Dual Band Thermal Videography: Separating Time-Varying Reflection and Emission Near Ambient Conditions

Sriram Narayanan Mani Ramanagopal Srinivasa G. Narasimhan

Carnegie Mellon University

Abstract

Long-wave infrared radiation captured by a thermal camera includes (a) emission from an object governed by its temperature and emissivity, and (b) reflected radiation from the surrounding environment. Separating these components is a long-standing challenge in thermography. Even when using multiple bands, the problem is under-determined without priors on emissivity. This difficulty is amplified in near ambient conditions, where emitted and reflected signals are of comparable magnitude. We present a dual-band video thermography framework that reduces this ambiguity by combining two complementary ideas at a per-pixel level: (i) spectral cues (ratio of emissivity between bands is unknown but fixed), and (ii) temporal cues (object radiation changes smoothly while background radiation changes rapidly). We derive an image formation model and an algorithm to jointly estimate the object's emissivity at each band, and the time-varying object and background temperatures. Experiments with calibrated and uncalibrated emissivities in everyday scenes (e.g., coffee pot heating up, palm print on mirrors, reflections of moving people), demonstrate robust separation and recovery of temperature fields.

1. Introduction

Long-Wave Infrared (LWIR, 8-14 μm) radiation received by a thermal camera consists of two components: (1) the reflection or transmission of radiation originating from the surrounding environment, akin to traditional light transport, and (2) the object's own thermal emission, governed by internal heat transport and heat exchange with its surroundings. Accurately disentangling these factors is crucial to correctly interpret thermal imagery for a range of scientific and engineering applications [2, 41].

In general, separating the reflected and emitted components of thermal radiation is a highly under-constrained problem [15]. As a result, most prior approaches assume that one of the components is negligible or known *a priori* [5, 47] . For example, in industrial scenarios like welding and defect inspection, objects are significantly hotter than their surroundings, allowing background component to be neglected [17, 40]. Other studies employ controlled environments where the background component is spatially uniform or measured ahead of time [27] or inferred from polarization [37]. Some methods assume stationary backgrounds [23, 38] or steady-state object emissions that can be subtracted out [20]. However, most scenes of relevance to computer vision are near-ambient conditions where the reflected and emitted components are comparable in magnitude, time-varying, and spatially complex, making such assumptions unreliable or impractical.

In this work, we present a novel approach to disentangle the time-varying reflected and emitted radiation components using thermal videos captured within two sub-bands of LWIR. Our method builds on two key insights: (1) an object's spectral emissivity can vary with wavelength even within LWIR - contrary to the conventional "gray-body" assumption [2] — and (2) heat transport within objects evolves smoothly over time, while background changes abruptly.

We derive the dual-band image formation model as a linear combination of the reflection and emission components. Given thermal videos captured with two spectral bandpass filters, we consider both calibrated and uncalibrated scenarios. In the calibrated setting, we estimate material emissivity *a priori* from contact measurements of temperature using thermocouples where the background is a blackbody at known temperature. We then use the calibrated emissivity to separate reflection and emission in new scenes with those materials. To address the ill-posed uncalibrated case, our optimization exploits the differences in temporal variations in reflection vs. emission using a novel reconstruction loss to jointly estimate object emissivity and time varying object temperature and background radiation.

We validate our calibrated and uncalibrated approaches through both simulation and real experiments. We show that our calibrated emissivities of materials match those in the ECOSTRESS spectral library [3, 21]. We also show that our results match the reflector approach recommended

by FLIR Systems [10] for higher emissivity materials and improve on that approach for lower emissivity materials. Using simulations, we compare our uncalibrated method against existing baselines [2, 9, 39] across diverse materials under varying noise conditions.

Finally, we demonstrate the separation of dynamic emitted and reflected thermal components in real-world videos. Using dual-band thermal imaging, we recover emission and reflection videos in challenging scenes. The slow emission due to heat transport from a coffee pot containing a hot liquid is separated from reflections of people moving near it. A glass plate heated by a hot air gun increases its emission and is separated from a reflection of text written on a surface with cold and wet fingers. A palm print dissipating on a cold bulb is separated from higher reflection intensities due to a person. Please watch the supplementary video for the best visualizations. Our results show that thermal imaging, when analyzed spectrally and temporally, can reveal rich scene information inaccessible to visible-light cameras.

2. Related Work

Point measurements: Separating reflected and emitted radiation by an object has a long history [8, 9, 27, 37, 45, 46] but with no universally accepted technique applicable in general situations [2, 15]. The survey in [2] concludes that the best works are still the early ones performing dual [39] and multi-band [30] measurements. However, these methods make a "gray body" assumption and focus on measurements at a single point and do not demonstrate spatially varying signals of complex scenes relevant to computer vision. Our method outperforms such approaches.

Pyrometry: Our work differs from studies in pyrometry [2, 26], remote sensing [31], and related non-contact thermometry [5, 47] in several aspects. These methods often assume single, dual or multi-band measurements with either known/parameterized emissivity or conditions where emission dominates. Many works focus on hotter targets (200–2000 °C) and MWIR sensing where reflectivity is lower [2, 17, 40, 41]. We instead operate near ambient temperatures with LWIR (8–14 µm) sensors, where reflectance is higher and background radiation strongly contaminates measurements. Our method explicitly models and compensates for temporally varying background radiation.

Reducing reflections and enhancing contrast: Recent approaches reduce reflections using polarization [37] or enhance image contrast using priors [4, 9, 24, 27, 48]. For e.g., Blackbody Channel Prior (BCP) [9] boosts contrast assuming locally brightest pixels approximate blackbody emitters. However, BCP's assumption breaks for cold objects near ambient temperatures and where both object and background vary over time. In contrast, our work is focused on physically-based separation of reflection and emission components that are both dynamic near ambient conditions. We demonstrate that our approach performs better than BCP.

Physics Based Vision with Thermal: The use of thermal cameras for vision tasks is gaining attention, particularly those exploiting the dominant radiative components, either reflection [11, 20] or emission [6, 22, 23, 28, 38] enabling novel ways to solve many vision tasks. In particular, Dashpute et al. [6] recover properties such as emissivity and diffusivity but the method requires access to surface temperatures and ignores the influence of background. Similarly, many works use the emitted or reflected light to recover material or geometric properties such as shape [14, 16, 22, 23], albedo [28, 44], pose [20] and trajectory [33] but assume a thermally static background or known object temperatures and hence can subtract the first frame to ignore the effect of reflections [38] or ambient emissions. [13] train a neural network to estimate transparent shape from computed degree of polarization in the presence of both reflection and emission. But they avoid reflections of backgrounds (e.g. people) similar in magnitude to the object temperatures.

Light Transport Decomposition: Our research is similar in spirit to many physically based works that separate different light transport components including diffuse vs. specular [7, 12, 19, 35], reflection vs. transmission [18, 34, 42, 43], reflectance vs. shading [32], direct vs. global [25, 49], and more [36]. But we operate in thermal domain where both heat and light transports are involved. Tanaka et al. [38] decompose the time varying heat transport emissions into ad-hoc diffuse and global components with a constant background reflection that can be subtracted out while our physically-grounded approach estimates the emissivity and temperatures of the object and background.

3. Thermal Image Formation Model

Thermal cameras convert the incoming radiation in the Long Wave Infrared (LWIR) spectrum to digital images. The radiation emitted by a blackbody at temperature T and wavelength λ follows Planck's spectral distribution [8]:

$$L_b(\lambda, T_{bb}) = \frac{c_1}{\lambda^5 \exp\left(\frac{c_2}{\lambda T_{bb}}\right) - 1}, \tag{1}$$

where, c_1, c_2 are known radiation constants. However, real-world objects emit only a fraction of this radiation, determined by their spectral emissivity $\epsilon(\lambda)$, where $0 \leq \epsilon(\lambda) \leq 1$. Thus, the emitted radiation from a real object is:

$$L(\lambda, T) = \epsilon(\lambda) L_b(\lambda, T) \tag{2}$$

A thermal camera uses this emitted radiation to infer the temperature of an object, therefore it is crucial to account for the object's emissivity to obtain accurate measurements. But, the image formation model of a thermal camera is more involved than directly measuring the emitted radiation.

