

3D Surface Reconstruction via the Photometric Stereo Method

JC Vaught and Ty Dangerfield

Department of Mechanical Engineering, University of South Carolina

December 7, 2025

Abstract

We present an analytical and numerical validation of the photometric stereo method used in industrial applications for defect detection. This paper details six primary contributions toward a synthetic photometric pipeline: (1) eight unique surface geometries, (2) rigorous mathematical solutions for both numerical and analytical approaches, (3) detailed algorithmic explanations for each primary component of the pipeline, (4) experimental validation across eight shapes and sixteen light-source locations, (5) ablation studies on number of lights, noise robustness, and resolution impacts, and (6) a comparison of two alternative solver methods and their impact on results. Three-dimensional reconstruction depth errors range from 0.022 through 0.147, well within acceptable tolerances and demonstrating robust recovery. Angular errors of the normals reconstruction remain below 3.5° for smooth surfaces and 2° for polyhedral shapes.

Table 1: Individual Contributions

Contribution	Primary Contributor(s)
MATLAB implementation of complete program	Ty Dangerfield
Python implementation of complete program	JC Vaught
Photometric principle derivation	Ty Dangerfield & JC Vaught
Gradient integration with Poisson solver	Ty Dangerfield
Separation of variables and numerical derivations	JC Vaught
Report chapters 1, 3, and 4 (initial)	JC Vaught
Report chapters 1, 2, and 5 (initial)	Ty Dangerfield
Report restructuring and enhanced documentation	JC Vaught & Ty Dangerfield
Code architecture refinement and expansion	JC Vaught & Ty Dangerfield

Table 2: Meeting Participation

Date	JC Vaught	Ty Dangerfield
19 Nov 2025	Present	Present
20 Nov 2025	Present	Present
07 Dec 2025	Present	Present

Contents

1	Introduction	3
1.1	What We Want to Model	3
1.2	Why This Problem is Important	3
1.3	Comparison with Alternative 3D Reconstruction Methods	4
1.4	The Core PDE Problem	4
2	Mathematical Foundation	5
2.1	Photometric Stereo Principle	5
2.1.1	Image Formation Model	5
2.1.2	Lambertian Reflectance	6
2.1.3	Least-Squares Normal Recovery	6

1 Introduction

1.1 What We Want to Model

Three-dimensional surface reconstruction is a fundamental problem in manufacturing applications and mechanical engineering more broadly. Photometric stereo is one possible solution among many, yielding highly accurate surface reconstruction compared to a dual-camera setup, which is more often used in robotic applications.

Given multiple 2D images of an object illuminated from different directions, we derive the PDE, analytical solution, and numerical solution that create an accurate reconstruction of the 3D depth map or surface geometry.

1.2 Why This Problem is Important

Photometric stereo has impact across multiple domains summarized in Table 3, including industrial inspection, where it supports surface defect detection and tolerance verification; medical imaging, where it enables high-fidelity reconstructions of anatomical surfaces for planning; archaeology, where non-destructive digitization preserves fragile artifacts; robotics and autonomous systems, where detailed surface geometry informs grasping and navigation; and reverse engineering, where recovered shapes seed CAD models for subsequent design and analysis.

Table 3: Representative Application Domains for Photometric Stereo

Domain	Example Use
Industrial inspection	Surface defect detection, geometric tolerance checks
Medical imaging	Preoperative surface reconstruction for planning
Archaeology	Non-destructive digitization of artifacts
Robotics	Shape-aware grasp planning and manipulation
Autonomous vehicles	High-fidelity surface maps for navigation
Reverse engineering	Recovering geometry for CAD modeling

Photometric stereo is an ideal solution for many surface extraction tasks owing to the reduction in moving parts. For example, in structure-from-motion methods, the movement of the camera must be precisely monitored and controlled, necessitating careful calibration; even minor calibration errors can severely impact reconstruction quality. Stereo cameras avoid motion but still suffer from calibration issues because the distance, angle, and focal length of each camera—the intrinsic parameters—must be known precisely to estimate depth. Any error there can distort the recovered geometry due to the trigonometric basis of the method. Photometric stereo, on the other hand, is a completely solid-state approach with a single camera, operates at a higher rate than structure-from-motion, and has lower data requirements than stereo systems. Because it leverages illumination direction to infer depth from shading and then reconstruct normals, it is inherently more tolerant to noise and avoids multi-camera calibration drift.

