

Shape from Heat Conduction

Sriram Narayanan¹, Mani Ramanagopal¹, Mark Sheinin²,
Aswin C. Sankaranarayanan¹, and Srinivasa G. Narasimhan¹

¹ Carnegie Mellon University, USA

² Weizmann Institute of Science, Israel

Abstract. Thermal cameras measure the temperature of objects based on radiation emitted in the infrared spectrum. In this work, we propose a novel shape recovery approach that exploits the properties of heat transport, specifically heat conduction, induced on objects when illuminated using simple light bulbs. Although heat transport occurs in the entirety of an object’s volume, we show a surface approximation that enables shape recovery and empirically analyze its validity for objects with varying thicknesses. We develop an algorithm that solves a linear system of equations to estimate the intrinsic shape Laplacian from thermal videos along with several properties including heat capacity, convection coefficient, and absorbed heat flux under uncalibrated lighting of arbitrary shapes. Further, we propose a novel shape from Laplacian objective that aims to resolve the inherent shape ambiguities by drawing insights from absorbed heat flux images using two unknown lights sources. Finally, we devise a coarse-to-fine refinement strategy that faithfully recovers both low- and high-frequency shape details. We validate our method by showing accurate reconstructions, to within an error of 1-2 mm (object size ≤ 13.5 cm), in both simulations and from noisy thermal videos of real-world objects with complex shapes and material properties including those that are transparent and translucent to visible light. We believe leveraging heat transport as a novel cue for vision can enable new imaging capabilities.

Keywords: Shape-from-X · Thermal Imaging · Heat Transport

1 Introduction

Estimating the 3D shape of an object from its images is central to computer vision. Traditional techniques for shape recovery such as photometric stereo [54], structured light [43] and structure from motion [10] rely on surface reflectance and often suffer when confronted with dark objects or those transparent to visible light. However, only a handful of works leverage the emissive properties of the surface [28, 29]. Every object with a temperature above absolute zero emits electromagnetic radiation and, for objects at room temperature, this primarily lies in the Long-Wave Infrared Spectrum (LWIR) detectable by thermal cameras. Further, temperature change within an object is governed by well-known laws of heat transfer physics related to conduction, convection, and radiation. A useful but often overlooked aspect of this heat transport is conduction, which is shape-dependent. This raises a fundamental question: *Can we infer an object’s shape by observing its heat flow?*

Recovering shapes in the thermal spectrum presents numerous advantages over traditional techniques that operate in the visible spectrum, particularly in reconstructing objects that are transparent or translucent to visible light (Fig. 1). However, modeling an object’s appearance in thermal wavelengths is complex due to its dependence on temperature, emissivity, and the properties of the surrounding medium. Accurately modeling thermal appearance requires considering light transport, such as reflections, and heat transport within and around the object. Unfortunately, inverting the full volumetric heat transport for an unknown scene and lighting is challenging, given only surface temperature measurements from a thermal camera.

In this work, we propose a surface approximation to volumetric heat transport that enables us to faithfully model heat flow across the objects’ extent. This approximation is accurate to within a few percent of the full model even for thick objects and we show it can be used to recover several material properties and surface geometry (Fig. 1). Our approximation modifies the heat conduction PDE [52] relating it to the object’s shape via a surface Laplace operator Δ_s . Interestingly, this results in a PDE that is non-linear to shape but linear with respect to the unknowns of the Laplace operator. Solving the linear system over time allows us to recover the Laplace operator and other properties such as absorbed heat flux, convection coefficient, and scaled heat capacity.

Estimating shape from Laplacian involves ambiguities due to the non-convex nature of the objective function, which has multiple global minima. Previous work has demonstrated effective shape recovery only when approximate initial shape estimates are employed [4]. In this paper, we demonstrate that the space of ambiguities in shape estimation is locally binary and that the camera ray constraint restricts the possible isometric deformations. We address this ambiguity using two uncalibrated light sources, leveraging the novel insight that absorbed heat flux resembles shading. Our approach is validated by recovering accurate shapes (avg. 1-2 mm error for object sizes ≤ 13.5 cm) of both simulated and real-world objects with complex geometries, varying thicknesses, and diverse visible reflectances using noisy temperature observations from a thermal camera.

In summary, we propose a novel shape-from-X technique using the first principles of heat conduction, which effectively recovers shapes for diverse visible reflectances while addressing ambiguities inherent in the shape from the Laplacian objective, albeit with known emissivity. We believe techniques utilizing thermal

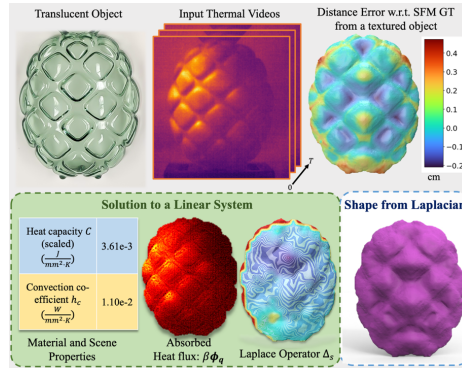


Fig. 1: A thermal video of a translucent object with heat conduction is used to estimate absorbed heat flux, heat capacity, convection coefficient, and Laplace operator Δ_s (shown as geodesic distance obtained from Δ_s) without knowledge of shape. The full shape is then estimated from the Laplace operator resolving ambiguities.

cameras can have a significant advantage over pure visible light methods due to passive sensing capabilities with complementary reflectance characteristics. With cheaper and better thermal cores, inverting the physics of thermal image formation can unlock new imaging capabilities that hold potential applications to robotics [45], manufacturing [16] and medical thermography [20].

2 Related Work

We briefly review the utility of thermal imaging in computer vision, the efforts in recovering shape from intrinsic attributes and relevant approaches in the physics-based shape-from-X literature.

Thermal imaging in computer vision. Recently, there has been considerable interest in utilizing the thermal modality to address diverse computer vision tasks [5, 15, 27, 38, 41, 45, 46, 48, 50, 56]. Moreover, research efforts have delved into the segmentation [21] and scanning [11, 32] of opaque glass objects in the thermal spectrum. Leveraging their passive imaging capabilities [51], thermal cameras have proven valuable for pose estimation [7, 22, 28, 29] and re-identification [14, 35, 53]. These cameras have also played a crucial role in material classification studies [25, 40, 42, 44]. Notably, Dashpute et al. [9] recently employed the heat equation to deduce the diffusivity and emissivity of objects with flat geometries.

Shape from intrinsic attributes. Intrinsic attributes, such as the Laplace operator, serve as the fundamental quantity in various physical simulations, including fluid dynamics, wave propagation, and heat transfer. These attributes facilitate the computation of geometrical measures such as lengths, areas, and volumes without requiring the object’s spatial embedding. Our work is aligned with previous efforts [18, 23] that addressed shape recovery from sound waves, concluding that, in two dimensions, multiple shapes could produce the same set of harmonics. Recent works [4, 8] have also explored recovering shapes given the intrinsic properties of objects but mention that the recovered shape highly depends on the initialization and the quality of the intrinsic attributes. It is important to note how our work differentiates on both these fronts, where initializations to our optimizations are flat planes compared to initial shape estimates in [4] and our estimate of intrinsic quantities comes from noisy temperature measurements. Moreover, [4] focuses on style transfer without needing to address non-rigid edge length preserving deformations by having good initial shape estimates.