Figure 1 shows various components of thermal radiation that contribute to the image formation. The thermal radiation at wavelength λ comprises of five main components: radiation $\Phi_s(\lambda)$ emitted by the target surface, radiation $\Phi_b(\lambda)$ emitted by the background objects and reflected or transmitted through the target, radiation $\Phi_t(\lambda)$ emitted by the transmission path, radiation $\Phi_o(\lambda)$ emitted by the optics and radiation $\Phi_i(\lambda)$ emitted by the internal components of the thermal camera. The total radiation received at the camera is a weighted sum of these components, where the weights depend on the emissivity $\epsilon(\lambda)$, transmissivity $\tau(\lambda)$ and reflectivity $r(\lambda)$. From Kirchoff's law:

$$\epsilon(\lambda) + \tau(\lambda) + r(\lambda) = 1 \tag{3}$$

Then, we can write the total spectral radiation leaving the object's surface as a weighted combination of an object's emitted, reflected or transmitted radiation:

$$\Phi_A(\lambda) = \epsilon_s(\lambda)\Phi_s(\lambda) + (1 - \epsilon_s(\lambda))\Phi_b(\lambda) \tag{4}$$

Similarly, the radiation reaching the optics is a combination of attenuated radiation from the object's surface and radiation emitted by the transmission path,

$$\Phi_B(\lambda) = \tau_t(\lambda)\big(\Phi_A(\lambda)\big) + (1 - \tau_t(\lambda))\Phi_t(\lambda) \tag{5}$$

A portion of the camera's own thermal radiation is also reflected by the optical elements, an effect known as the narcissus effect [22]. Accounting for this, the total radiation reaching the detector at wavelength λ is:

$$\Phi_C(\lambda) = \tau_o(\lambda)\Phi_B(\lambda) + \epsilon_o(\lambda)\Phi_o(\lambda) + r_o(\lambda)\Phi_i(\lambda) \tag{6}$$

here, ϵ_o, τ_o and r_o denote the emissivity, transmissivity and reflectivity of the optical components respectively. The total incoming radiation Φ_{tot} at the detector is obtained by integrating across the camera's sensitivity spectrum

$$\Phi_{tot} = \int_{\lambda_{\min}}^{\lambda_{\max}} \Phi_C(\lambda)\mathrm{d}\lambda \tag{7}$$

We make the following assumptions based on our imaging setup: (i) the camera-to-scene distance is small enough to neglect atmospheric attenuation ($\tau_t \approx 1$), (ii) radiation from the optics and internal camera elements remains constant across experiments, and (iii) we consider averaged values of ϵ, τ, and r within a given sensitivity spectrum. Under these conditions, the total incoming radiation is:

$$\Phi_{tot} = g\Phi_A + f \tag{8}$$

where $g = \tau_o$ and $f = \epsilon_o\Phi_o + r_o\Phi_i$ are the gain and offset corrections for the narcissus effect introduced by optical filters and internal camera elements. Unlike [22], which corrects only for the offset by subtracting a reference frame, we apply both gain and offset corrections for better accuracy.

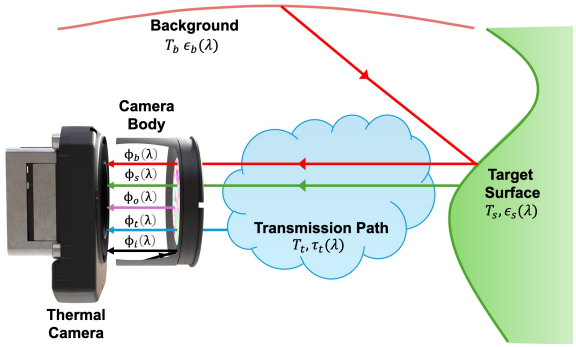

Figure 1. Image formation in a thermal camera comprises of radiation from the object Φ_s, reflection Φ_b, optics Φ_o, transmission path Φ_t and its internal components Φ_i.

From Planck's law (Eq. 1), any radiation can be interpreted as blackbody radiation at some temperature T. After correcting for gain and offset, we express the thermal camera signal $U(T_{bb})$ for incoming radiation Φ_A (Eq. 4) as:

$$U(T_{bb}) = \epsilon_s U(T_o) + (1 - \epsilon_s)U(T_b) \tag{9}$$

where, T_o is the object temperature, T_b denotes an effective background temperature and $U(T_{bb})$ relates pixel intensity to blackbody temperatures and is approximated using the Sakuma-Hattori model:

$$U(T_{bb}) = \frac{c_1}{\exp\left(\frac{c_2}{c_3 T_{bb} + c_4}\right) - 1}, \tag{10}$$

with calibration constants c_1, c_2, c_3, and c_4 determined through curve fitting [29]. Given that the Sakuma-Hattori function is nearly linear for small temperature differences around ambient conditions, *i.e.* $U(T) = aT + b$. The calibration constants a and b are obtained using a blackbody at a known temperature (see supplementary).

Band Limited Thermal Videos: We capture M multispectral thermal bands that lie within the camera's sensitivity spectrum of 8-14 µm. The captured pixel intensity within each band m across an entire video is expressed as:

$$I_m(t) = \epsilon_m U_m(T_o(t)) + (1 - \epsilon_m)U_m(T_b(t)). \tag{11}$$

As the background is composed of various objects with different emissivities and temperatures, we assume their effective contributions average out such that T_b is independent of wavelength. For simplicity, we denote the measured pixel intensity $U(T_{bb})$ in Eq. 9 as I and band emissivity as ϵ_m. Function $U_m(T)$ that converts blackbody temperature to pixel intensity varies across bands $U_m(T) = a_m T + b_m$ due to varying optical properties and camera response function within the sensitivity spectrum. Further, t denotes the frame of the video indicating a time varying object temperature $T_o(t)$ and effective background temperature $T_b(t)$.

4. Method

In this section, we analyze thermal videos captured in two different spectral bands, where emissivity varies between them. Given thermal video data over N timesteps, represented as $\{I_m^1, ..., I_m^N\}$ for each band $m \in \{1, 2\}$, our goal is to decompose the thermal radiation into: the band emissivity ϵ_m, the time-varying object's temperature $T_o(t)$, and the effective background temperature $T_b(t)$ that governs the image formation process described in Eq. 11.