However, photometric stereo does suffer from one critical constraint: the illumination sequence must be precisely controlled. As a result, it cannot be used effectively in uncontrolled environments such as UAVs or general-purpose robotics, which is why stereo cameras remain the de facto standard outside manufacturing despite their calibration sensitivity. Nevertheless, the fundamental challenge in photometric stereo still lies in *integrating* noisy normal estimates into a globally consistent 3D surface, which requires solving a partial differential equation: the **Poisson equation**.

1.3 Comparison with Alternative 3D Reconstruction Methods

Table 4 compares photometric stereo with other common 3D reconstruction techniques across key performance dimensions. The comparison highlights why photometric stereo is particularly well-suited for controlled industrial environments despite requiring specialized lighting infrastructure.

Table 4: Comparison of 3D Reconstruction Methods

Method	Advantages	Disadvantages	Typical Use Cases
Photometric Stereo	Single camera; high spatial resolution; dense surface normals; solid-state (no moving parts)	Requires controlled lighting; assumes Lambertian reflectance; limited to near-field	Defect detection, material inspection, quality control
Stereo Vision	Passive (no special lighting); works outdoors; real-time capable	Requires precise calibration; texture-dependent; ambiguity in textureless regions	Robotics, autonomous vehicles, SLAM
Structure from Motion	Single camera; scales to large scenes; robust to illumination changes	Requires camera motion; computationally intensive; drift accumulation	3D scanning, mapping, archaeological digitization
Structured Light	High accuracy; active illumination; fast acquisition	Sensitive to ambient light; limited range; projector-camera calibration required	Industrial metrology, 3D printing, facial scanning
Time-of-Flight	Direct depth measurement; real-time; works in low texture	Lower resolution; sensitive to multipath reflections; expensive hardware	Gesture recognition, indoor navigation, human-computer interaction

Photometric stereo excels in industrial settings where lighting can be tightly controlled and surface finish is relatively uniform. The method’s ability to capture dense normal fields from a single viewpoint makes it ideal for in-line inspection systems where throughput and repeatability are critical. By contrast, methods like stereo vision or time-of-flight are better suited to unstructured environments (e.g., outdoor robotics) where lighting variability and scene dynamics preclude the use of synchronized illumination sequences.

1.4 The Core PDE Problem

The fundamental problem reduces to solving the 2D Poisson equation in a rectangular domain:

$$\nabla^2 z(x, y) = f(x, y), \quad (x, y) \in \Omega = [0, L_x] \times [0, L_y] \quad (1)$$

where

$$z(x, y) = \text{unknown surface height over the image domain}$$

$$f(x, y) = \frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} \quad (\text{divergence of Poisson-derived gradient estimates})$$

$$\Omega = [0, L_x] \times [0, L_y] \quad (\text{rectangular image domain})$$

This project demonstrates both analytical (separation of variables) and numerical (FFT-based and finite-difference) methods to solve this fundamental PDE and validate the solutions on synthetic 3D surfaces.

2 Mathematical Foundation

This chapter establishes the theoretical foundations for photometric stereo reconstruction. We begin with the photometric stereo principle, deriving the relationship between observed image intensities and surface normals under Lambertian reflectance. We then formulate the gradient integration problem as a Poisson equation, providing both variational and Euler-Lagrange derivations. Next, we develop the theory of boundary conditions—Dirichlet, Neumann, and periodic—and their impact on solution uniqueness. Finally, we present an analytical solution via separation of variables, including eigenfunction expansions and worked examples that validate our numerical methods.