Physics-based shape from X. Recovering shape has been of keen interest to the vision community with methods recovering shape from shading [13], polarization [24, 31], texture [12], specular flow [1] and others. While these methods rely on inputs from a visible camera, only a handful of works have explored the thermal regime. Nagase et al. [33] present a method to recover shape from multi-spectral thermal radiation based on the attenuation of light due to air. Tanaka et al. [49] show photometric stereo with thermal cameras under steady-state conditions (takes approx. ten minutes) but don’t use the knowledge of shape-dependent heat conduction. Additionally, their method depends on calibrated distant lighting. In contrast, our approach utilizes uncalibrated light sources for heat induction and uses quicker transient temperature measurements that takes 15 seconds or less.

3 Background

We introduce the heat transport equation and discuss the discrete Laplace operator that is necessary to computationally evaluate them.

Transient Heat Equation. Heat transfer in an object can occur due to three main phenomena, namely, conduction, convection, and radiation. While conduction can occur through the entire solid object, convection, and radiation are surface phenomena. Given a domain Ω , we can define its Laplace operator Δ that encodes shape information, and $\mathbb{1}_{\partial\Omega}$ is an indicator function that is 1 if the volume is exposed to the surface and 0 otherwise. Now, the heat transfer equation [3, 52] can be written as:

Table 1: Properties in the heat equation (Eq. 1) and their dimensions.

Property	Description	Unit
h_c	Convection co-efficient	$W/(m^2 \cdot K)$
ρ	Density	kg/m^3
ϵ	Emissivity	–
β	Energy Absorption factor	–
ϕ_q	Input heat flux density	W/m^2
A	Surface Area	m^2
c_p	Specific heat capacity	$J/(kg \cdot K)$
σ	Stefan-Boltzmann constant	$W/(m^2 K^4)$
T	Temperature	K
α	Thermal diffusivity	m^2/s

$$\frac{dT}{dt} = \alpha \Delta T + \mathbb{1}_{\partial\Omega} \frac{1}{\rho c_p} \frac{d}{dv} \left(\sigma \epsilon A (T_{surr}^4 - T^4) + h_c A (T_{surr} - T) + A \beta \phi_q \right) \quad (1)$$

The first term ΔT represents conduction, and the second term represents the sum of radiation, convection, and input heat source; other quantities are explained in Table 1 and a complete derivation is provided in the supplementary.

Discrete Laplace Operator. Evaluating the heat equation computationally requires discrete approximations of quantities in Eq. 1. Consider the domain Ω discretized into finite elements N , defined using V vertices and F faces. The discrete version of the Laplace operator Δ on the domain Ω is denoted by Δ_Ω . The Laplacian $\mathbf{L} \in \mathbb{R}^{|V| \times |V|}$ is a matrix that acts on a function $\mathbf{f} \in \mathbb{R}^{|V|}$ to compute its second derivative over Ω . This can be expressed of the form,

$$(\mathbf{L}\mathbf{f})_i = \sum_{j \in \mathcal{N}(i)} w_{ij} (f_j - f_i) \quad (2)$$

where, w_{ij} is the weight between the vertices i and j . Here, the sum is taken over the neighbors $\mathcal{N}$ of vertex i . If the finite elements are made of triangular mesh elements then the most commonly used weights for w_{ij} are cotangent weights [30, 37], defined as

$$L_{ij} = \begin{cases} w_{ij} = \frac{\cot \alpha_{ij} + \cot \beta_{ij}}{2} & \text{if } i \neq j \\ -\sum_{j \in \mathcal{N}(i)} w_{ij} & \text{if } i = j \end{cases} \quad (3)$$

where α_{ij} and β_{ij} are the angles opposite to the edge connecting the vertices i and j . To perform volumetric simulations of heat transport, we use tetrahedral

Laplacians based on cotangent weights that has a complex dependence with dihedral angles as mentioned in [2, 6]. The discrete Laplace operator is defined as $\Delta_\Omega = \mathbf{M}^{-1}\mathbf{L}$ where $\mathbf{M} \in \mathbb{R}^{|V| \times |V|}$ is a mass matrix [47]. A common choice for $\mathbf{M}$ in the case of triangular meshes is the lumped mass system $\mathbf{M}_a$ and a mass matrix $\mathbf{M}_v$ containing volumes in case of tetrahedral mesh elements.

4 Surface Approximation of Volumetric Heat Transfer

A thermal camera only provides temperature measurements of a surface based on the radiation emitted by it. Since all real-world objects have some thickness, inverting the heat transport in an unknown volume with unknown material properties, heated by an uncalibrated source and without interior temperature measurements is a daunting task. To tackle this, we propose an approximation for the volumetric heat flow that aids in shape recovery just from surface temperature measured using thermal cameras.

Discrete heat equation. Consider a temperature field $\{\mathbf{T}_k, k \in [0, \dots, K]\}$ on the domain Ω , observed with a sampling period t producing a total number of observations K . If the scene has a temporally constant surrounding temperature $\mathbf{T}_{surr}$ and an input heat flux ϕ_q is induced across the mesh surface then the heat equation (Eq. 1) can be discretized and written in the form,

$$\frac{\mathbf{T}_{k+1} - \mathbf{T}_k}{t} = \alpha \Delta_\Omega \mathbf{T}_{k+1} + \mathbb{1}_{\partial\Omega} \frac{\mathbf{M}_v^{-1}}{\rho c_p} \left(\sigma \epsilon \mathbf{M}_a (\mathbf{T}_{surr}^4 - \mathbf{T}_{k+1}^4) + h_c \mathbf{M}_a (\mathbf{T}_{surr} - \mathbf{T}_{k+1}) + \beta \mathbf{M}_a \phi_q \right) \quad (4)$$

Volumetric heat flow. To illustrate the impact of object thickness on volumetric heat transport, we conducted tetrahedral heat simulations using cuboidal shapes of varying thicknesses, as shown in Figure 2 (left). The objects are heated uniformly at each vertex with a 5 mW light source and a 15 second simulated video is recorded. The thicknesses range from 2 mm to 1 m, and we compared their video outputs to simulations with a small thickness of 1 mm. The simulations indicate that beyond a thickness near 1 cm, the difference between the thermal videos saturates for a wide range of materials (polystyrene, steel, glass, and brick) with varying thermal conductivity. In other words, the volumetric simulation does not change the surface temperature on the object appreciably after a certain thickness.

