1. Lookup Table for Passband-Integrated Planck Radiance

We precompute a lookup table (LUT) for the passband-integrated Planck radiance $\mathcal{B}(T)$ and its temperature derivative $\mathcal{B}'(T)$ over the temperature range of interest. Starting from Planck's law in the wavelength domain,

$$B_\lambda(T) = \frac{2hc^2}{\lambda^5} \frac{1}{\exp\left(\frac{hc}{\lambda kT}\right) - 1}, \tag{1}$$

with Planck constant h, Boltzmann constant k, speed of light c, wavelength λ (in meters), and temperature T (in Kelvin), the passband-integrated radiance for an LWIR camera with an (approximately) rectangular passband is

$$\mathcal{B}(T) = \int_{\lambda_{\min}}^{\lambda_{\max}} B_\lambda(T)\, \mathrm{d}\lambda, \quad \lambda_{\min} = 7.5 \times 10^{-6}\ \mathrm{m},\ \lambda_{\max} = 13.5 \times 10^{-6}\ \mathrm{m}. \tag{2}$$

For backpropagation through temperature we also need the derivative of $\mathcal{B}(T)$ with respect to T. Differentiating Eq. (1) gives the temperature derivative of the spectral radiance:

$$\frac{\partial B_\lambda(T)}{\partial T} = \frac{2h^2c^3}{\lambda^6 kT^2} \frac{\exp\left(\frac{hc}{\lambda kT}\right)}{\left(\exp\left(\frac{hc}{\lambda kT}\right) - 1\right)^2}. \tag{3}$$

Hence,

$$\mathcal{B}'(T) = \frac{\mathrm{d}\mathcal{B}(T)}{\mathrm{d}T} = \int_{\lambda_{\min}}^{\lambda_{\max}} \frac{\partial B_\lambda(T)}{\partial T}\, \mathrm{d}\lambda. \tag{4}$$

Because the integrals in Eqs. (2) and (4) have no simple closed form, we evaluate them numerically on a discrete (T, λ) grid and store the results in a LUT.

Discretization. Let $\{T_n\}_{n=0}^{N_T-1}$ be a uniform temperature grid covering the range of interest,

$$T_n = T_{\min} + n\,\Delta T, \qquad \Delta T = \frac{T_{\max} - T_{\min}}{N_T - 1}.$$

Similarly, we discretize the passband $[\lambda_{\min}, \lambda_{\max}]$ into N_λ samples:

$$\lambda_j = \lambda_{\min} + j\,\Delta\lambda, \Delta\lambda = \frac{\lambda_{\max} - \lambda_{\min}}{N_\lambda - 1}, j = 0, \ldots, N_\lambda - 1.$$

For each temperature T_n and wavelength sample λ_j we evaluate $B_{\lambda_j}(T_n)$ using Eq. (1) and $\partial B_{\lambda_j}(T_n)/\partial T$ using Eq. (3). We then approximate the passband integrals by a simple composite quadrature rule, the trapezoidal rule:

$$\mathcal{B}(T_n) \approx \sum_{j=0}^{N_\lambda - 1} w_j\, B_{\lambda_j}(T_n), \tag{5}$$

$$\mathcal{B}'(T_n) \approx \sum_{j=0}^{N_\lambda - 1} w_j\, \frac{\partial B_{\lambda_j}(T_n)}{\partial T}, \tag{6}$$

where w_j are the trapezoidal weights ($w_0 = w_{N_\lambda - 1} = \frac{1}{2}\Delta\lambda$ and $w_j = \Delta\lambda$ otherwise). Higher-order rules (e.g. Simpson's rule) can be used if desired, but in practice N_λ in the range $[256, 1024]$ already yields negligible quadrature error.

LUT construction and use. The LUT stores pairs $(\mathcal{B}(T_n), \mathcal{B}'(T_n))$ for all grid temperatures $\{T_n\}$. During training, given a predicted temperature T we obtain $\mathcal{B}(T)$ and $\mathcal{B}'(T)$ by 1D interpolation on this table. This provides a numerically stable, differentiable approximation to the passband-integrated Planck radiance within the working range of the LWIR camera.

2. Per-Scene Pseudo-Color Calibration

Many off-the-shelf thermal cameras only provide pseudo-colored outputs together with a color bar, rather than calibrated radiance or temperature images. As stated in the main paper, we perform a one-off per-scene calibration: we sample the on-image color bar to build a color→normalized grayscale (and corresponding display-temperature) lookup, and convert pseudo-color to a single-channel thermal signal per pixel.

