 18th Eurographics Conference on Rendering Techniques*, ser. EGSR'07. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2007, pp. 195–206.
- [50] P. Beckmann, "Shadowing of random rough surfaces," *IEEE Transactions on Antennas and Propagation*, vol. 13, no. 3, pp. 384–388, May 1965.
- [51] B. T. Phong, "Illumination for computer generated pictures," *Commun. ACM*, vol. 18, no. 6, pp. 311–317, Jun. 1975.
- [52] H. E. Bennett and J. O. Porteus, "Relation between surface roughness and specular reflectance at normal incidence," *J. Opt. Soc. Am.*, vol. 51, no. 2, pp. 123–129, Feb 1961. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=josa-51-2-123>
- [53] K. Tanaka, N. Ikeya, T. Takatani, H. Kubo, T. Funatomi, and Y. Mukaigawa, "Time-resolved light transport decomposition for thermal photometric stereo," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.
- [54] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1. IEEE, 2005, pp. 886–893.
- [55] Y. Yang and D. Ramanan, "Articulated pose estimation with flexible mixtures-of-parts," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 1385–1392.
- [56] W. Jakob, "Mitsuba renderer," 2010, <http://www.mitsuba-renderer.org>.

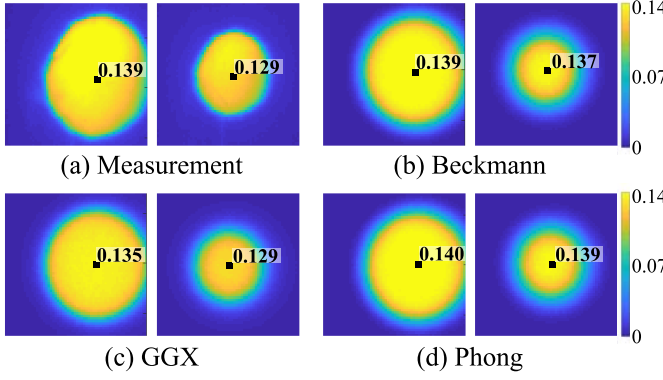


Fig. 11: **Comparing BRDF models with real measurements.** (a)-(d) show actual and fitted simulation measurements of reflection images of a disk heat source 12.7cm, 25.4cm away from the wall. (b), (d) Beckmann and Phong distribution model fails to explain intensity drop off near the center of the reflection image.

[57] <https://github.com/wg-perception/PartsBasedDetector>, [Online].

APPENDIX A: CHOICE OF ANALYTIC MODEL FOR BRDF

While the GGX specular BRDF model approximates the measured BRDF well, other specular models such as the Beckmann and Phong model also fit the measured angle-varying intensity. Fig. 11 compares actual measurements and fitted simulated measurements. Beckmann and Phong distributions fail to explain the intensity drop near the center of reflection image when a disk heat source moves further away from the wall. This motivated us to choose the GGX BRDF models for long-wave IR.

APPENDIX B: COOK-TORRANCE MODEL WITH GGX DISTRIBUTION

Eq. 7 is an analytic expression of Cook-Torrance BRDF model [48]. Here, we elaborate each term of this equation with GGX distribution [49]. The normal distribution function $D(\vec{h}, \vec{n})$, which is parameterized by 'roughness' α , describes distribution of microfacets on the surface:

$$D(\vec{h}, \vec{n}) = \frac{\alpha^2}{\pi((\vec{h} \cdot \vec{n})^2(\alpha^2 - 1) + 1)^2} \quad (18)$$

The shadowing factor $G(\vec{l}, \vec{v}, \vec{h}, \vec{n})$ describes shadowing of microfacets, and is also parameterized α with Smith approximation:

$$G(\vec{l}, \vec{v}, \vec{h}, \vec{n}) \approx G'(\vec{l}, \vec{n})G'(\vec{v}, \vec{n}), \text{ where} \\ G'(\vec{a}, \vec{n}) = \frac{2(\vec{a} \cdot \vec{n})}{(\vec{a} \cdot \vec{n}) + \sqrt{\alpha^2 + (1 - \alpha^2)(\vec{a} \cdot \vec{n})^2}} \quad (19)$$

The Fresnel function describes the reflection intensity based on the index of refraction. Because expression for the Fresnel function is complex, we refer readers to [48] for the details. For simplicity, we approximated Fresnel term as $F(\vec{v}, \vec{h}) \approx 1$ to fit the measurement for a narrow specular BRDF.

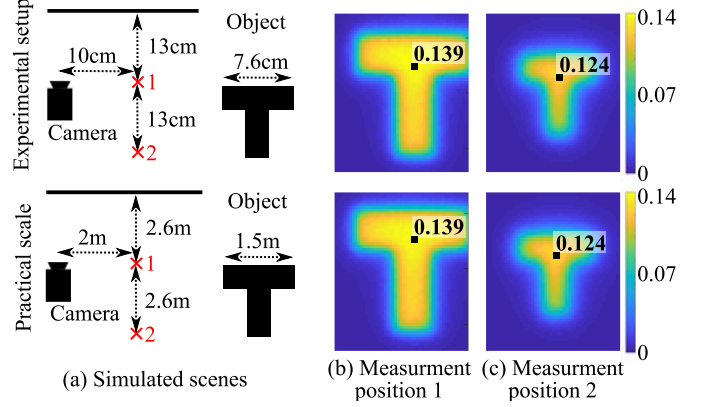


Fig. 12: **Our reconstruction algorithm scales up to practical scenarios.** (a) Simulation setup in the scale of our experimental setup and practical scene. (b)(c) shows simulated intensity measurements. When both objects and scene geometry are scaled by a same factor, measurement doesn't change. Hence, our reconstruction algorithm should be applicable for scenes at realistic scales.

APPENDIX C: SCALABILITY OF RECONSTRUCTION

Our experimental setup for BRDF-based NLOS scene reconstruction is in a small scale due to the availability of the controlled heat source. Section VI-A discusses that depth recoverability depends on the object's size. Here, we simulate a scene at two different scales – the scale of our experimental setup and realistic NLOS scenario. The simulated object temperature, ambient temperature, and reflectance were same as the experimental setup.

Fig. 12 shows that the intensity measurements do not change when both the object size and scene geometry are scaled by the same factor. This is because $\vec{l}$, $\vec{v}$ and $\vec{n}$ are preserved with the scaling. Hence, our reconstruction algorithm and experimental validation should scale up to realistic scenes as well.