2.1 Photometric Stereo Principle

Photometric stereo recovers surface normals from multiple images of a static scene illuminated by different light sources. The key insight is that shading variations across images encode the local surface orientation. Table 5 summarizes the key modeling assumptions underlying this approach.

Table 5: Photometric Stereo Modeling Assumptions

Assumption	Description
Lambertian reflectance	Surface reflects light diffusely in all directions; no specular highlights or glossy reflections
Orthographic projection	Camera uses parallel projection (unit focal length); no perspective distortion
Known light directions	Light source directions \mathbf{L}_i are calibrated and known a priori
Static scene	Surface geometry and camera remain fixed across all images
Uniform albedo	In our derivations, albedo $\rho(x, y)$ is assumed constant, though the general formulation allows spatially varying albedo
No inter-reflections	Light reaches the surface directly without bouncing off other surfaces
Attached shadows only	Self-shadowing is handled via the max operator; cast shadows from other objects are not modeled

2.1.1 Image Formation Model

By assuming an ideal camera with a unit focal length, we map a 3D point \mathbf{X} to image coordinates $\mathbf{x} = (x, y)$ for any arbitrary surface with depth $z(x, y)$. The measured image intensity I_i at pixel (x, y) under light source i is the product of three components: the camera calibration factor κ_i (encapsulating exposure, gain, and light source intensity), the material albedo $\rho(x, y) \in [0, 1]$ (the fraction of light reflected by the surface), and the geometric factor $\max(0, \mathbf{n}(x, y) \cdot \mathbf{L}_i)$ (the cosine of the angle between the surface normal \mathbf{n} and light direction \mathbf{L}_i). These combine to yield the image formation equation:

$$I_i(x, y) = \kappa_i \rho(x, y) \max(0, \mathbf{n}(x, y) \cdot \mathbf{L}_i) \quad (2)$$

The max operator enforces attached shadows: when a light illuminates the surface from behind ($\mathbf{n} \cdot \mathbf{L}_i < 0$), the contribution is zero.

2.1.2 Lambertian Reflectance

The Lambertian assumption states that a surface appears equally bright from all viewing directions—light is reflected diffusely in all directions. Mathematically, the reflected radiance is proportional to $\cos \theta = \mathbf{n} \cdot \mathbf{L}$, where θ is the angle between the surface normal and the incident light direction.

To connect surface geometry to shading, we must express the surface normal in terms of the depth map $z(x, y)$. We begin by computing the tangent vectors to the surface. These tangent vectors represent the local orientation of the surface along the x and y coordinate directions:

$$\frac{\partial \mathbf{X}}{\partial x} = (1, 0, p), \quad \text{where } p = \frac{\partial z}{\partial x} \quad (3)$$

$$\frac{\partial \mathbf{X}}{\partial y} = (0, 1, q), \quad \text{where } q = \frac{\partial z}{\partial y} \quad (4)$$

Here, p and q represent the local slopes of the surface in the x and y directions, respectively. Note that the first two components of each tangent vector are $(1, 0)$ and $(0, 1)$ because we move one unit along the respective image coordinate while the surface height changes by p or q .

The surface normal is perpendicular to both tangent vectors and can be computed via their cross product:

$$\tilde{\mathbf{n}} = \frac{\partial \mathbf{X}}{\partial x} \times \frac{\partial \mathbf{X}}{\partial y} = (-p, -q, 1) \quad (5)$$

This unnormalized normal $\tilde{\mathbf{n}}$ points away from the surface but does not necessarily have unit length. The negative signs on p and q arise naturally from the cross product and ensure the normal points outward (away from the surface) rather than inward.

To obtain a unit normal vector (required for computing the geometric factor $\mathbf{n} \cdot \mathbf{L}$), we divide by the magnitude:

$$\mathbf{n} = \frac{\tilde{\mathbf{n}}}{\|\tilde{\mathbf{n}}\|} = \frac{(-p, -q, 1)}{\sqrt{p^2 + q^2 + 1}} \quad (6)$$

Equation (6) explicitly links the integrable gradient field (p, q) to the observed shading—this is the key relationship that enables gradient recovery from images.