Fig. 1(a) illustrates the procedure. We first select a reference pseudo-colored frame with the embedded color bar and manually specify a quadrilateral that roughly encloses the bar. We warp this quadrilateral to a canonical vertical rectangle, extract the central strip of the warped bar, and compute the per-row median RGB value. A 1D Gaussian filter along the vertical axis is then applied to remove small oscillations and contamination from ticks and labels.

From the smoothed per-row colors, we construct a 256-entry lookup table (LUT) indexed by a normalized scalar $s \in [0, 1]$. By construction, $s = 0$ corresponds to the bottom of the bar (the lowest temperature $T_{\min}$) and $s = 1$ to the top of the bar (the highest temperature $T_{\max}$). Fig. 1(b, left) plots the resulting LUT channels as functions of s, i.e., $R(s), G(s), B(s)$; Fig. 1(b, right) shows the induced grayscale–temperature mapping

$$T(s) = T_{\min} + s\,(T_{\max} - T_{\min}),$$

where $(T_{\min}, T_{\max}) = (17.5°\mathrm{C}, 56.3°\mathrm{C})$ for this scene. This plot confirms that our calibration defines a monotonic mapping in which $s = 1$ corresponds to the hottest (nearly white) end of the palette and $s = 0$ to the coldest (dark) end.

Given this LUT, we invert pseudo-colored frames framewise. For each pixel, we measure the distance between its RGB value and all LUT entries in CIE–Lab space, find the nearest LUT index and its neighbor, and use distance-weighted interpolation to obtain a continuous scalar $\hat{s}(x, y) \in [0, 1]$. We treat $\hat{s}(x, y)$ as a linearized single-channel thermal signal, and optionally map it back to pseudo-color via the same LUT for visualization. On