Surface approximated heat equation. Based on the above empirical evidence, we approximate the volumetric heat flow as a surface phenomenon assuming that the object is made of a finite volume with uniform/average thickness δ directly below it. This implies that the interior volumes are represented by nodes on the object’s surface. The mass matrix $\mathbf{M}_v$ representing this volume can then be approximated as $\mathbf{M}_v = \mathbf{M}_a \delta$. This approximation allows us to calculate

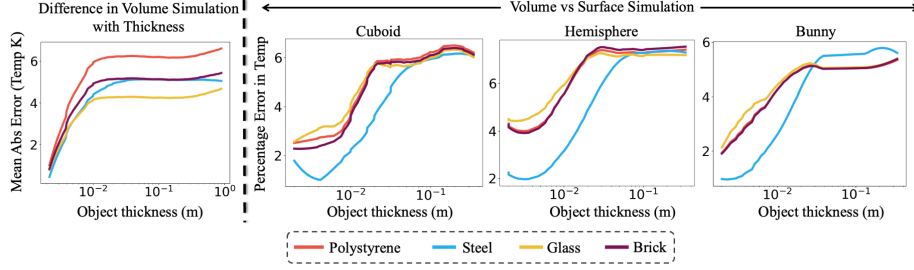


Fig. 2: The plot on the left shows average absolute error (across time) within thermal videos between tetrahedral simulations of varying thickness (0.001m to 1m) compared to the video from simulations of least thickness (0.001m). The plots on the right showcase the percentage error in temperature between true volumetric transport and its surface approximation with varying shapes and material properties for objects like cuboid, hemisphere and bunny shown up to a thickness of 20cm.

Laplacians with neighboring volumes using the cotangent approximation for triangular meshes (Eq. 3), instead of the more complex tetrahedral version. As we are interested in solving a linear system of equations in Section 5, we further linearize Eq. 1 using a first-order approximation for the radiation term as [52],

$$\sigma\epsilon\mathbf{M}_a(\mathbf{T}_{surr}^4 - \mathbf{T}^4) \approx 4\sigma\epsilon\mathbf{M}_a\mathbf{T}_{surr}^3(\mathbf{T}_{surr} - \mathbf{T}). \quad (5)$$

Substituting Eq. 5 in Eq. 4 and rearranging the terms we obtain,

$$\delta\rho c_p(\mathbf{T}_{k+1} - \mathbf{T}_k) - t\alpha\delta\rho c_p\tilde{\mathbf{M}}_v^{-1}\mathbf{L}\mathbf{T}_{k+1} - t\beta\phi_q + th_c(\mathbf{T}_{k+1} - T_{surr}) - 4t\sigma\epsilon(T_{surr}^4 - T_{surr}^3\mathbf{T}_{k+1}) = 0. \quad (6)$$

Here, the quantity $\tilde{\mathbf{M}}_v^{-1}\mathbf{M}_a = 1/\delta$ and the indicator function $\mathbb{1}_{\partial\Omega}$ has been dropped since we deal with only surfaces. Convection, radiation, and absorption are assumed to occur at all finite elements throughout the mesh. We use this surface approximation (Eq. 6) in our shape recovery and simulation experiments.

Validating surface approximation. We analyze the validity of Eq. 6 in Figure 2 (right) by calculating the percentage error in forward heat transport between tetrahedral (volumetric) simulations for varying object thicknesses and our corresponding approximated surface simulations. The percentage error in the thermal simulation videos is lower than 5% for thicknesses up to 1 cm and has an upper bound of 8% even for complex shapes with thicknesses up to 20 cm. Common objects such as cups, bowls, vessels and bottles typically have thicknesses much less than 1 cm, so our surface approximation is applicable for a diverse range of real-world objects. In Section 7, we show that this surface approximation is crucial to recover shapes, using both simulations and experiments with noisy thermal videos, resulting in average shape errors of less than 2 mm for objects of many shapes, thicknesses and materials.

5 Estimating Material and Geometric Properties

Optimizing for the shape information directly from the heat equation (Eq. 6) is non-convex and difficult with several unknown material properties. We now describe an approach to first estimate the shape-dependent Laplacian information, and while doing so, obtain several quantities such as absorbed heat flux, convection coefficient, and scaled heat capacity for objects with arbitrary shapes without the knowledge of their embedding in space. This is a notable contribution compared to works like [9] which estimate material properties for planar objects.

Our approach relies on some reasonable assumptions about the object and its environment. We assume that the object is made of a homogeneous material with properties such as specific heat c_p , density ρ , and thermal diffusivity α constant throughout the object. The emissivity ϵ of the object is known (they are typically above 0.9 for a range of painted objects) and is assumed to be independent of wavelength. The object is in a stable air environment at a known constant temperature T_{surr} with a convection coefficient h_c that is constant during the video capture process.

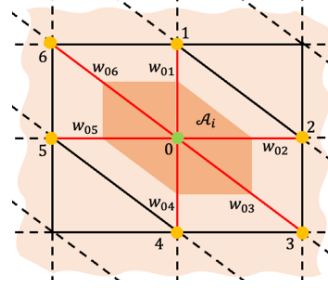
Minimum number of frames required. After linearizing the surface approximated heat equation (Eq. 6) in terms of temperature, we aim to provide theoretical insights into the constraints governing the minimum number of thermal image pairs that would be needed to recover unknown quantities. To start, the heat equation (6) can be re-written in the form,

$$\mathcal{E}(\mathbf{T}) \equiv \underline{\delta\rho c_p}(\mathbf{T}_{k+1} - \mathbf{T}_k) - t\underline{\alpha\delta\rho c_p}\tilde{\mathbf{M}}_v^{-1}\mathbf{L}\mathbf{T}_{k+1} - t\underline{\beta}\phi_q + \underline{th_c}(\mathbf{T}_{k+1} - T_{surr}) - 4t\sigma\epsilon(T_{surr}^4 - T_{surr}^3\mathbf{T}_{k+1}). \quad (7)$$

Here, the quantities underlined in orange indicate the unknown coefficients.

Analyzing the number of unknown coefficients across each term of Eq. 7, the first term provides a single unknown quantity $C \equiv \delta\rho c_p$ which is the heat capacity of the object per unit squared area. The second term $\Delta_s \equiv \alpha\delta\rho c_p\tilde{\mathbf{M}}_v^{-1}\mathbf{L}$ is a scaled version of the Laplace operator Δ_Ω . If mesh M had regular grid connectivity then the maximum valence of the surface is limited to six. Laplacian's non-diagonal elements will have non-zero entries only at the ij^{th} locations where an edge ij exists between the vertex v_i and v_j .

Further, the rest of the coefficients $\alpha\delta c_p\tilde{\mathbf{M}}_v^{-1}$ do not change the number of coefficients as the lumped mass matrix $\tilde{\mathbf{M}}_v$ contains non-zero entries only along the diagonal and other quantities are scalar values. Hence, a maximum of six unknown coefficients are introduced by the second term. Finally, the third and fourth terms provide two more unknown coefficients, namely, the absorbed heat flux $\beta\phi_q$ and the convection coefficient h_c . This results in a total of nine unknown coefficients per pixel. Evaluating Eq. 7 once, requires a pair of consecutive images



$\{k, k+1\}$. Thus, estimating the unknown quantities for every vertex individually would require at least ten consecutive image (video frame) pairs. We perform the following global optimization where quantities such as C and h_c are assumed spatially constant.

