GUSLO: General and Unified Structured Light Optimization

Tinglei Wan[1,3], **Zhongjie Wang**[1], **Tonghua Su**[1,2,4*]
[1]Harbin Institute of Technology, Harbin, China
[2]Guangdong Laboratory of Artificial Intelligence and Digital Economy (SZ), Shenzhen, China
[3]Harbin Institute of Technology Zhengzhou Research Institute, Zhengzhou, China
[4]Chongqing Research Institute of HIT, Chongqing, China
23b936043@stu.hit.edu.cn, rainy@hit.edu.cn, thsu@hit.edu.cn

Abstract

Structured light (SL) 3D reconstruction captures the precise surface shape of objects, providing high-accuracy 3D data essential for industrial inspection and cultural heritage digitization. However, existing methods suffer from two key limitations: reliance on scene-specific calibration with manual parameter tuning, and optimization frameworks tailored to specific SL patterns, limiting their generalizability across varied scenarios. We propose **G**eneral and **U**nified **S**tructured **L**ight **O**ptimization (**GUSLO**), a novel framework addressing these issues through two coordinated innovations: (1) single-shot calibration via 2D triangulation-based interpolation that converts sparse matches into dense correspondence fields, and (2) artifact-aware photometric adaptation via explicit transfer functions, balancing generalization and color fidelity. We conduct diverse experiments covering binary, speckle, and color-coded settings. Results show that GUSLO consistently improves accuracy and cross-encoding robustness over conventional methods in challenging industrial and cultural scenarios.

Datasets — https://github.com/bragi17/GUSLO

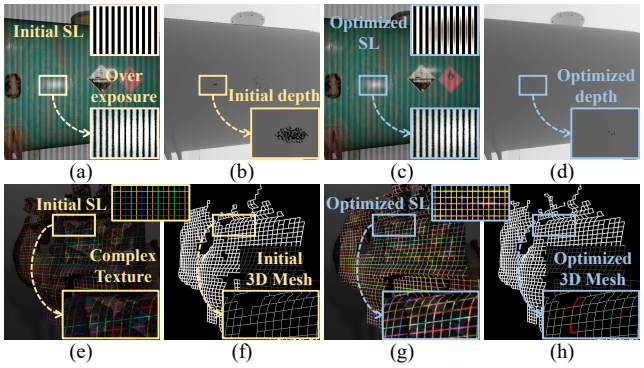

Figure 1: Robust 3D reconstruction with our framework. (a–b) Overexposure disrupts pattern coding, leading to missing geometry. (c–d) Our projection compensates illumination loss, enabling complete reconstruction. (e–f) Surface texture degrades color-coded decoding, causing partial mesh loss. (g–h) Texture-aware chromatic optimization restores decoding and improves reconstruction completeness (highlighted in red).

Introduction

Structured light (SL) 3D reconstruction has become a vital technique in industrial and cultural heritage applications due to its high precision and non-contact nature. In industry, SL enables rapid reverse engineering (Qian et al. 2019; Bak 2003) and real-time quality control (Qian et al. 2021; Sansoni, Trebeschi, and Docchio 2009). In cultural heritage preservation, SL is widely adopted for digitizing and monitoring fragile artifacts, such as historical clothing (Montusiewicz et al. 2021; Ding and Liang 2024), architectural surfaces under climatic change (Holl, Pallas, and Bellendorf 2021), and detailed edge visualization in 3D point clouds (Yamada et al. 2024). It also supports deformation analysis of organic materials (Stelzner et al. 2022), contributing to long-term conservation workflows.

A core challenge in SL lies in optimizing projection patterns under complex lighting and surface textures. Traditional methods focus on two non-adaptive strategies: post-processing camera measurements (e.g., denoising (Dodda et al. 2023), phase correction (Wang et al. 2024)) and heuristic pattern design (e.g., composite (Nguyen et al. 2022), frequency (Li et al. 2022), or color multiplexing (Fu et al. 2024)). These hand-crafted approaches fail to adapt to photometric variations, limiting robustness in uncontrolled environments, as illustrated in Fig. 1.

Projection compensation methods attempt to address this by modeling the projection process to adaptively refine input patterns (Shih et al. 2020; Sugimoto et al. 2021). Though effective in reducing ambient and reflectance-induced color shifts (Fujii, Grossberg, and Nayar 2005; Li et al. 2023), they face two key limitations: (1) they rely on extensive calibration to map surface reflectance to the projector plane, which restricts their applicability to simple geometries; and (2) they require multiple projections per scene, making them impractical for real-time or general-use deployment.

To overcome these challenges, we propose a general and unified SL optimization (GUSLO) framework (Fig. 2) that achieves joint photometric calibration and SL pattern optimization using a single projection. Our approach is based on two key components: (1) continuous surface parameter-

*Corresponding author

ization via barycentric interpolation over triangulated discrete matches, and (2) physics-aware photometric transfer through enhanced thin-plate spline (TPS) warping with adaptive color mapping.

