Carnegie Mellon University
vt-intrinsic.github.io

VT-Intrinsic: Physics-based Decomposition of Reflectance and Shading from a Single Visible-Thermal Image Pair

Zeqing (Leo) Yuan, Mani Ramanagopal, Aswin Sankaranarayanan, Srinivasa Narasimhan

CVPR JUNE 3-7, 2026 DENVER COLORADO

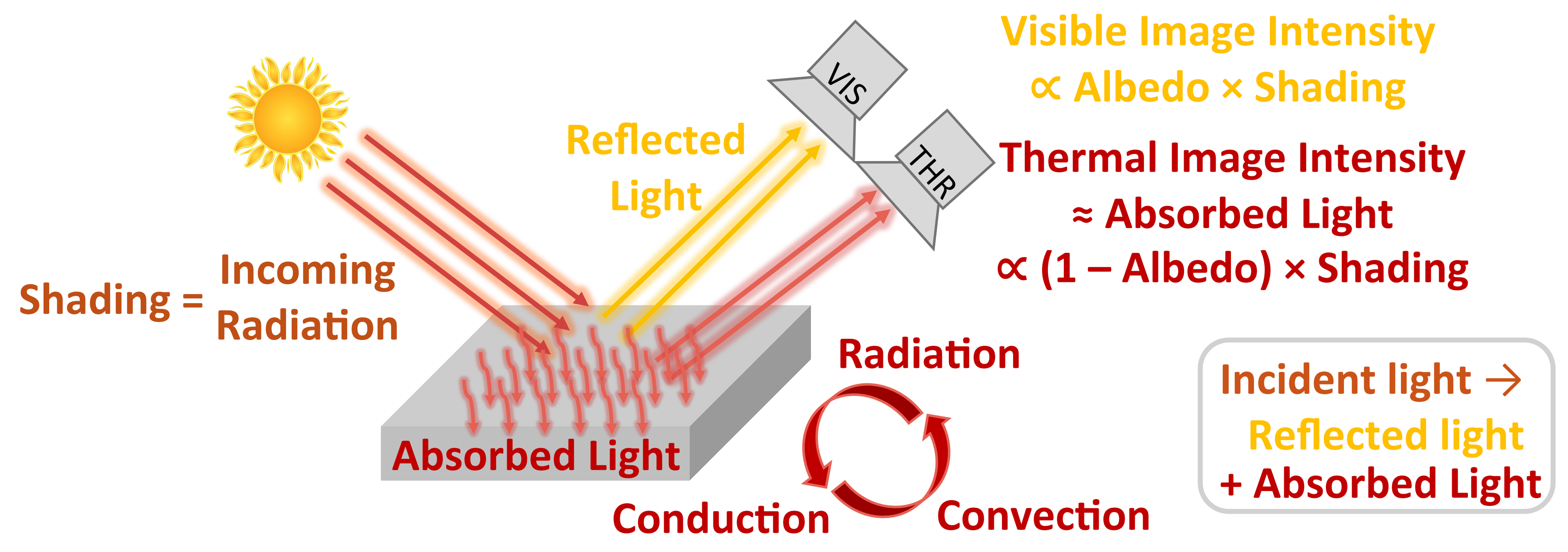
Task & Motivation

Goal: Decomposing Scene Appearance into Lighting and Material Reflectance

Input image: $I_v = \rho \cdot \eta \rightarrow$ albedo (ρ) and shading (η)