Optimization procedure. Given a thermal video $\{\mathbf{T}_k, k \in [1, \dots, K]\}$ and a 2D mask of the object, we create a flat triangular mesh M with V vertices and F faces within the mask where every pixel $p_i \in \{0, \dots, H \times W\}$ is associated to vertices $v_i \in V$. The residual of Eq. 7 is used as our objective function to perform a gradient descent minimization to obtain the unknown quantities. We remove the reflections and fixed pattern noise from the captured video by subtracting the first frame, but the system noise characterized by Noise Equivalent Temperature Difference (NETD) still persists. To address this, we allow the temperature to be optimized while remaining close to the observed values. This provides a noise-reduced estimate of temperature values allowing the objective to be minimized better. The overall cost function can be written as follows,

$$C, \Delta_s, h_c, \beta\phi_q, \hat{\mathbf{T}} = \underset{C, h_c, \Delta_s, \beta\phi_q, \hat{\mathbf{T}}}{\operatorname{argmin}} |\mathcal{E}(\hat{\mathbf{T}})| + \lambda_r \|\mathbf{T} - \hat{\mathbf{T}}\|^2 \quad (8)$$

here, λ_r is the weight for the regularization term and is set to 1.0 in our experiments. We use Adam optimizer [26] from Pytorch [36] for performing the optimization. This optimization estimates Laplace operator Δ_s that is crucial for shape recovery and other properties such as absorbed heat flux $\beta\phi_q$, convection coefficient h_c and the scaled heat capacity C of the material.

6 Shape from Laplace Operator

We now describe our approach towards computing shape from the estimated Laplace operator Δ_s . The estimated Δ_s contains an unknown scale factor α that was not estimated in the previous stage. Therefore, our shape estimates are inherently subject to an unknown scale. As recognized by [4], this optimization highly depends on the quality of the intrinsic quantities and initial shape estimates due to the possibility of non-rigid edge lengths preserving deformations. While [4] tackled challenges by focusing on the problem of style transfer with good initial shape estimates, it is crucial to note our estimates of the operator come from noisy temperature measurements captured using a thermal camera, and our initialization to the shape optimization is a flat plane.

6.1 Ambiguities in Shape

It is crucial to note that the Laplace operator is intrinsic in nature, i.e., it is independent of the specifics of the object’s embedding in 3D space. For example, the Laplace operator is invariant to the rigid body transformations of the object. Unfortunately, the equivalent class of shapes that have the same Laplace operator is quite large. We provide an intuitive characterization of this class using the notion of Gaussian curvature.

Gaussian Curvature is also an intrinsic property of a surface. At any point on the surface, it is given by the product of principal curvatures [19] $K = \kappa_1 \kappa_2$. In its discrete form, it is described as the deviation of vertex angle sum $\Theta_i = \sum \theta_i^k$ from the Euclidean angle sum 2π around a vertex. For a vertex with dual area $\mathcal{A}_i$ this is formally written as, $K_i = (2\pi - \Theta_i)/\mathcal{A}_i$. Concave and convex surfaces take values $K > 0$, developable surfaces have $K = 0$, and saddle points have $K < 0$.

Possible ambiguities in shape. A well-known fact in differential geometry is that the Laplace operator is fully determined by the discrete Riemannian metric up to a scale factor [17, 55]. Computing Gaussian curvatures from the Riemannian metric is straightforward, as it describes edge lengths in a triangular mesh. Knowing the sign of Gaussian curvature reduces the combination of principal curvature signs to two, implying, if $K > 0$ then $\kappa_1, \kappa_2 > 0$ or $\kappa_1, \kappa_2 < 0$. Obtaining an embedding by optimizing the least squares objective with respect to the reference Laplacian can yield one of two possible principal curvature directions based on the initialization. For smooth objects, this implies that the recovered shapes could either be locally convex or concave when $K > 0$ or one of two possible hyperbolic or developable surfaces when $K < 0$ or $K = 0$ respectively.

6.2 Resolving Shape Ambiguities

We introduce camera and normal constraints to resolve the above ambiguities.

Camera ray constraints. Every vertex v_i corresponding to pixel p_i must lie along the camera ray direction. To satisfy this constraint, we optimize depth $\mathbf{D}$ at every pixel p_i . The depth $\mathbf{D}$ is used to obtain the 3D vertex locations of the mesh along the camera rays. This avoids the problem of self-intersections during mesh optimization as camera rays don't intersect beyond the origin. Further, any isometric deformation is also constrained within the object's silhouette along the camera ray, significantly narrowing ambiguous deformations in the shape.

Normal constraints using absorbed heat flux. If all the active heating to the object is only due to light absorption then the recovered heat flux from Eq. 8 loosely resembles the shading (or irradiance) as shown in Fig. 3(left b). Normals that point toward the light source receive more energy (due to Lambert's cosine law) and hence absorb more energy compared to the ones that point away.

If ambiguous shapes have normals $\mathbf{n}_1, \mathbf{n}_2$, then their absorbed heat flux values $\beta\phi_{\mathbf{q}}$ for a light source $\mathbf{s}$ is going to be different. Knowing which side of the camera the light source was placed, looking at $\beta\phi_{\mathbf{q}}$ values within small patches can reveal if the retrieved shape has the right sign of principal curvatures κ_1, κ_2 . But we take an even simpler approach and collect two videos $\mathbf{T}^1, \mathbf{T}^2$ with light sources placed on opposite sides of the camera $\{s_x^1 < 0, s_x^2 > 0\}$ ³. Then a pixel-wise comparison of the normalized heat flux image can be used to assign a half-space

³ Ideally, four light sources are required to resolve the per-pixel ambiguity but we found that using two light sources to correct a few normals is sufficient for the optimization.

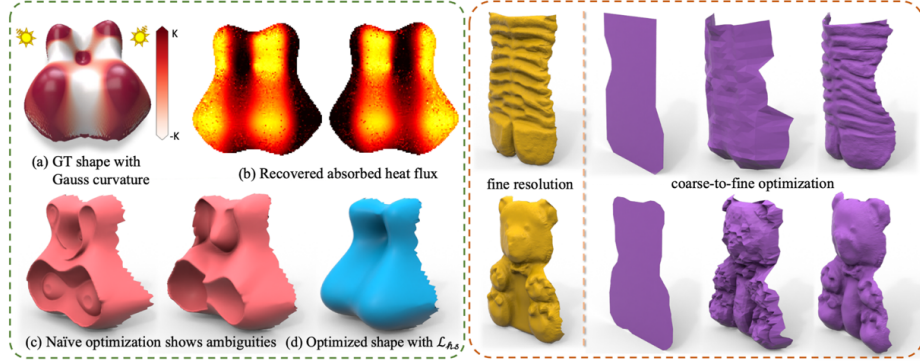


Fig. 3: Left figures illustrate that absorbed heat flux from light sources resembles shading information in visible images. Incorporating constraints like $\mathcal{L}_{hs}$ derived from heat flux images enhances shape recovery, whereas naive optimization leads to shape ambiguities. Figures on the right demonstrate the impact of coarse-to-fine optimization on retrieving low-frequency information, in comparison to optimization at fine resolution.