4.1. Calibrated Case

A priori, we calibrate the spectral emissivity of the materials of interest. Then, we use the calculated emissivity to separate the dynamic reflected background and the emission in a new scene with the same materials. To calibrate, we use a thermocouple and measure temperature at a point on the material. Also, we place a blackbody at known temperature near the material to serve as background. We then capture thermal images of the material with different spectral filters (see Fig. 4). Here, the object and background temperatures are known using the thermocouple and the blackbody. Then, emissivity ϵ_m is computed as:

$$\epsilon_m = \frac{I_m - U_m(T_b)}{U_m(T_o) - U_m(T_b)} \tag{12}$$

Estimating T_o and T_b from Known Emissivities: The emissivities ϵ_m for two spectral bands ($m \in \{1, 2\}$) obtained from calibration are then used for a new scene with unknown object and background temperature. The object's time-varying temperature T_o and the background temperature T_b are estimated using closed-form expressions. Expanding Eq. 11 for two bands ($m \in \{1, 2\}$), we have:

$$\begin{aligned} I_1 &= \epsilon_1 a_1 (T_o - T_b) + a_1 T_b + b_1 \\ I_2 &= \epsilon_2 a_2 (T_o - T_b) + a_2 T_b + b_2 \,. \end{aligned} \tag{13}$$

Solving for T_o and T_b form the above equation, we get:

$$\begin{aligned} T_o &= \frac{a_1(I_2 - b_2)(\epsilon_1 - 1) - a_2(I_1 - b_1)(\epsilon_2 - 1)}{a_1 a_2 (\epsilon_1 - \epsilon_2)} \\ T_b &= \frac{a_1(I_2 - b_2)\epsilon_1 - a_2(I_1 - b_1)\epsilon_2}{a_1 a_2 (\epsilon_1 - \epsilon_2)} \,. \end{aligned} \tag{14}$$

4.2. Uncalibrated Case

When object emissivities are unknown, jointly estimating object and background temperatures becomes ill-posed. Although the thermal light transport in Eq. 9 is in principle linearly independent across bands with differing emissivities, the near-linear behavior of the Sakuma–Hattori function $U(T)$ close to ambient conditions collapses the effective rank of the system. As a result, the two-band image formation model can be rewritten compactly as:

$$\begin{bmatrix} \dfrac{I_1 - b_1}{a_1} \\ \dfrac{I_2 - b_2}{a_2} \end{bmatrix} = \begin{bmatrix} \epsilon_1 & 1 - \epsilon_1 \\ \epsilon_2 & 1 - \epsilon_2 \end{bmatrix} \begin{bmatrix} T_\mathrm{o} \\ T_\mathrm{b} \end{bmatrix}, \tag{15}$$

where the constants a_i and b_i are absorbed into the left-hand side. Writing this compactly as $\mathbf{I} = \mathbf{ET}$ makes the structure apparent: the matrix $\mathbf{E}$ has rank 2. This follows from our reasonable assumption that T_b is independent of wavelength. Extending this formulation to M spectral bands and N timesteps yields $\mathbf{I} \in \mathbb{R}^{M \times N}$, $\mathbf{E} \in \mathbb{R}^{M \times 2}$, and $\mathbf{T} \in \mathbb{R}^{2 \times N}$, with $\mathbf{E}$ still of rank 2 by construction. Hence, additional constraints are required to recover T_o and T_b in the uncalibrated case.

4.2.1 Thermography in Dynamic Scenes

The appearance of an object is often more dynamic in the thermal spectrum than in the visible spectrum. Even when the object remains stationary and light transport is constant at millisecond timescales, it may not be in thermal equilibrium with its surroundings, leading to continuous variations in thermal emission over time. Fortunately, heat transport evolves gradually, while background reflections typically change abruptly and lack correlation with the object's emission. Leveraging this observation, we introduce novel constraints that disentangle reflection and emission.

Prior on object temperature T_o: The change in object temperature is governed by the heat transfer process that occurs within and around it governed by the standard heat transport equation [6, 23, 41]. Ramanagopal et al. [28] show that with short capture timescales, conduction is negligible, simplifying the system to an ordinary differential equation (ODE). The solution to this ODE follows an exponential form serving as a valuable prior for addressing the ill-posed nature of the system in Eq. 15.

Constraints from a static background: In many practical scenarios, the change in background radiation observed by the camera occurs in smaller portions of the scene. This is especially true when the target surface is glossy/specular. In such cases, one can write a differential form of Eq. 11,

$$\frac{1}{a_m}\frac{\partial I_m}{\partial t} = \epsilon_m \frac{\partial T_o}{\partial t} + (1 - \epsilon_m)\frac{\partial T_b}{\partial t} \tag{16}$$

At pixels where the change in background radiation is insignificant, the differential change is negligible. Taking the ratio of the dual bands in Eq. 16 yields a constant ratio k_1:

$$k_1 = \frac{\epsilon_2}{\epsilon_1} = \frac{a_1 \dfrac{\partial I_2}{\partial t}}{a_2 \dfrac{\partial I_1}{\partial t}} \tag{17}$$

Constraints from a dynamic background: From Eq 11 we can decompose the image for any spectral band as,

$$I_m(t) = \epsilon_m U(T_o(t)) + (1-\epsilon_m)U(\bar{T}_b + T_b^*(t)) \quad (18)$$

where $\bar{T}_b$ is some mean background temperature and $T_b^*(t)$ represents the differential change in the background from the mean. This can be separated into:

$$I_m(t) = \underbrace{\epsilon_m U(T_o(t)) + (1-\epsilon_m)U(\bar{T}_b)}_{\tilde{I}_m(t)} + (1-\epsilon_m)a_m T_b^*(t) \quad (19)$$

where $\tilde{I}_m(t)$ represents a smoothly varying, emission-dominant signal derived from the object's temperature prior. Here we assume that object and background temperatures vary in an un-correlated fashion. From this, given smooth signals in two spectral bands, we can express:

$$I_1(t) - \tilde{I}_1(t) = (1-\epsilon_1)a_1 T_b^*(t) \quad (20)$$

$$I_2(t) - \tilde{I}_2(t) = (1-\epsilon_2)a_2 T_b^*(t) \quad (21)$$

Ratio between Eq. 20 and 21 yields a constant ratio k_2 across the thermal video,

$$k_2 = \frac{1-\epsilon_2}{1-\epsilon_1} = \frac{a_1}{a_2}\frac{I_2(t) - \tilde{I}_2(t)}{I_1(t) - \tilde{I}_1(t)} \quad (22)$$

Solution in dynamic scenes: The constraints described in Eq. 17 and 22 show that having access to robust ratios k_1 and k_2 based on locations that undergo changes in background radiation (either static or dynamic) along with a smooth signal $\tilde{I}_m(t)$, yields spectral band emissivity:

$$\epsilon_1 = \frac{k_2 - 1}{k_2 - k_1}, \quad \epsilon_2 = k_1\frac{k_2 - 1}{k_2 - k_1} \quad (23)$$

4.2.2 Optimization Framework

Algorithm 1 summarizes our optimization procedure. Our goal is to jointly estimate emissivities ϵ_m and smooth signals $\tilde{I}_m(t)$ using the constraints in Eq. 17 and Eq. 22. The recovered emissivities are then used in Eq. 14 to compute object and background temperatures $T_o(t)$ and $T_b(t)$.

Noise Modeling: Thermal videos captured with spectral filters suffer from low SNR. We explicitly model per-pixel, per-timestep noise $I_m^\varepsilon(t)$ and regularize it to follow a zero-mean distribution:

$$\mathcal{L}_{\text{noise}}^{L2} = \|I_m^\varepsilon(t)\|_2^2, \quad \mathcal{L}_{\text{noise}}^{M} = \frac{1}{T}\sum_{i=1}^{T} I_m^\varepsilon(i). \quad (24)$$

For the rest of our optimization we use a noise-corrected signal $\ddot{I}_m(t)$, where, $\ddot{I}_m(t) = I_m(t) - I_m^\varepsilon(t)$.