Issue: the lack of real-world radiometric ground truth

Theory: Energy Conservation in Light and Heat Transport



Light & Heat Transport ⇒ Albedo / Shading Ordinalities

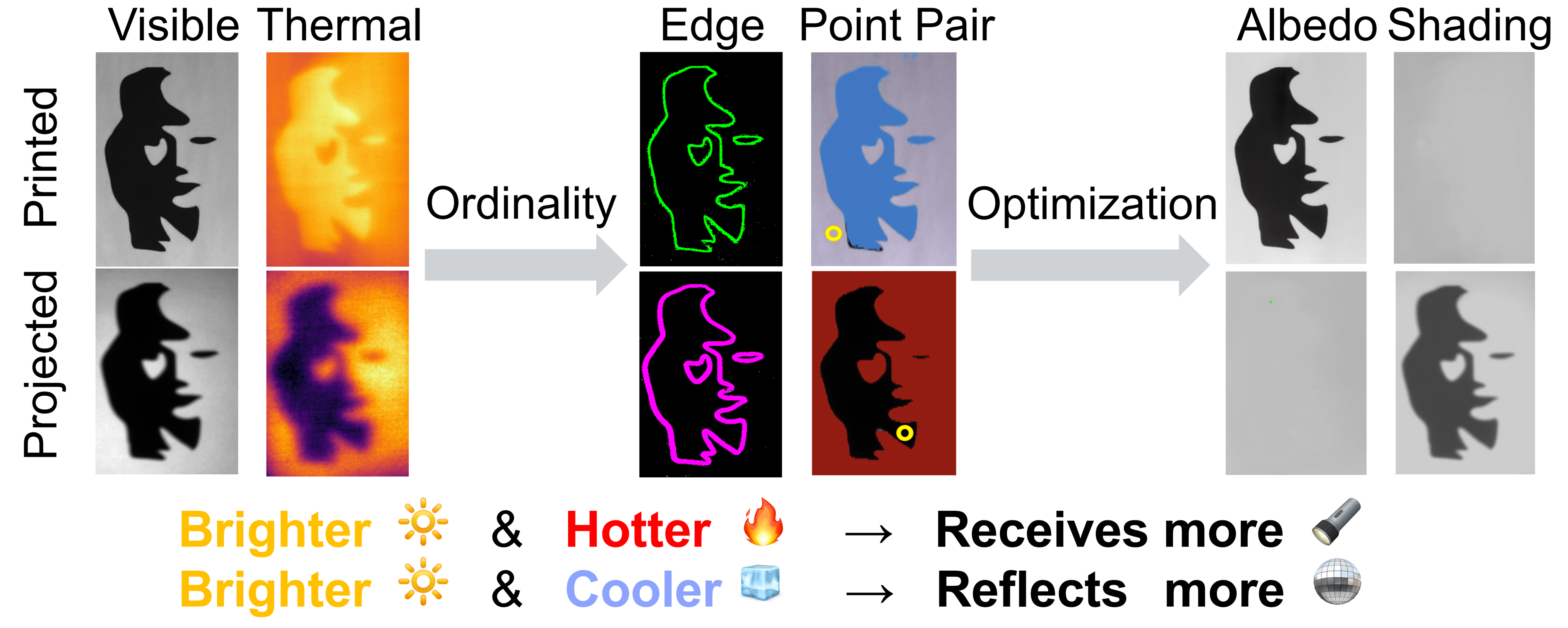
$$\begin{cases} I_v = g \cdot \rho \cdot \eta, & \text{Visible image} \\ S = (1-\rho) \cdot \eta, & \text{Absorbed heat} \end{cases} \Rightarrow \begin{cases} I_v(x_i) > I_v(x_j), S(x_i) > S(x_j) \Rightarrow \eta(x_i) > \eta(x_j) \\ I_v(x_i) > I_v(x_j), S(x_i) < S(x_j) \Rightarrow \rho(x_i) > \rho(x_j) \end{cases}$$

$$\begin{cases} \overset{\text{Heat capacity}}{H} \frac{\partial T}{\partial t} = S + \underbrace{\boxed{hc\,(T_a - T)}}_{\text{Convection}} + \underbrace{\boxed{4\epsilon\sigma T_s^{\,3}\,(Ts - T)}}_{\text{Radiation}} + \underbrace{\boxed{\kappa\,\Delta T}}_{\text{Conduction}}, & \text{Heat transport} \\ I_t = \underbrace{\boxed{\epsilon\, U(T)}}_{\text{Emission}} + \underbrace{\boxed{(1-\epsilon) U(T_s)}}_{\text{Reflection}}, & \text{Thermal image formation} \end{cases}$$

$\Rightarrow S = c_1 I_t + c_2 \Delta I_t + c_3$, where c_1, c_2, c_3 are locally constant and $I_t \gg \Delta I_t$.