direction $\mathbf{N}_{hs}^x$ along an axis as follows,

$$\mathbf{N}_{hs}^x = \begin{cases} 1 & \text{if } \widehat{\beta\phi_q}^1 - \widehat{\beta\phi_q}^2 \geq \epsilon_{hs} \\ -1 & \text{if } \widehat{\beta\phi_q}^1 - \widehat{\beta\phi_q}^2 \leq -\epsilon_{hs} \end{cases}. \quad (9)$$

Here, $\widehat{\beta\phi_q}^1, \widehat{\beta\phi_q}^2$ represent normalized heat flux images as shown in Figs. 3b on the left panel. Only pixels that have ϵ_{hs} difference in the normalized heat flux image get assigned a half-space direction. The following hinge penalty is added to nudge the optimization towards shapes that have components of normals along the right direction,

$$\mathcal{L}_{hs} = \frac{1}{V} \sum \max(-\mathbf{n}_x * \mathbf{N}_{hs}^x, 0). \quad (10)$$

Here, a vertex will only incur a loss when the $\text{sign}(n_x)$ does not obey $\text{sign}(n_{hs}^x)$. Please note that in theory, one would require four videos to account for pixels that do not have a normal component along the axis of lighting. In practice, we found that only a few normals have to be corrected for the entire shape to snap in the correct convex/concave configuration.

Optimization procedure. Our objective for shape from the Laplace operator is:

$$\mathbf{D}^* = \underset{\mathbf{D}}{\text{argmin}} \mathcal{E}(\hat{\mathbf{T}}) + \lambda_{sm} \mathcal{L}_{sm} + \lambda_{hs} \mathcal{L}_{hs} + \lambda_{sy} \mathcal{L}_{sy}. \quad (11)$$

Here, $\mathcal{L}_{sm}$ is the normal consistency smoothness [39], $\mathcal{L}_{sy}$ is symmetry loss along an axis used in the coarsest stage of our optimization for regularization and $\lambda_{sm}, \lambda_{hs}, \lambda_{sy}$ are weights to the respective loss terms. Typically, our optimization takes less than a minute, using a AdamUniform optimizer from Nicolet et al. [34] which diffuses the gradients over the entire domain allowing for a smoother optimization of shapes.

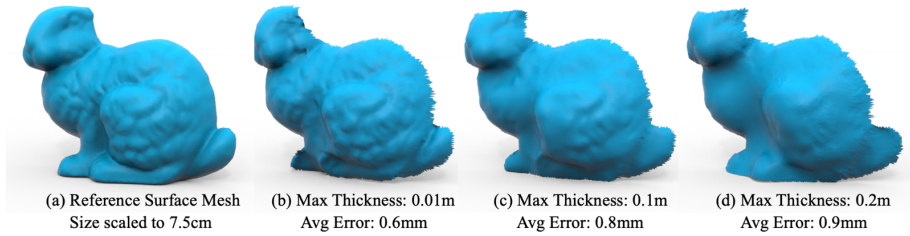


Fig. 4: Accuracy of shape recovery using simulations for the Stanford Bunny of varying thicknesses. Notice the shape details (bumps, normal discontinuities, crevices, etc.) recovered on the Bunny even with a thickness of 1 cm, and good overall shape (albeit with reduced detail) even for large thicknesses of 10 cm and 20 cm, demonstrating the effectiveness of our surface approximation and optimization method.

6.3 Coarse-to-Fine Refinement

To enhance our reconstruction results we further implement a hierarchical refinement approach as shown in Fig. 3 (right) Given the inherent noise in thermal videos, the accurate reconstruction of low-frequency information at finer resolutions poses a challenge. The Laplacian being a high-pass filter with a response that is quadratic in frequency, primarily captures local shape details. Incorporating this localized information at finer resolutions hinders the faithful reconstruction of low-frequency features. Therefore, we initially reduce the image resolution through downsampling and conduct reconstructions using a coarser mesh. The estimated coarse mesh is then used as an initial guess for further refinement.

7 Experimental Results

Hardware setup. We use a FLIR Boson thermal camera that is radiometrically calibrated and has a pixel resolution of 512×640 . The camera is operated in TStable mode with radiometry enabled. Further, we place two light sources at uncalibrated locations and capture two 15-second videos, each with one light source active at a time. Before capture, we do a Flat Field Correction (FFC) and disable FFC for the video’s remaining duration. The first ten video frames are used for removing reflections and fixed pattern noise, while our method uses the remaining footage at 6FPS.

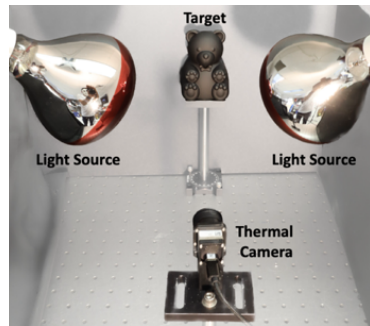


Fig. 5: Setup where the object is illuminated with incandescent lights

Simulation results with varying thicknesses. Figure 4 shows shape recovery results from thermal videos obtained from volumetric simulations of heat transport, while our shape optimization uses the approximated heat equation (Eq. 7). We verify that our surface approximation can faithfully recover shape

Table 2: Average absolute shortest distance error in mm w.r.t. Kinect and SFM scans. The gross heights of these objects are 6 cm, 13.5 cm, 10.5 cm, 10.5 cm, and 10.2 cm for the bears respectively. The low errors show the effectiveness of our surface model, optimization, coarse-to-fine refinement, and ambiguity resolution strategy.

	Ball	Soda-Can	Bottle-1	Bottle-2	Bear-black	Bear-Tex.	Bear-Transp.
Kinect	0.7	1.3	2.0	2.2	2.8	2.8	3.0
SFM	1.2	0.8	1.2	1.3	1.4	1.6	1.7

details up to a object thickness of 1 cm and maintains overall shape, albeit with reduced detail even at 20 cm thickness.

Details on real-world objects. We evaluated the effectiveness of our method using a variety of real-world objects that exhibit diverse shapes, including intricate ridges and grooves, as shown in Figures 6 and 7. These objects are made of different materials: acrylic plastic for the ball and the bear, PET plastic for packaged bottles, aluminum alloys for the soda can, and plastic filament for the 3D-printed bunny. The objects in Figure 6 were painted black to enhance light absorption, with an approximately known emissivity. In contrast, Figure 7 shows the effectiveness of our method with varying visible reflectances. The objects also vary in thickness, and the bunny, the thickest among them, is 4.5 cm wide, having a 2mm shell thickness with honeycomb structures inside for solid geometry.

Qualitative results. The reconstructed shape outputs are displayed in the third and fourth columns of Figures 6 and 7. The reconstructions faithfully recover the overall object geometry, including high-frequency details such as the grooves in the bottles and the bow tie on the bear. In addition, Figure 7 shows reconstructions of identically shaped bears painted black, painted to have strong texture, and without any paint (i.e., transparent). These results shows that our method works across a wide range of visible BRDFs, including objects with *spatially-varying albedo*, *translucent* (Fig. 1), and *transparent* appearances. Note that shape recovery for such dark, transparent, or translucent objects fails terribly with SFM or Kinect scans, highlighting the benefits of capturing surface emission using thermal cameras. Although our results are far from perfect due to poor signal-to-noise ratio from thermal cameras, we believe that the obtained results have significantly higher quality compared to previous efforts towards shape recovery [33, 49] using thermal cameras (see Fig. 13 in [49], Fig. 7 in [33]).