Algorithm 1 Dual-band separation (uncalibrated)

procedure DBVT($\mathbf{I_1}, \mathbf{I_2}, a_1, a_2, b_1, b_2$)
  // Initialize Variables
  $\tilde{I}_1(t), \tilde{I}_2(0), I_m^\epsilon(t) \leftarrow 0, \epsilon_m \leftarrow rand(0,1)$
  while $\|\Delta\mathcal{L}\| > \delta_L$ **do**
    $\ddot{I}_m(t) \leftarrow I_m(t) - I_m^\epsilon(t)$ ▷ noise-corrected
    Estimate $\tilde{I}_2(t)$ using Eq. 25 ▷ Smooth Signal
    Estimate $\hat{I}_2(t)$ using Eq. 28 ▷ complete Signal
    Compute: $\mathcal{L}_{\text{smooth}}, \mathcal{L}_{\text{Huber}}, \mathcal{L}_{\text{MSE}}, \mathcal{L}_{\text{noise}}$
  end while
  Recover $T_o, T_b \leftarrow$ Eq. 14
  return ϵ_m, T_o, T_b
end procedure

Smooth Signal Estimation: We estimate smooth, emission-dominant signals $\tilde{I}_m(t)$ by leveraging the ratio constraint k_1. Specifically, $\tilde{I}_2(t)$ is recursively constructed from $\tilde{I}_1(t)$ as:

$$\tilde{I}_2(t) = \tilde{I}_2(t-1) + k_1\frac{a_2}{a_1}(\tilde{I}_1(t) - \tilde{I}_1(t-1)). \quad (25)$$

We optimize for $\tilde{I}_1(t)$ and an unknown offset $\tilde{I}_2(0)$ such that $\tilde{I}_m(t)$ remains smooth while staying close to $\ddot{I}_m(t)$. Smoothness is enforced using a second-order prior:

$$\mathcal{L}_{\text{smooth}} = \|\tilde{I}_m(t-1) - 2\tilde{I}_m(t) + \tilde{I}_m(t+1)\|^2. \quad (26)$$

To ensure that estimated smooth signal remains consistent with the measured quantities while being robust to abrupt changes in the background, we introduce a normalized Huber loss with a threshold δ as follows:

$$\mathcal{L}_{\text{Huber}} = \begin{cases} \frac{1}{2}\|\ddot{I}_m^n - \tilde{I}_m^n\|^2 & |\cdot| \le \delta \\ \delta(|\ddot{I}_m^n - \tilde{I}_m^n| - \frac{1}{2}\delta) & \text{otherwise.} \end{cases} \quad (27)$$

Here, superscript n represents the normalization of the signals based on $\ddot{I}_m$ by subtracting from its mean and dividing by its standard deviation.

Reconstruction via k_2: We further regularize the estimates by reconstructing the measured signal from the smooth signal using the constraint k_2. The reconstructed signal $\hat{I}_m(t)$ is given by:

$$\hat{I}_2(t) = \tilde{I}_1(t) + k_2\frac{a_2}{a_1}(\ddot{I}_1(t) - \tilde{I}_1(t)). \quad (28)$$

We minimize the discrepancy between reconstructed and observed signals using a mean squared error objective:

$$\mathcal{L}_{\text{MSE}} = \|\ddot{I}_m^n - \hat{I}_m^n\|^2. \quad (29)$$

Emissivity Optimization: Although emissivities can be analytically recovered from k_1 and k_2 (Eq. 23), this is unstable under noise and may yield invalid solutions. Instead, we directly optimize ϵ_m values as proxies for k_1 and k_2.

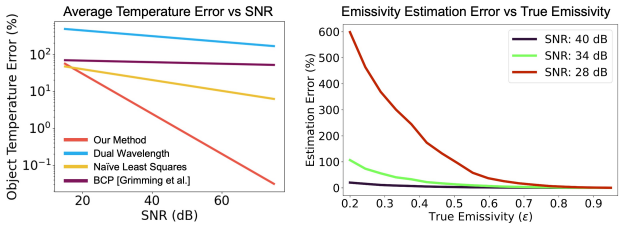

Figure 2. **[Left]** Comparison of our method with a recent BCP [9] technique, naive multi-wavelength, and dual-wavelength pyrometry [2, 39] on simulated thermal videos of different materials sourced from the spectral library [3, 21]. The naive least squares result uses the best of five initializations selected based on the one that achieved the least objective. All methods degrade at high noise; ours improves significantly at moderate noise (log scale). See our plot in the supplementary for per-material comparisons. **[Right]** Shows the error in estimated emissivities from our method for a range of emissivities under varying noise levels in simulation.

Overall Objective: We jointly optimize $\tilde{I}_1(t)$, $\tilde{I}_2(0)$, ϵ_m, and $I_m^{\varepsilon}(t)$ (for $m \in \{1, 2\}$) by minimizing:

$$\underset{\tilde{I}_1(t),\ \tilde{I}_2(0),\ \epsilon_m,\ I_m^{\varepsilon}(t)}{\arg\min} \begin{aligned} &\gamma_1 \mathcal{L}_{\text{smooth}} + \gamma_2 \mathcal{L}_{\text{Huber}} + \gamma_3 \mathcal{L}_{\text{MSE}} \\ &+ \gamma_4 \mathcal{L}_{\text{noise}}^{L2} + \gamma_5 \mathcal{L}_{\text{noise}}^{M}. \end{aligned} \tag{30}$$

Here, γ_i denote the corresponding loss weights. This formulation balances temporal smoothness, measurement fidelity, and robustness to noise, enabling reliable recovery of emissivity and temperature in dynamic scenes.

5. Experimental Results

We demonstrate our calibrated and uncalibrated approaches using both simulations and real data with time-varying emissions and reflections and compare to three baselines.

Baselines: We compare our method against representative approaches for reflection reduction and emission estimation from both pyrometry and computer vision literature [2, 9, 39], along with a naive least-squares optimization baseline. The recent BCP method [9] assumes that the brightest pixels within a local patch approximate near-blackbody emitters, using this prior to suppress reflections. We also include a traditional dual-wavelength pyrometry baseline [2, 39], which neglects background radiation and models object emissivity as a gray body at the selected wavelengths. Lastly, our naive baseline jointly estimates emissivity, object temperature, and background temperature by minimizing an image reconstruction loss from Eq. 11 in the least-squares sense. To ensure stability, we perform optimization from five random initializations and retain the result with the lowest objective value for this method.

Experimental Setup: Our setup uses two FLIR Boson thermal cameras-a performance-grade unit ($\leq$ 40 mK NETD) and an industrial-grade unit ($\leq$ 20 mK NETD)-operating at 640×512 resolution with a 24° HFOV lens. The cameras are placed side by side to minimize parallax; we opt against using a beam splitter due to its significant SNR penalty. We capture data using spectral filters centered at 8.5 μm, 9.5 μm, 10.6 μm, and 12.1 μm with FWHM values of 0.5 μm, 0.5 μm, 1.5 μm, and 0.5 μm, respectively. The filters are mounted on an FW103H/M motorized wheel, with the industrial-grade camera positioned in front of the wheel. Ground-truth surface temperatures are recorded using a TC-08 data logger with Type-T thermocouples.

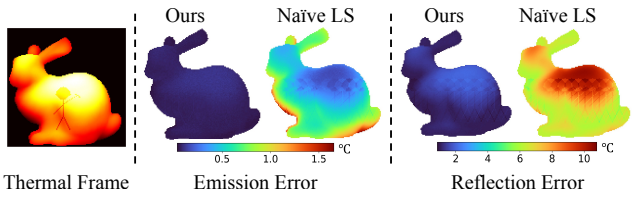

Figure 3. Spatial temperature errors (in $^{\circ}C$) averaged over video frames for our method and naive least squares (LS). The thermal video (left) was obtained through a combination of heat simulation and a moving toy background based on Eq. 11 with added noise.

Loss-terms	Reconstruction term	Smoothing term	Huber loss	Noise terms
% reduction	90.10%	56.73%	11.31%	11.64%

Table 1. Percentage error reduction in temperature estimates by adding the loss terms individually (while including the other three loss terms), averaged over a sweep of SNR levels in simulation.