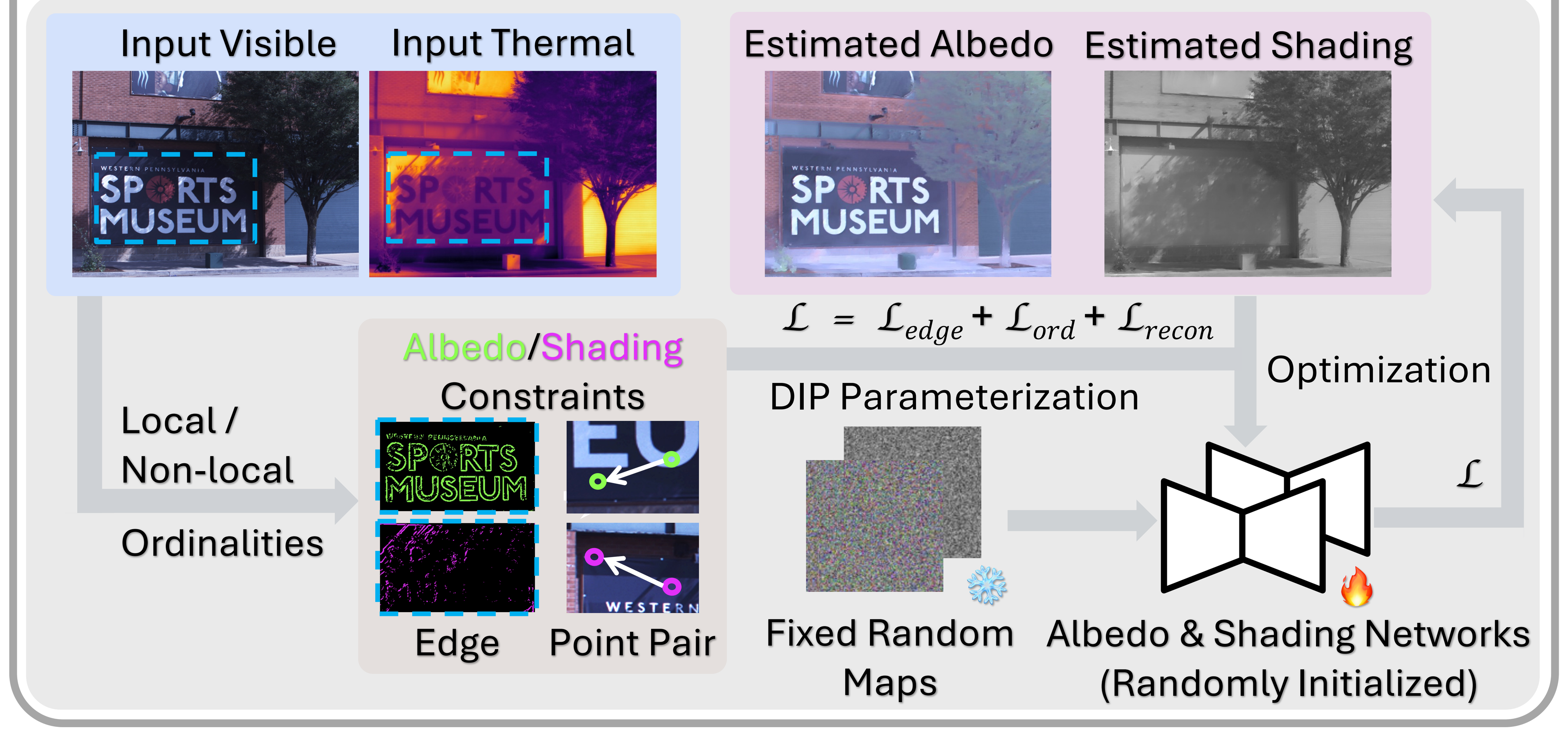
Hence, ordinalities preserve when substituting S with I_t.

Method: Visible/Thermal → Albedo/Shading Ordinalities

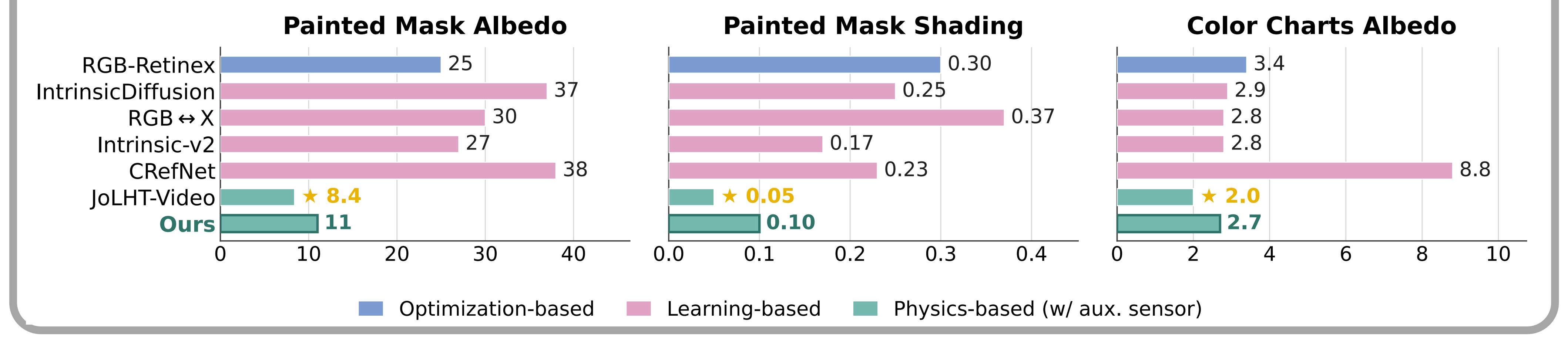


Brighter & Hotter → Receives more

Brighter & Cooler → Reflects more



Quantitative Results



Evaluation on JoLHT-Video Dataset (si-MSE $\times 10^{-2}$)

Method	Painted Mask Albedo	Painted Mask Shading	Color Charts Albedo
RGB-Retinex	25	0.30	3.4
IntrinsicDiffusion	37	0.25	2.9
RGB ↔ X	30	0.37	2.8
Intrinsic-v2	27	0.17	2.8
CRefNet	38	0.23	8.8
JoLHT-Video	★ 8.4	★ 0.05	★ 2.0
Ours	**11**	**0.10**	**2.7**

Optimization-based (RGB-Retinex); Learning-based (IntrinsicDiffusion, RGB ↔ X, Intrinsic-v2, CRefNet); Physics-based (w/ aux. sensor) (JoLHT-Video, Ours)

Qualitative Results