Quantitative comparisons. The last column in Figures 6 and 7 show the signed distance error for the obtained reconstructions compared to the ground truth dense SFM scans for these objects painted with dense textures (Table 2 right). Comparisons are done after ICP alignment of obtained reconstructions with their respective SFM GTs. For the 3D printed bunny we make comparisons with the true 3D model. Table 2 reports the absolute average closest-distance error (unsigned) compared with dense SFM and also scans with Kinect. Our

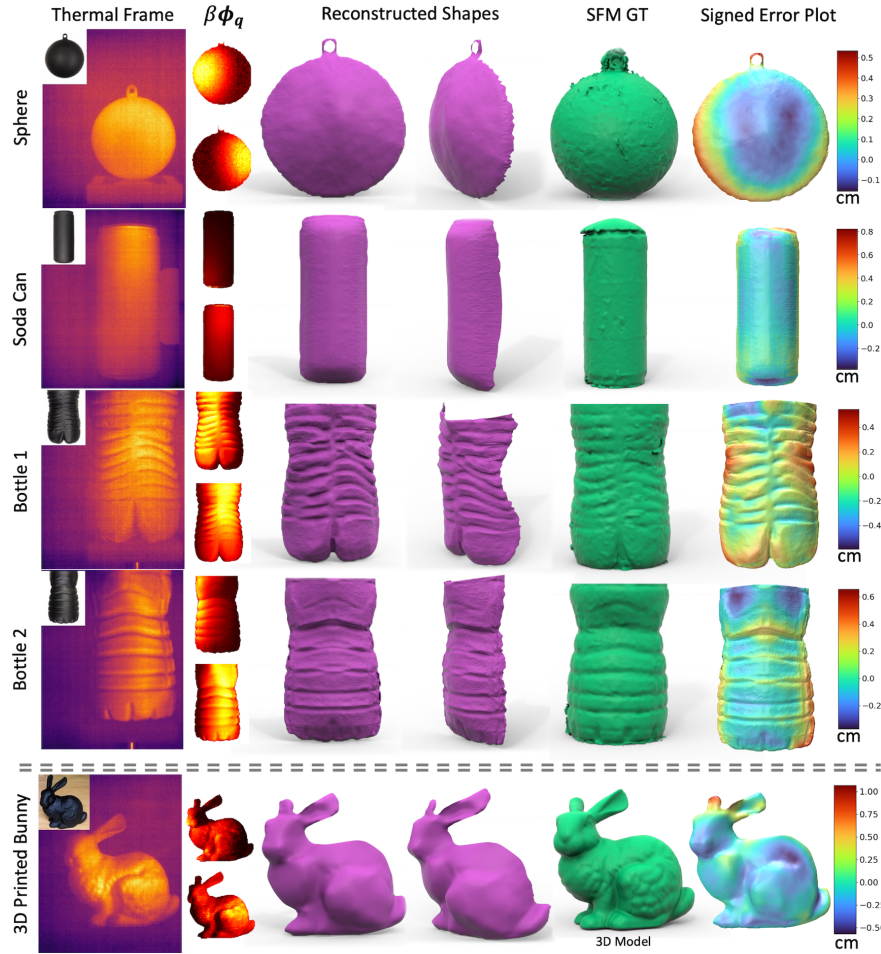


Fig. 6: Shape optimization results for real-world objects with thermal video of heat conduction as inputs. The second column shows the absorbed heat flux at every vertex location obtained in Sec. 5, visualized in the form of an image for video inputs with light sources placed at two sides of the camera. The subsequent columns show reconstructed shapes from flat mesh initializations obtained by following the optimization procedure in Sec. 6. The last column shows signed distance error of the reconstructions with their corresponding SFM scans (computed using 50 views) shown in the previous column after ICP alignment (absolute errors in Table 2). We also show shape recovery with a 3D printed bunny which achieves an mean absolute (unsigned) error of 1.4mm, albeit with less surface detail because of the dense internal honeycomb structures.

reconstructions obtain an average distance error of less than 1.8 mm with the SFM scans for objects as tall as 13.5 cm. The method also achieves a mean absolute error of 1.4mm when compared to the 3D printed Bunny. These quantitative comparisons strongly validate our model and approach for 3D shape recovery. In all cases, the overall structure is faithfully recovered with mean absolute reconstruction errors (MAE) of less than 1.8 mm.

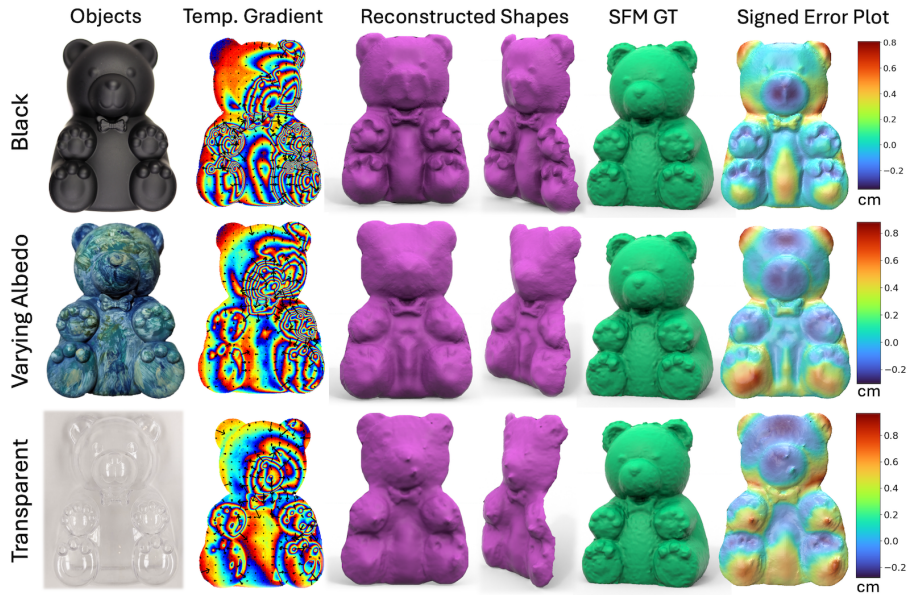


Fig. 7: Versatility of our approach in reconstructing various object materials (transparent, textured, homogeneous & black). The second column shows iso-contour lines of temperature with spatial gradient direction (black arrows) indicating observable heat flow in the thermal spectrum for varying visible reflectance properties. Notably, the reflectance properties of many visibly transparent objects appear opaque in thermal enabling heat flow observations in such transparent or translucent objects. Tab. 2 shows the absolute average distance error for these reconstructions.

Limitations. Materials with high thermal conductivity require light sources with high spatial frequency for good SNR due to rapid heat distribution. Low thermal conductivity materials can use lower-frequency sources but need longer heating. Further, our method assumes known emissivity, which varies with material properties, surface roughness, and viewing angle, affecting the temperature measurement. Please refer to the supplementary material for further discussion.