5.1. Simulations

We evaluate our method using controlled heat-transfer simulations [23] on objects (e.g., Stanford Bunny) against a dynamically varying background. Thermal sequences are generated via Eq. 11 over a wide range of emissivities sampled from spectral libraries [3, 21] (see supplementary for details). As shown in Fig. 2 (left), our method consistently outperforms all baselines in recovering object temperature across noise levels. Spatial error maps in Fig. 3 show accurate reconstruction, with minor errors near object boundaries due to reduced thermal contrast from reflected background. Fig. 2 (right) demonstrates that our uncalibrated approach can also recover emissivity across noise levels, performing well at moderate-to-high emissivities, while low-emissivity and high-noise regimes remain challenging.

Ablation of Loss Terms: Table 1 quantifies the contribution of each loss term by measuring performance gains when added to the remaining losses. Reconstruction loss is the most critical component, followed by smoothing and Huber losses, while the noise term becomes beneficial under high-noise conditions.

Filter Selection: Spectral filters are selected based on two factors: SNR of the thermal image captured and the change in emissivity across the filters. For the first, we choose the image without a filter (8-14μm) since the SNR is highest.

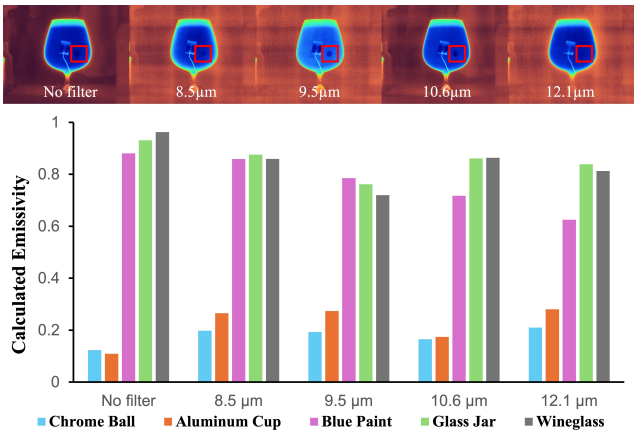

Figure 4. **[Top]** Images captured with different spectral filters used for calibration. The red bounding box highlights the reflected blackbody on the object. A thermocouple is placed next to this reflection to measure the object's temperature. **[Bottom]** Calibrated emissivity values for various materials using our technique align the closely with reported values [1, 3] as shown in Tab. 2.

Methods	Chrome Ball	Al. Cup	Blue Paint	Glass Jar	Wineglass
Reference Emissivity	0.1	0.05	0.87	0.95	0.95
Reflector Method [10]	0.43	0.16	0.88	0.93	0.97
Ours	**0.12**	**0.10**	**0.88**	**0.93**	**0.96**

Table 2. Comparison of emissivity estimates obtained using our calibration technique and the reflector method [10] for various materials using 8–14μm range. Our estimates closely match those from the reflector method and show improved accuracy for low-emissivity materials. Reference emissivity values are from [1, 3] and may not reflect the exact emissivities of the tested objects.

Then, we choose the second filter, $9.5\mu m$, that has the highest condition number of the resulting emissivity matrix $\mathbf{E}$ from Eq. 15 using emissivities sourced from the spectral library [21]. *See full ablation in Supplementary.*

5.2. Real Experiments

Emissivity Calibration and Comparison: Figure 4 illustrates our emissivity calibration setup. A thermocouple is placed near a specular background reflection to measure the object's temperature. To obtain a robust estimate of the object's emissivity using Eq. 12, we capture images of the object and its background reflections at different temperatures. This is achieved by heating the object with hot water or a hot air gun and varying the reflected blackbody temperature. The resulting spectral-band emissivities for multiple objects are shown in Figure 4. As summarized in Table 2, our estimates closely match the reference values in [1, 3]. Our method also outperforms the FLIR Systems reflector method [10] for low-emissivity surfaces, because our method leverages specular reflections and is less sensitive to inaccuracies compared to the reflector approach.

Reflection-Emission Separation Results: First, we illustrate the difference between reflection due to light transport and emission due to heat transport with the same object: the reflection of fingers moving in front of a glass plate vs. the finger prints dissipating on the glass plate. Our method separates these components well, as shown in Figure 5. We demonstrate separation for several interesting and challenging scenes whose descriptions are given in Figure 6. Our results demonstrate the effectiveness of our optimization in isolating background reflections from object emissions, even in visually imperceptible cases such as the glass plate, while simultaneously providing meaningful temperature estimates validated via contact thermometry. While our calibrated approach tends to be noisier in low-signal regions due to its analytical nature and the absence of a temperature prior, it achieves decent separation results for high-emissivity objects with good signal, as shown in Figure 6. The estimated emissivities from our uncalibrated approach deviate by an average of $0.17, 0.15, 0.19$, and 0.15 from our calibrated method across the dual spectral bands. Also, our uncalibrated approach has an average temperature difference of 1.72% and 5.34% with wineglass and coffee pot videos compared to contact-measurements with a thermocouple, highlighting the ability of our method to perform thermography in complex, dynamic real-world scenarios for the first time. Best viewed in the Supplementary video.

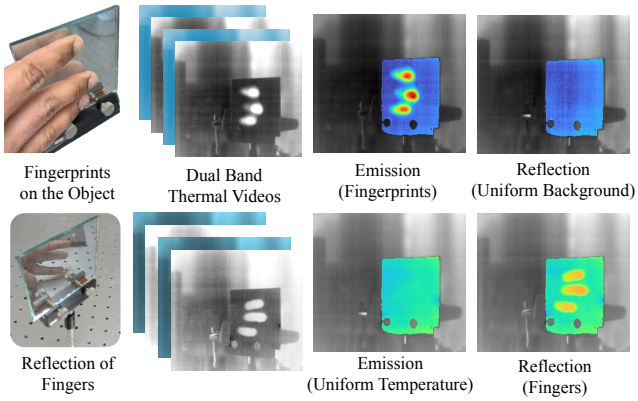

Figure 5. Finger print versus Finger reflection: [Top] Separating finger prints on a glass plate that emit light (heat transport) from the reflection of the uniform background (light transport). [Bottom] Separating the reflection of the fingers (light transport) by the glass plate at constant room temperature (heat transport).

6. Limitations

Thermographic separation of reflection and emission is a highly ill-posed problem, with no widely accepted standard for analysis in general scenarios. While our work makes significant progress in resolving ambiguities between emitted and reflected radiation, several challenges remain: (i) We assume that changes in background radiation are uncorrelated with the object signal. While this holds for reflections caused by object motion, it may not be valid in scenar-