2.1.3 Least-Squares Normal Recovery

Having established the relationship between surface normals and image intensities, we now formulate the inverse problem: given multiple images under different illumination, recover the surface normal at each pixel.

The Linear System Formulation. We introduce the scaled normal vector $\mathbf{g}(x, y) = \rho(x, y)\mathbf{n}(x, y)$, which combines the geometric information (normal direction) with the material property (albedo). This factorization is convenient because \mathbf{g} appears linearly in the image formation equation.

Substituting the Lambertian reflectance model (6) into the image formation equation (2), we obtain for a single light source i :

$$I_i(x, y) = \kappa_i \mathbf{L}_i^T \mathbf{g}(x, y) \quad (7)$$

where we have absorbed the max operator into the formulation (assuming no self-shadowing for now; we address this later). This is a scalar equation linear in \mathbf{g} .

For m different light sources, we obtain m such equations at each pixel. Stacking these into a single matrix equation yields:

$$\underbrace{\begin{bmatrix} I_1(x, y) \\ I_2(x, y) \\ \vdots \\ I_m(x, y) \end{bmatrix}}_{\mathbf{I}(x, y) \in \mathbb{R}^m} = \underbrace{\begin{bmatrix} \kappa_1 \mathbf{L}_1^T \\ \kappa_2 \mathbf{L}_2^T \\ \vdots \\ \kappa_m \mathbf{L}_m^T \end{bmatrix}}_{S \in \mathbb{R}^{m \times 3}} \underbrace{\mathbf{g}(x, y)}_{\in \mathbb{R}^3} \quad (8)$$

Compactly, we write:

$$S\mathbf{g}(x, y) = \mathbf{I}(x, y) \quad (9)$$

The matrix S is called the *illumination matrix* and depends only on the known light directions and calibration factors. The vector $\mathbf{I}(x, y)$ contains the observed intensities. Our goal is to solve for $\mathbf{g}(x, y)$ at every pixel.

For a unique solution to exist, S must have full column rank: $\text{rank}(S) = 3$. This requires at least three images ($m \geq 3$) with non-coplanar light directions. In practice, we use $m > 3$ to obtain an overdetermined system, which provides robustness against noise.

Least-Squares Solution via Pseudoinverse. When $m > 3$, the system (9) is overdetermined and generally has no exact solution due to measurement noise. We instead seek the least-squares solution that minimizes the residual:

$$\hat{\mathbf{g}}(x, y) = \arg \min_{\mathbf{g}} \|S\mathbf{g} - \mathbf{I}(x, y)\|_2^2 \quad (10)$$

Expanding the objective function:

$$\|S\mathbf{g} - \mathbf{I}\|_2^2 = (S\mathbf{g} - \mathbf{I})^T (S\mathbf{g} - \mathbf{I}) = \mathbf{g}^T S^T S \mathbf{g} - 2\mathbf{I}^T S \mathbf{g} + \mathbf{I}^T \mathbf{I} \quad (11)$$

Taking the gradient with respect to \mathbf{g} and setting it to zero yields the *normal equations*:

$$\nabla_{\mathbf{g}} \|S\mathbf{g} - \mathbf{I}\|_2^2 = 2S^T S \mathbf{g} - 2S^T \mathbf{I} = 0 \Rightarrow S^T S \mathbf{g} = S^T \mathbf{I} \quad (12)$$

Assuming S has full column rank, $S^T S$ is invertible, and we obtain:

$$\mathbf{g}(x, y) = \underbrace{(S^T S)^{-1} S^T}_{S^+ \text{ (pseudoinverse)}} \mathbf{I}(x, y) \quad (13)$$

The matrix $S^+ = (S^T S)^{-1} S^T$ is the Moore–Penrose pseudoinverse of S . This solution minimizes the L^2 norm of the residual and is unique when $\text{rank}(S) = 3$.