8 Conclusion

Heat conduction happens all around us, from pouring hot coffee into a mug to walking barefoot on the floor, and it is not just limited to light absorption. Our work takes a step forward in using this novel cue to extract shape information from thermal videos. To our knowledge, this is the first attempt to recover the intrinsic surface Laplace operator from noisy thermal data and extract shape using the first principles of heat conduction. We derived and validated a surface approximation to the volumetric heat transport equation for objects with varying thicknesses and material properties. We resolve the ambiguity in shape recovery and demonstrate accurate reconstructions for complex geometries with varying visible BRDFs. As thermal cameras become cheaper and more sensitive, exploring heat conduction as a novel cue is an exciting research avenue.

Acknowledgements: This work was partly supported by NSF grants IIS-2107236, CCF-1730147, and NSF-NIFA AI Institute for Resilient Agriculture. Thanks to Keenan Crane for useful discussions.

References

1. Adato, Y., Vasilyev, Y., Zickler, T., Ben-Shahar, O.: Shape from specular flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **32**(11), 2054–2070 (2010). <https://doi.org/10.1109/TPAMI.2010.126>
2. Alexa, M., Herholz, P., Kohlbrenner, M., Sorkine-Hornung, O.: Properties of Laplace Operators for Tetrahedral Meshes. *Computer Graphics Forum* (2020). <https://doi.org/10.1111/cgf.14068>
3. Bergman, T.L.: *Introduction to Heat Transfer*. Wiley (2011)
4. Boscaini, D., Eynard, D., Kourounis, D., Bronstein, M.M.: Shape-from-operator: Recovering shapes from intrinsic operators. *Computer Graphics Forum* (2), 265–274. <https://doi.org/https://doi.org/10.1111/cgf.12558>
5. Brahmabhatt, S., Ham, C., Kemp, C.C., Hays, J.: ContactDB: Analyzing and predicting grasp contact via thermal imaging. In: *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (6 2019)
6. Calcagni, G., Oriti, D., Thürigen, J.: Laplacians on discrete and quantum geometries. *Classical and Quantum Gravity* **30**(12), 125006 (May 2013). <https://doi.org/10.1088/0264-9381/30/12/125006>
7. Chen, I.C., Wang, C.J., Wen, C.K., Tzou, S.J.: Multi-person pose estimation using thermal images. *IEEE Access* **8**, 174964–174971 (2020). <https://doi.org/10.1109/ACCESS.2020.3025413>
8. Chern, A., Knöppel, F., Pinkall, U., Schröder, P.: Shape from metric. *ACM Trans. Graph.* **37**(4) (jul 2018). <https://doi.org/10.1145/3197517.3201276>
9. Dashpute, A., Saragadam, V., Alexander, E., Willomitzer, F., Katsaggelos, A., Veeraraghavan, A., Cossairt, O.: Thermal spread functions (tsf): Physics-guided material classification. *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (Jun 2023). <https://doi.org/10.1109/cvpr52729.2023.00164>
10. Dellaert, F., Seitz, S., Thorpe, C., Thrun, S.: Structure from motion without correspondence. In: *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662)*. vol. 2, pp. 557–564 vol.2 (2000). <https://doi.org/10.1109/CVPR.2000.854916>
11. Eren, G., Aubreton, O., Meriaudeau, F., Secades, L.S., Fofi, D., Naskali, A.T., Truchetet, F., Ercil, A.: Scanning from heating: 3d shape estimation of transparent objects from local surface heating. *Opt. Express* **17**(14), 11457–11468 (Jul 2009). <https://doi.org/10.1364/OE.17.011457>
12. Forsyth, D.A.: Shape from texture without boundaries. In: *Proceedings of the 7th European Conference on Computer Vision-Part III*. p. 225–239. *ECCV '02*, Springer-Verlag, Berlin, Heidelberg (2002)
13. Frankot, R., Chellappa, R.: A method for enforcing integrability in shape from shading algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **10**(4), 439–451 (1988). <https://doi.org/10.1109/34.3909>
14. Fu, C., Hu, Y., Wu, X., Shi, H., Mei, T., He, R.: Cm-nas: Cross-modality neural architecture search for visible-infrared person re-identification. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. pp. 11823–11832 (October 2021)

15. Gan, L., Lee, C., Chung, S.J.: Unsupervised rgb-to-thermal domain adaptation via multi-domain attention network. In: 2023 IEEE International Conference on Robotics and Automation (ICRA). IEEE (May 2023). <https://doi.org/10.1109/icra48891.2023.10160872>
16. Garrido, I., Lagüela, S., Otero, R., Arias, P.: Thermographic methodologies used in infrastructure inspection: A review—post-processing procedures. *Applied Energy* **266**, 114857 (2020). <https://doi.org/10.1016/j.apenergy.2020.114857>
17. Goes, F.d., Memari, P., Mullen, P., Desbrun, M.: Weighted triangulations for geometry processing. *ACM Trans. Graph.* **33**(3) (jun 2014). <https://doi.org/10.1145/2602143>
18. Gordon, C., Webb, D.L., Wolpert, S.: One cannot hear the shape of a drum. *Bulletin of the American Mathematical Society* **27**(1), 134–138 (1992). <https://doi.org/10.1090/s0273-0979-1992-00289-6>
19. Gray, A.: *Modern Differential Geometry of Curves and Surfaces with Mathematica*. CRC Press, Inc., USA, 1st edn. (1996)
20. Hildebrandt, C., Raschner, C., Ammer, K.: An overview of recent application of medical infrared thermography in sports medicine in austria. *Sensors* **10**(5), 4700–4715 (May 2010). <https://doi.org/10.3390/s100504700>
21. Huo, D., Wang, J., Qian, Y., Yang, Y.H.: Glass segmentation with rgb-thermal image pairs. *Trans. Img. Proc.* **32**, 1911–1926 (jan 2023). <https://doi.org/10.1109/TIP.2023.3256762>
22. Iwasawa, S., Ebihara, K., Ohya, J., Morishima, S.: Real-time estimation of human body posture from monocular thermal images. In: *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. pp. 15–20 (1997). <https://doi.org/10.1109/CVPR.1997.609290>
23. Kac, M.: Can one hear the shape of a drum? *The American Mathematical Monthly* **73**(4P2), 1–23 (1966). <https://doi.org/10.1080/00029890.1966.11970915>
24. Kadambi, A., Taamazyan, V., Shi, B., Raskar, R.: Polarized 3d: High-quality depth sensing with polarization cues. In: *2015 IEEE International Conference on Computer Vision (ICCV)*. pp. 3370–3378 (2015). <https://doi.org/10.1109/ICCV.2015.385>
25. Kerr, E., McGinnity, T., Coleman, S.: Material classification based on thermal properties — a robot and human evaluation. In: *2013 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. pp. 1048–1053 (2013). <https://doi.org/10.1109/ROBIO.2013.6739602>
26. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization (2014)
27. Kütük, Z., Algan, G.: Semantic segmentation for thermal images: A comparative survey. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 286–295 (2022)
28. Liu, R., Vondrick, C.: Humans as light bulbs: 3d human reconstruction from thermal reflection. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 12531–12542 (June 2023)
29. Maeda, T., Wang, Y., Raskar, R., Kadambi, A.: Thermal non-line-of-sight imaging. In: *2019 IEEE International Conference on Computational Photography (ICCP)*. pp. 1–11 (2019). <https://doi.org/10.1109/ICCPHOT.2019.8747343>
30. Meyer, M., Desbrun, M., Schröder, P., Barr, A.H.: Discrete differential-geometry operators for triangulated 2-manifolds. In: Hege, H.C., Polthier, K. (eds.) *Visualization and Mathematics III*. pp. 35–57. Springer Berlin Heidelberg, Berlin, Heidelberg (2003)
31. Miyazaki, Tan, Hara, Ikeuchi: Polarization-based inverse rendering from a single view. In: *Proceedings Ninth IEEE International Conference on Computer Vision*. pp. 982–987 vol.2 (2003). <https://doi.org/10.1109/ICCV.2003.1238455>