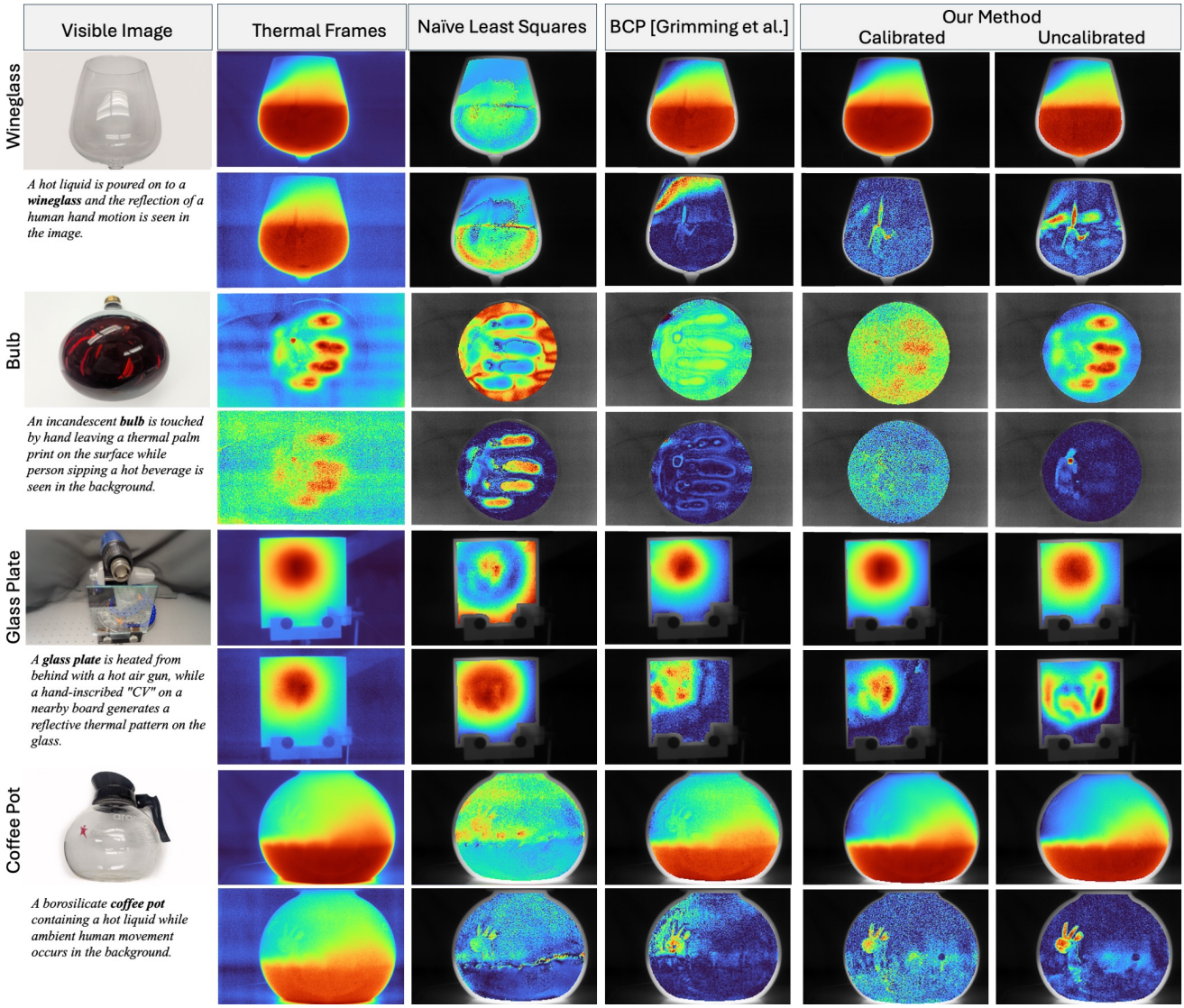

Figure 6. Reflection–emission separation results for objects heated using different methods (hot liquid, hand contact, and hot air). For each object, the first row shows the estimated emission, and the second row shows the mean-subtracted reflection. Input thermal frames from two spectral bands are shown on the left. Since per-pixel temperature ground truth is unobtainable, we record temperatures at sparse locations using a thermocouple. The wineglass and coffee pot reached peak temperatures of $63.6^{\circ}C$ and $63.1^{\circ}C$, respectively. The corresponding estimation errors were 1.72%, 5.04%, 14.6%, and 31.68% for the uncalibrated, calibrated, BCP, and naive approaches on the wineglass video, and 5.34%, 0.36%, 6.62%, and 45.5% for the coffee pot video. Our method cleanly separates emission and reflection even in challenging cases, such as isolating the "CV" reflection on glass or distinguishing a palm print on a bulb from a person sipping coffee.

ios such as a room warming uniformly. (ii) Low-cost microbolometers with spectral filters have limited sensitivity, making it difficult to detect small temperature differences in low-emissivity objects due to poor signal-to-noise ratio.

7. Conclusion

Heat and light transport jointly encode the material and geometric properties of a scene [6, 23, 38]. Disentangling these components is essential for robust thermal interpretation in unconstrained environments. We presented a dual-band approach with both calibrated and uncalibrated formulations, supported by a calibration procedure that estimates spectral-band emissivities from measured temperatures. We further introduced a theory and optimization strategy that leverages variations in emissivity and temperature to add structure to an otherwise ill-posed decomposition. As thermal cameras become increasingly accessible, exploring the interaction between light and heat transport remains a promising direction for advancing thermal imaging and its applications.

Acknowledgements: This work was partly supported by NSF grants IIS-2107236, and NSF-NIFA AI Institute for Resilient Agriculture.

References

[1] Design 1st. Thermal emissivity values. Online PDF, n.d. Available: https://www.design1st.com/Design-Resource-Library/engineering_data/ThermalEmissivityValues.pdf. Accessed: 2025-11-09. 7

[2] António Araújo. Multi-spectral pyrometry—a review. *Measurement Science and Technology*, 28(8):082002, 2017. 1, 2, 6, 12

[3] A.M. Baldridge, S.J. Hook, C.I. Grove, and G. Rivera. The aster spectral library version 2.0. *Remote Sensing of Environment*, 113(4):711–715, 2009. 1, 6, 7, 11, 12

[4] Fanglin Bao, Xueji Wang, Shree Hari Sureshbabu, Gautam Sreekumar, Liping Yang, Vaneet Aggarwal, Vishnu N Boddeti, and Zubin Jacob. Heat-assisted detection and ranging. *Nature*, 619(7971):743–748, 2023. 2

[5] Krzysztof Chrzanowski. Non-contact thermometry measurement errors. 2001. 1, 2

[6] Aniket Dashpute, Vishwanath Saragadam, Emma Alexander, Florian Willomitzer, Aggelos Katsaggelos, Ashok Veeraraghavan, and Oliver Cossairt. Thermal spread functions (tsf): Physics-guided material classification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1641–1650, 2023. 2, 4, 8

[7] Akshat Dave, Yannick Hold-Geoffroy, Miloš Hašan, Kalyan Sunkavalli, and Ashok Veeraraghavan. Snapshot polarimetric diffuse-specular separation. *Opt. Express*, 30(19):34239–34255, 2022. 2

[8] D. P. DeWitt and F. P. Incropera. *Physics of Thermal Radiation*, chapter 1, pages 19–89. John Wiley & Sons, Ltd, 1988. 2

[9] Robert Grimming, Orges Furxhi, Ronald Driggers, and Kyle Renshaw. Multiband longwave infrared reflectance removal using blackbody channel prior. *Optical Engineering*, 62(3): 033101, 2023. 2, 6, 12

[10] FLIR Systems Inc. Thermography reference documentation. Online, 2019. Available: https://support.flir.com/DSDownload/Assets/T810442-en-US_A4.pdf. Accessed: 2025-11-09. 2, 7

[11] Masaki Kaga, Takahiro Kushida, Tsuyoshi Takatani, Kenichiro Tanaka, Takuya Funatomi, and Yasuhiro Mukaigawa. Thermal non-line-of-sight imaging from specular and diffuse reflections. *IPSJ Transactions on Computer Vision and Applications*, 11(1):8, 2019. 2

[12] Soma Kajiyama, Taihe Piao, Ryo Kawahara, and Takahiro Okabe. Separating partially-polarized diffuse and specular reflection components under unpolarized light sources. In *2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 2548–2557, 2023. 2

[13] Kazuma Kitazawa and Tsuyoshi Takatani. Shape from polarization of thermal emission and reflection. In *2025 IEEE International Conference on Computational Photography (ICCP)*, 2025. 2

[14] Kazuma Kitazawa and Tsuyoshi Takatani. Shape from Polarization of Thermal Emission and Reflection . In *2025 IEEE International Conference on Computational Photography (ICCP)*, pages 1–11, Los Alamitos, CA, USA, 2025. IEEE Computer Society. 2

[15] Jean-Claude Krapez. Multiwavelength thermometry without a priori emissivity information: from promise to disillusionment. In *Thermosense: Thermal Infrared Applications XLVII*, page 1347006. International Society for Optics and Photonics, SPIE, 2025. 1, 2