Once $\mathbf{g}(x, y)$ is computed, we recover the unit normal and albedo by exploiting the factorization $\mathbf{g} = \rho \mathbf{n}$:

$$\hat{\rho}(x, y) = \|\mathbf{g}(x, y)\|_2 \quad (14)$$

$$\hat{\mathbf{n}}(x, y) = \frac{\mathbf{g}(x, y)}{\|\mathbf{g}(x, y)\|_2} \quad (15)$$

The *direction* of \mathbf{g} encodes the surface normal while its *magnitude* equals the albedo.

Conditioning and Noise Sensitivity. The accuracy of the recovered normals depends critically on the *condition number* of S :

$$\kappa(S) = \frac{\sigma_{\max}(S)}{\sigma_{\min}(S)} \quad (16)$$

where σ_{\max} and σ_{\min} are the largest and smallest singular values of S , respectively.

To understand the impact of conditioning, consider the singular value decomposition (SVD) of S :

$$S = U\Sigma V^T \quad (17)$$

where $U \in \mathbb{R}^{m \times 3}$ and $V \in \mathbb{R}^{3 \times 3}$ are orthogonal matrices, and $\Sigma = \text{diag}(\sigma_1, \sigma_2, \sigma_3)$ contains the singular values in descending order.

The pseudoinverse can be expressed as:

$$S^+ = V\Sigma^{-1}U^T \quad (18)$$

If the measured intensities contain noise $\Delta\mathbf{I}$ such that $\|\Delta\mathbf{I}\|/\|\mathbf{I}\| = \epsilon$, the relative error in the recovered \mathbf{g} is bounded by:

$$\frac{\|\Delta\mathbf{g}\|}{\|\mathbf{g}\|} \leq \kappa(S) \cdot \epsilon \quad (19)$$

This shows that the condition number $\kappa(S)$ controls the worst-case amplification of measurement noise. A large condition number indicates that small errors in \mathbf{I} can lead to large errors in \mathbf{g} .

Optimal Light Placement: To minimize noise sensitivity, we seek to minimize $\kappa(S)$, or equivalently, maximize $\sigma_{\min}(S)$. This is achieved when the light directions $\{\mathbf{L}_i\}$ are as uniformly distributed as possible over the unit sphere, ensuring that the columns of S span \mathbb{R}^3 evenly. In practice, lights are often arranged in a hemispherical configuration with uniform azimuthal and elevation spacing.

Overdetermined Systems: Using $m > 3$ images provides two benefits: (1) the least-squares solution averages out random noise, and (2) the system becomes more robust to individual measurements corrupted by saturation or outliers.

Shadow Handling and Robust Estimation. The linear system (9) assumes that all lights illuminate the surface from the front ($\mathbf{n} \cdot \mathbf{L}_i > 0$). When a light source illuminates from behind, *self-shadowing* occurs, and the image formation equation becomes:

$$I_i(x, y) = 0 \quad \text{if } \mathbf{n}(x, y) \cdot \mathbf{L}_i < 0 \quad (20)$$

These shadowed measurements violate the linear model and must be excluded. The standard approach is:

1. For each pixel, identify light sources i for which $I_i(x, y) \approx 0$ (or below a threshold).
2. Remove the corresponding rows from S and entries from $\mathbf{I}(x, y)$.
3. Solve the reduced least-squares problem on the remaining (valid) measurements.

This row deletion strategy preserves the least-squares structure and ensures that only valid measurements contribute to the normal estimate. However, it requires at least three non-shadowed lights at every pixel to maintain $\text{rank}(S) = 3$.

Alternative: Weighted Least Squares. For pixels near shadow boundaries where intensities are low but non-zero, a weighted least-squares approach can be used:

$$\hat{\mathbf{g}} = \arg \min_{\mathbf{g}} \sum_{i=1}^m w_i (I_i - \kappa_i \mathbf{L}_i^T \mathbf{g})^2 \quad (21)$$

where weights w_i are set based on confidence (e.g., $w_i = \max(0, \mathbf{n} \cdot \mathbf{L}_i)$ or based on intensity magnitude). This provides a smooth transition between fully illuminated and shadowed regions.