32. Miyazaki, D., Saito, M., Sato, Y., Ikeuchi, K.: Determining surface orientations of transparent objects based on polarization degrees in visible and infrared wavelengths. *J. Opt. Soc. Am. A* **19**(4), 687–694 (Apr 2002). <https://doi.org/10.1364/JOSAA.19.000687>
33. Nagase, Y., Kushida, T., Tanaka, K., Funatomi, T., Mukaigawa, Y.: Shape from thermal radiation: Passive ranging using multi-spectral lwir measurements. In: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 12651–12661 (2022). <https://doi.org/10.1109/CVPR52688.2022.01233>
34. Nicolet, B., Jacobson, A., Jakob, W.: Large steps in inverse rendering of geometry. *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia)* **40**(6) (Dec 2021). <https://doi.org/10.1145/3478513.3480501>
35. Park, H., Lee, S., Lee, J., Ham, B.: Learning by aligning: Visible-infrared person re-identification using cross-modal correspondences. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp. 12046–12055 (October 2021)
36. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., Chintala, S.: Pytorch: An imperative style, high-performance deep learning library (2019)
37. Pinkall, U., Polthier, K.: Computing discrete minimal surfaces and their conjugates. *Experimental Mathematics* **2**(1), 15 – 36 (1993)
38. Ramanagopal, M., Narayanan, S., Sankaranarayanan, A.C., Narasimhan, S.G.: A theory of joint light and heat transport for lambertian scenes. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 11924–11933 (June 2024)
39. Ravi, N., Reizenstein, J., Novotny, D., Gordon, T., Lo, W.Y., Johnson, J., Gkioxari, G.: Accelerating 3d deep learning with pytorch3d. arXiv:2007.08501 (2020)
40. Riba, J.R., Canals, T., Cantero, R.: Recovered paperboard samples identification by means of mid-infrared sensors. *IEEE Sensors Journal* **13**(7), 2763–2770 (2013). <https://doi.org/10.1109/JSEN.2013.2257943>
41. Rivadeneira, R.E., Sappa, A.D., Vintimilla, B.X., Kim, J., Kim, D., Li, Z., Jian, Y., Yan, B., Cao, L., Qi, F., Wang, H., Wu, R., Sun, L., Zhao, Y., Li, L., Wang, K., Wang, Y., Zhang, X., Wei, H., Lv, C., Sun, Q., Tian, X., Jia, Z., Hu, J., Wang, C., Zhong, Z., Liu, X., Jiang, J.: Thermal image super-resolution challenge results - pbvs 2022. In: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). pp. 417–425 (2022). <https://doi.org/10.1109/CVPRW56347.2022.00057>
42. Salamati, N., Fredembach, C., Süssstrunk, S.: Material classification using color and nir images. *Color and Imaging Conference* **17** (01 2009). <https://doi.org/10.2352/CIC.2009.17.1.art00040>
43. Salvi, J., Fernandez, S., Pribanic, T., Llado, X.: A state of the art in structured light patterns for surface profilometry. *Pattern Recognition* **43**(8), 2666–2680 (2010). <https://doi.org/10.1016/j.patcog.2010.03.004>
44. Saponaro, P., Sorensen, S., Kolagunda, A., Kambhamettu, C.: Material classification with thermal imagery. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 4649–4656 (2015). <https://doi.org/10.1109/CVPR.2015.7299096>
45. Sheinin, M., Sankaranarayanan, A., Narasimhan, S.G.: Projecting trackable thermal patterns for dynamic computer vision. In: Proc. IEEE/CVF CVPR (2024)

46. Shin, U., Park, J., Kweon, I.S.: Deep depth estimation from thermal image. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 1043–1053 (June 2023)
47. Solomon, J., Crane, K., Vouga, E.: Laplace-beltrami: The swiss army knife of geometry processing. In: Symposium on Geometry Processing Graduate School. Cardiff, UK (2014)
48. Srinivasan Ramanagopal, M., Zhang, Z., Vasudevan, R., Johnson Roberson, M.: Pixel-wise motion deblurring of thermal videos. In: Robotics: Science and Systems XVI. RSS2020, Robotics: Science and Systems Foundation (Jul 2020). <https://doi.org/10.15607/rss.2020.xvi.022>
49. Tanaka, K., Ikeya, N., Takatani, T., Kubo, H., Funatomi, T., Ravi, V., Kadambi, A., Mukaigawa, Y.: Time-resolved far infrared light transport decomposition for thermal photometric stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **43**(6), 2075–2085 (2021). <https://doi.org/10.1109/TPAMI.2019.2959304>
50. Tang, Z., Ye, W., Ma, W.C., Zhao, H.: What happened 3 seconds ago? inferring the past with thermal imaging. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 17111–17120 (June 2023)
51. Treptow, A., Cielniak, G., Duckett, T.: Real-time people tracking for mobile robots using thermal vision. *Robotics and Autonomous Systems* **54**(9), 729–739 (2006). <https://doi.org/10.1016/j.robot.2006.04.013>, selected papers from the 2nd European Conference on Mobile Robots (ECMR '05)
52. Vollmer, M., Mllmann, K.P.: Some Basic Concepts in Heat Transfer, chap. 4, pp. 351–392. John Wiley & Sons, Ltd (2017). <https://doi.org/10.1002/9783527693306.ch4>
53. Wei, Z., Yang, X., Wang, N., Gao, X.: Syncretic modality collaborative learning for visible infrared person re-identification. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp. 225–234 (October 2021)
54. Woodham, R.J.: Photometric Method for Determining Surface Orientation from Multiple Images, p. 513–531. MIT Press, Cambridge, MA, USA (1989)
55. Zeng, W., Guo, R., Luo, F., Gu, X.: Discrete heat kernel determines discrete riemannian metric. *Graph. Models* **74**(4), 121–129 (jul 2012). <https://doi.org/10.1016/j.gmod.2012.03.009>
56. Zhang, P., Zhao, J., Wang, D., Lu, H., Ruan, X.: Visible-thermal uav tracking: A large-scale benchmark and new baseline. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 8886–8895 (June 2022)