[16] Takahiro Kushida and Kenichiro Tanaka. Thermal polarimetric multi-view stereo. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 27390–27399, 2025. 2

[17] Thomas Lafargue-Tallet, Romain Vaucelle, Cyril Caliot, Abderezak Aouali, Emmanuelle Abisset-Chavanne, Alain Sommier, Raymond Peiffer, and Christophe Pradere. Active thermo-reflectometry for absolute temperature measurement by infrared thermography on specular materials. *Scientific Reports*, 12(1):7814, 2022. 1, 2

[18] Anat Levin and Yair Weiss. User assisted separation of reflections from a single image using a sparsity prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(9):1647–1654, 2007. 2

[19] Stephen Lin, Yuanzhen Li, Sing Bing Kang, Xin Tong, and Heung-Yeung Shum. Diffuse-specular separation and depth recovery from image sequences. In *Computer Vision — ECCV 2002*, pages 210–224, Berlin, Heidelberg, 2002. Springer Berlin Heidelberg. 2

[20] Tomohiro Maeda, Yiqin Wang, Ramesh Raskar, and Achuta Kadambi. Thermal non-line-of-sight imaging. In *2019 IEEE International Conference on Computational Photography (ICCP)*, pages 1–11, 2019. 1, 2

[21] Susan K. Meerdink, Simon J. Hook, Dar A. Roberts, and Elsa A. Abbott. The ecostress spectral library version 1.0. *Remote Sensing of Environment*, 230:111196, 2019. 1, 6, 7, 11, 12

[22] Yasuto Nagase, Takahiro Kushida, Kenichiro Tanaka, Takuya Funatomi, and Yasuhiro Mukaigawa. Shape from thermal radiation: Passive ranging using multi-spectral lwir measurements. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 12651–12661, 2022. 2, 3, 11

[23] Sriram Narayanan, Mani Ramanagopal, Mark Sheinin, Aswin C. Sankaranarayanan, and Srinivasa G. Narasimhan. Shape from heat conduction. In *Computer Vision – ECCV 2024: 18th European Conference, Milan, Italy, September 29–October 4, 2024, Proceedings, Part XXXVIII*, page 426–444, Berlin, Heidelberg, 2024. Springer-Verlag. 1, 2, 4, 6, 8

[24] Mariam Nawaz, Gulistan Raja, and Muhammad Qasim. A new approach for dehazing and enhancement of infrared images. In *2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST)*, pages 545–551, 2021. 2

[25] Shree K. Nayar, Gurunandan Krishnan, Michael D. Grossberg, and Ramesh Raskar. Fast separation of direct and

global components of a scene using high frequency illumination. *ACM Trans. Graph.*, 25(3):935–944, 2006. 2

[26] Fred E. Nicodemus. Directional reflectance and emissivity of an opaque surface. *Appl. Opt.*, 4(7):767–775, 1965. 2

[27] N. Ogasawara, C. Kobayashi, and H. Yamada. Infrared image correction for the reduction of background reflection. In *Thermosense: Thermal Infrared Applications XXXIX*, page 102141E. International Society for Optics and Photonics, SPIE, 2017. 1, 2

[28] Mani Ramanagopal, Sriram Narayanan, Aswin C. Sankaranarayanan, and Srinivasa G. Narasimhan. 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)*, pages 11924–11933, 2024. 2, 4

[29] F Sakuma. Establishing a practical temperature standard by using a narrow-bannd radiation thermometer with a silicon detector. *Temperature, Its Measurement and Control in Science and Industry*, 5:421–427, 1982. 3, 11

[30] Peter Saunders and D Rod White. Physical basis of interpolation equations for radiation thermometry. *Metrologia*, 40 (4):195, 2003. 2

[31] Thomas Schmugge, Andrew French, Jerry C Ritchie, Albert Rango, and Henk Pelgrum. Temperature and emissivity separation from multispectral thermal infrared observations. *Remote Sensing of Environment*, 79(2):189–198, 2002. Recent Advances in Remote Sensing of Biophysical Variables. 2

[32] Steven A. Shafer. Using color to separate reflection components. *Color Research & Application*, 10(4):210–218, 1985. 2

[33] Mark Sheinin, Aswin C. Sankaranarayanan, and Srinivasa G. Narasimhan. Projecting trackable thermal patterns for dynamic computer vision. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 25223–25232, 2024. 2

[34] YiChang Shih, Dilip Krishnan, Fredo Durand, and William T. Freeman. Reflection removal using ghosting cues. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015. 2

[35] Antonio C.S. Souza, Márcio C.F. Macedo, Verônica P. Nascimento, and Bruno S. Oliveira. Real-time high-quality specular highlight removal using efficient pixel clustering. In *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 56–63, 2018. 2

[36] Rui Sun, Tanbin Liao, Zhiguo Fan, Xudong Zhang, and Changxiang Wang. Polarization dehazing method based on separating and iterative optimizing airlight from the frequency domain for different concentrations of haze. *Appl. Opt.*, 61(35):10362–10373, 2022. 2

[37] Soshi Suzuki and Nagahisa Ogasawara. Infrared thermographic test for removing background reflection based on polarization theory. *NDT & E International*, 103:19–25, 2019. 1, 2

[38] Kenichiro Tanaka, Nobuhiro Ikeya, Tsuyoshi Takatani, Hiroyuki Kubo, Takuya Funatomi, Vijay Ravi, Achuta Kadambi, and Yasuhiro Mukaigawa. Time-resolved far infrared light transport decomposition for thermal photometric stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(6):2075–2085, 2021. 1, 2, 8

[39] B. K. Tsai, R. L. Shoemaker, D. P. DeWitt, B. A. Cowans, Z. Dardas, W. N. Delgass, and G. J. Dail. Dual-wavelength radiation thermometry: Emissivity compensation algorithms. *International Journal of Thermophysics*, 11(1):269–281, 1990. 2, 6, 12

[40] Simon Verspeek, Bart Ribbens, Xavier Maldague, and Gunther Steenackers. Spot weld inspections using active thermography. *Applied Sciences*, 12(11), 2022. 1, 2

[41] Michael Vollmer. Infrared thermal imaging. In *Computer Vision: A Reference Guide*, pages 1–4. Springer, 2020. 1, 2, 4

[42] Patrick Wieschollek, Orazio Gallo, Jinwei Gu, and Jan Kautz. Separating reflection and transmission images in the wild. In *Computer Vision – ECCV 2018*, pages 90–105, Cham, 2018. Springer International Publishing. 2

[43] Tianfan Xue, Michael Rubinstein, Ce Liu, and William T. Freeman. A computational approach for obstruction-free photography. *ACM Trans. Graph.*, 34(4), 2015. 2

[44] Zeqing Leo Yuan, Mani Ramanagopal, Aswin C. Sankaranarayanan, and Srinivasa G. Narasimhan. Ordinality of visible-thermal image intensities for intrinsic image decomposition, 2025. 2

[45] Kaihua Zhang, Yinxin He, Kun Yu, and Yufang Liu. Fast data processing method for multispectral radiation thermometry based on euclidean distance optimization. *Opt. Express*, 32(2):1342–1356, 2024. 2

[46] Nannan Zhang, Jian Xing, and Shuanglong Cui. Generalized inverse matrix-graphic deep learning algorithm for multispectral pyrometer temperature inversion. *Opt. Express*, 31 (23):38038–38048, 2023. 2

[47] Zhuomin M. Zhang and Graham Machin. Chapter 1 overview of radiation thermometry. In *Radiometric Temperature Measurements: I. Fundamentals*, pages 1–28. Academic Press, 2009. 1, 2