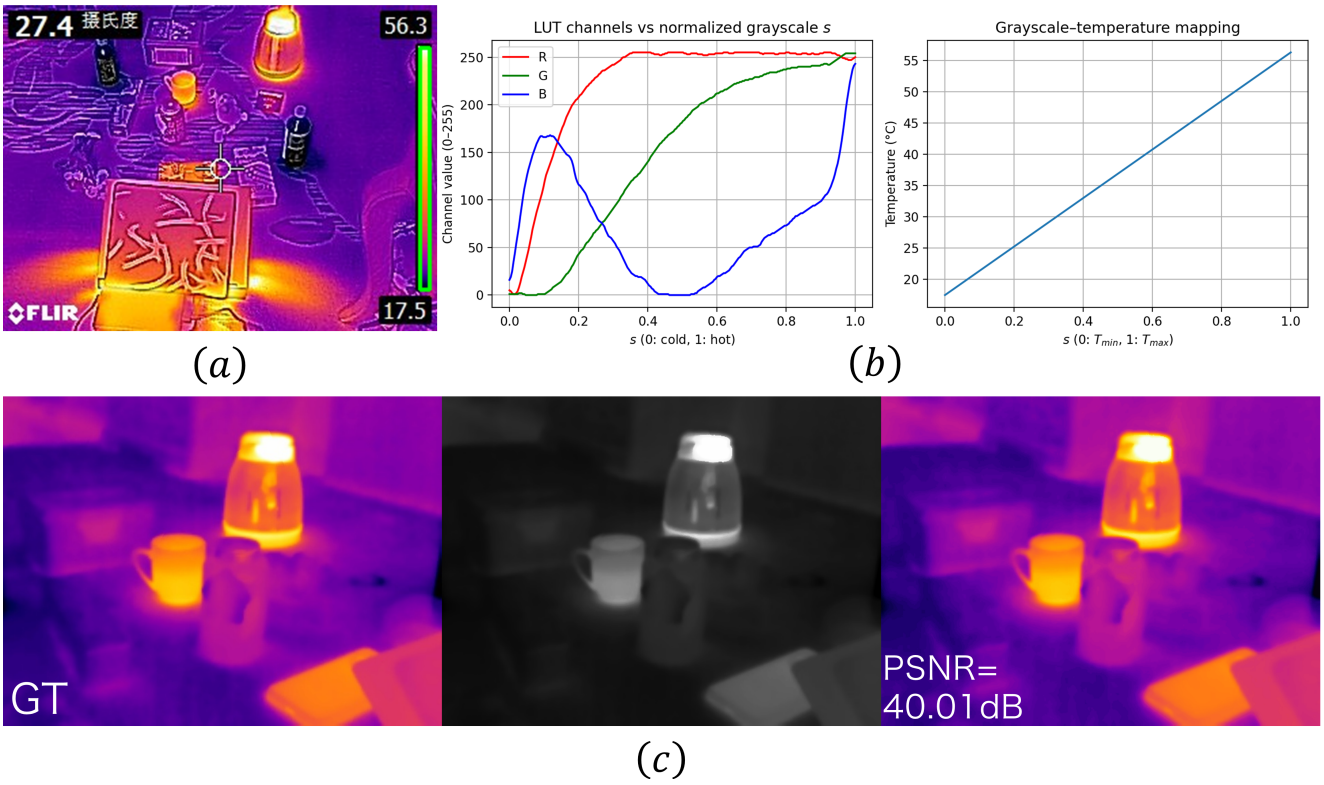

Figure 1. **Per-scene pseudo-color calibration.** (a) Reference pseudo-colored frame with the embedded color bar and the quadrilateral used for warping. (b) Fitted mapping from normalized grayscale $s \in [0, 1]$ to LUT channels $R(s), G(s), B(s)$ (left) and the induced linear grayscale–temperature relation $T(s)$ (right). Here $s = 0$ corresponds to $T_{\min} = 17.5°\mathrm{C}$ and $s = 1$ to $T_{\max} = 56.3°\mathrm{C}$. (c) Qualitative verification of the inverse mapping: original pseudo-color input (left), recovered single-channel thermal map $\hat{s}(x, y)$ (middle), and pseudo-colored reconstruction using our LUT (right).

the RGBT-Scenes dataset, this inverse mapping yields reconstructed pseudo-colored images whose per-scene PSNR with respect to the original camera outputs consistently exceeds 40 dB,indicating that the LUT inversion error is small for our purposes. Fig. 1(c) shows a qualitative check: from left to right, we display the original pseudo-colored input, the recovered grayscale thermal map $\hat{s}(x, y)$, and the pseudo-colored reconstruction obtained by re-applying the LUT. The close visual agreement between the first and third images further confirms that the colorbar-based calibration is sufficiently accurate for our pipeline.

3. Emissivity Consistency Validation: Per-Material Statistics

Figure 7 in the main text plots, for each material class, the estimated emissivity against a typical reference value in the 7.5–13 µm band. Most common diffuse or near-diffuse materials (e.g., concrete, brick, wood, soil, grass, plastic, porcelain) cluster close to the diagonal, while only a few classes with strong specular or partially transparent behavior (glass, steel) exhibit noticeable deviations. This indicates that our method provides broadly reasonable emissivity estimates relative to typical literature values for the majority of common diffuse materials in the scene.

Table 1 reports, for each material, the typical reference emissivity (Lit.Ref.), the mean estimated emissivity (Est.), the per-patch RMSE, and the number of sampled patches *N*. Across all diffuse materials (736 patches in total), the average RMSE is 0.089, and the bias |Est. − Lit.Ref.| is below 0.08 for all classes. In contrast, highly specular or partially transparent materials such as glass and steel (134 patches) exhibit a higher RMSE of 0.167. This behavior is expected, as our current formulation assumes predominantly diffuse reflectance and does not explicitly model specular reflection or transmission in the TIR band.

From the perspective of the underlying 3D Gaussian Splatting representation, strongly specular or (semi-)transparent materials remain challenging: the original 3DGS model is mainly optimized for mostly opaque, diffuse scenes with a single alpha-blended layer of Gaussians and low-order view-dependent color, and in practice it often fails to accurately reproduce mirror-like reflections, refraction, and transmission, which has motivated a number of

Table 1. Emissivity consistency validation on the RGBT-Scenes dataset. Lit.Ref. denotes a typical 7.5–13 µm emissivity value from standard material charts. Est. is the mean estimated emissivity over all patches of that class.

Material	Lit.Ref.	Est.	RMSE	*N*
Diffuse / near-diffuse materials				
Concrete	0.95	0.9621	0.0731	*148*
Porcelain	0.92	0.8557	0.0834	*93*
Brick	0.90	0.9069	0.0869	*74*
Wood	0.90	0.8948	0.1027	*107*
Soil	0.96	0.9325	0.0961	*127*
Grass	0.98	0.9950	0.0896	*129*
Plastic	0.95	0.8738	0.1053	*58*
Specular / partially transparent materials				
Glass	0.85	0.7791	0.1529	*68*
Steel	0.80	0.7382	0.1817	*66*

extended variants for reflective and transparent objects.

In our TIR setting, these glass and steel regions usually occupy only a small fraction of the scene (e.g., windows and metal fixtures), and the missing specular/transparent component is partly absorbed into the learned self-emission parameters (temperature and emissivity). As a result, the estimated emissivity for these materials should be interpreted as an effective parameter that balances emission and unresolved reflections, rather than an exact match to tabulated diffuse emissivity values, while the overall TIR radiance remains stable and consistent with that observed in Fig. 6 of the main text.