[48] Lintao Zheng, Hengliang Shi, and Ming Gu. Infrared traffic image enhancement algorithm based on dark channel prior and gamma correction. *Modern Physics Letters B*, 31(19-21):1740044, 2017. 2

[49] Xinyu Zhou, Peiqi Duan, Boyu Li, Chu Zhou, Chao Xu, and Boxin Shi. Evdig: Event-guided direct and global components separation. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9612–9621, 2024. 2

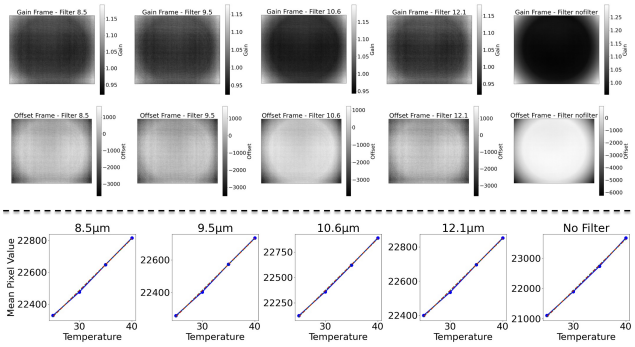

Figure 7. Top: Shows the gain and offset correction images for the various spectral bands. Unlike [22] which only performs offset correction, we perform both gain and offset correction for accurate recovery of incoming radiation. Bottom: Shows blackbodies at different temperatures and their corresponding pixel values in blue, with the exponential Sakuma-Hattori fit [29] and linear fit for the pixel values is shown in dotted red and green respectively. As seen, both temperature to camera counts curve looks linear and the difference between the Sakuma-Hattori and linear curve is less than a couple of counts, which is much below the noise floor of the thermal camera.

A. Calibration and Pre-Processing

In this section, we provide more details on our camera calibration procedure.

Gain-Offset Correction: We begin by applying gain and offset correction using the gain and offset frames shown in Figure 7. To calibrate, we position a blackbody at various temperatures in front of the camera, ensuring it fills the entire field of view. The true pixel value is determined as the mean pixel count across the entire image. For each pixel, we compute gain and offset values so that the corrected output matches this mean. This calibration is performed in a least-squares sense across multiple blackbody temperatures.

Thermal Camera Calibration: Similar to gain-offset correction, we use blackbody videos at different temperatures to perform radiometric calibration of thermal cameras. Figure 7 presents the calibration results across various spectral bands, comparing a Planckian Sakuma-Hattori fit [29] with a linear fit. As observed, the camera response is nearly linear for objects at ambient temperatures.

B. More Experimental Results

Ablation Studies: Figure 8 presents the ablation results of our optimization, where each loss term is individually removed from the pipeline. As shown, the reconstruction loss has the most significant impact, while the smoothing and Huber losses have a comparatively smaller effect on overall accuracy. Initially, optimizing noise parameters does not provide a noticeable improvement, but at lower SNR values, failing to account for noise leads to higher errors. Fig. 8 represents the SNR level breakdown of Tab. 1 shown in the main paper.

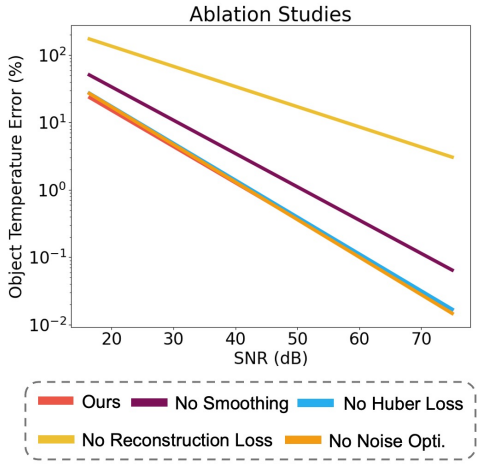

Figure 8. Ablating the loss terms in simulation: We individually disable each loss term from the full optimization pipeline. Note the log scale on the plots. Reconstruction loss is obviously the most important term followed by smoothing, huber and noise optimization. The noise term is useful at high noise levels.

Breakdown of Simulation Evaluation: Fig. 10 provides a detailed breakdown of the simulation results summarized in Fig. 2 in the main paper.

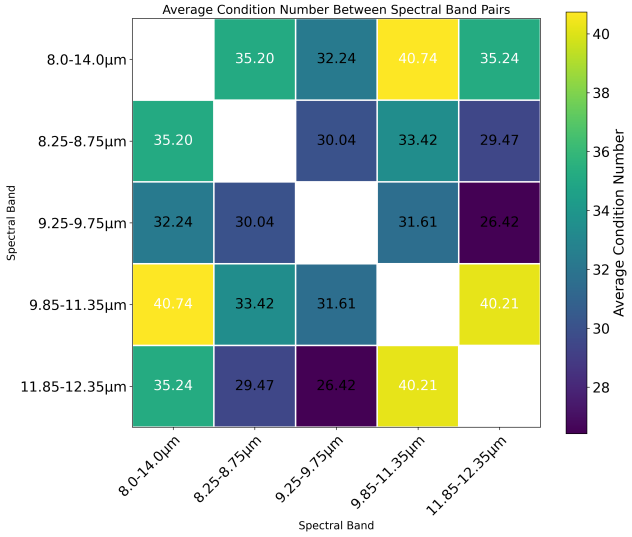

Figure 9. Condition numbers of emissivity matrix $\mathbf{E}$ in Eq. 15 for different spectral band pairs, averaged over materials in the spectral library [3, 21].

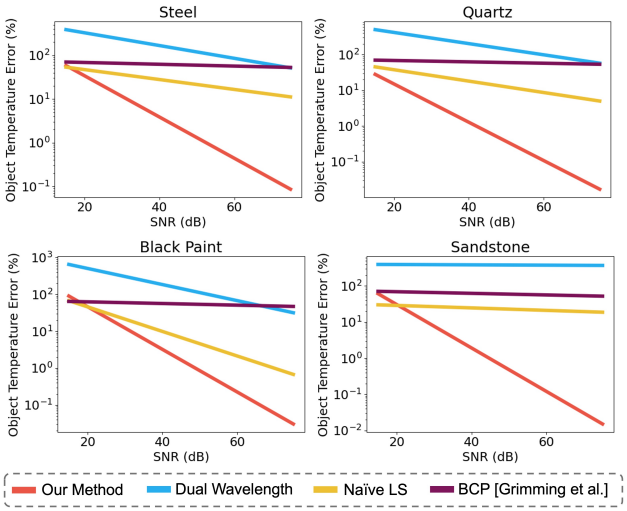

Figure 10. Comparison of our method with a recent BCP [9] technique to remove reflections, a naive multi-wavelength approach and a traditional dual-wavelength pyrometry technique [2, 39] using simulated thermal videos of different materials sourced from spectral library [3, 21]. For naive least squares, we run the optimization with five initializations and select one that achieved the least objective compared to ground truth (which we will not have access to at test time). At high noise levels, all methods have a large error as the problem is too under constrained. As noise decreases to more reasonable levels, our method performs significantly better. Note the log scale on the plots.

Choice of Spectral Filters: The selection of spectral filters for thermography plays a crucial role in distinguishing emitted and reflected light. Figure 9 illustrates the condition number of the matrix $\mathbf{E}$ from Eq. 15 for different spectral filter choices in our experiments, using emissivities sourced from the spectral library [21]. The choice of spectral bands involves a trade-off between the condition number of the resulting emissivity matrix and the noise introduced by each filter. Generally, wider spectral bands reduce noise levels; hence, we select 8–14μm as one of the spectral bands. Additionally, Figure 9 shows that the 9.5μm central wavelength filter achieves the lowest condition number, indicating a better-conditioned linear system for our application.