First, the projector-camera global matching process employs planar triangulation with barycentric interpolation to parameterize 3D surfaces into a continuous UV mapping space. This geometric foundation facilitates single-shot calibration via projection-plane conformal mapping. Second, the SL projection compensation process employs an enhanced thin-plate spline (TPS) operator to achieve adaptive artifact-aware photometric transfer. This framework enables dynamic color mapping while preserving gradient coherence across illumination conditions. Our work introduces fundamental advancements in structured light optimization through three key innovations:

- **General and Unified SL Optimization Architecture**: We propose **GUSLO**, a novel structured light optimization framework that jointly performs geometric calibration and photometric compensation under a single-projection setup. Extensive experiments on binary, speckle, and color-coded SL patterns demonstrate GUSLO's superior cross-encoding generality and significant accuracy improvements across diverse industrial and cultural heritage scenarios.
- **Single-Projection Geometric-Photometric SL Optimization**: To the best of our knowledge, GUSLO is the first SL optimization system that enables *simultaneous geometric and photometric calibration using a single structured pattern*. Our design resolves the long-standing adaptability-accuracy trade-off in SL systems by integrating continuous projective-plane parameterization with dynamic, artifact-suppressed color correction.
- **Open and Reproducible Benchmarking System**: We establish an open-source simulation and evaluation platform for SL research, featuring a physically-grounded calibration setup, configurable virtual scenes, and standardized protocols. Built upon Blender and parametric scripting, the system enables automated dataset generation and reproducible benchmarking for future research.

Related Work

Our proposed unified framework for SL optimization consists of two main components: projector-camera global matching and projection compensation. We first introduce related work in these two areas. Additionally, we introduce various recent attempts by researchers to generate SL patterns that are better adapted to the capturing scene.

Projector-Camera Global Matching

Matching methods can be broadly categorized by surface continuity. For planar surfaces, Raskar et al. (Raskar, van Baar, and Chai 2002) introduced a 3×3 response matrix, while Huang et al. (Huang et al. 2018) extended it to curved surfaces via TPS and affine transformations, though requiring scene-specific training data.

For non-continuous surfaces, Gupta et al. (Gupta 2007) proposed composite Gray code patterns, and Pages et al. (Pages et al. 2003) used sinusoidal patterns for phase computation, both requiring multiple projections and incurring time overhead.

Hardware-assisted methods, such as coaxial optics (Fujii, Grossberg, and Nayar 2005) or RGBD cameras with differentiable rendering (Park, Jung, and Moon 2022), can accelerate matching but depend on specialized setups, limiting practical deployment.

Projector Compensation

Traditional compensation approaches modify projector input based on environment lighting. For example, Raskar et al. (Raskar, van Baar, and Chai 2002) derived a color-space compensation function via input-output mapping, but such methods struggle with complex reflectance and often rely on discrete Gray code mappings.

Deep learning approaches improve adaptability. Huang et al. (Huang, Sun, and Ling 2021) separated projection and surface components, and Park et al. (Park, Jung, and Moon 2022) used differentiable rendering to optimize input images. However, both require scene-specific retraining or additional hardware, limiting real-time and general-purpose SL applications.

Structured Light Projection Optimization

Recent work focuses on tailoring SL patterns to capture conditions. Xu et al. (Xu et al. 2023) used LED arrays and LCD masks to learn optimal encoding sets, combining results across LEDs. Dong et al. (Dong, Ling, and Huang 2023) proposed ambient-light-adaptive color SL using MAP-based color detection. Jia et al. (Jia et al. 2024) designed depth-adaptive speckle SL by adjusting gray level, density, and size across sub-regions.

While effective for specific inputs, these methods lack generality. Xu's method targets LED-based SL, Dong's is unsuitable for grayscale, and Jia's focuses solely on speckle. As such, none of them generalize well across diverse SL types or scenes.

Methodology

Our GUSLO framework (Fig. 2) is structured around two core processes: (1) Delaunay-barycentric interpolation to generate dense geometric mappings from sparse correspondences, and (2) a physics-aware compensation scheme that integrates thin-plate spline warping with photometric modeling to address projector nonlinearities and ambient illumination. The optimization flow first builds continuous surface parameterization via the **Projector-Camera Global Matching** stage, then refines SL inputs through the **Structured Light Projection Compensation** stage, enabling cross-scene and cross-encoding optimization from a single projection. To support reproducible evaluation, our **Open Structured Light Benchmarking System** offers hardware-agnostic validation through virtual co-calibration and automated dataset generation.

Projector-camera Global Matching

Matching Pattern Design. Accurate pixel localization between projector and camera is essential for SL pattern

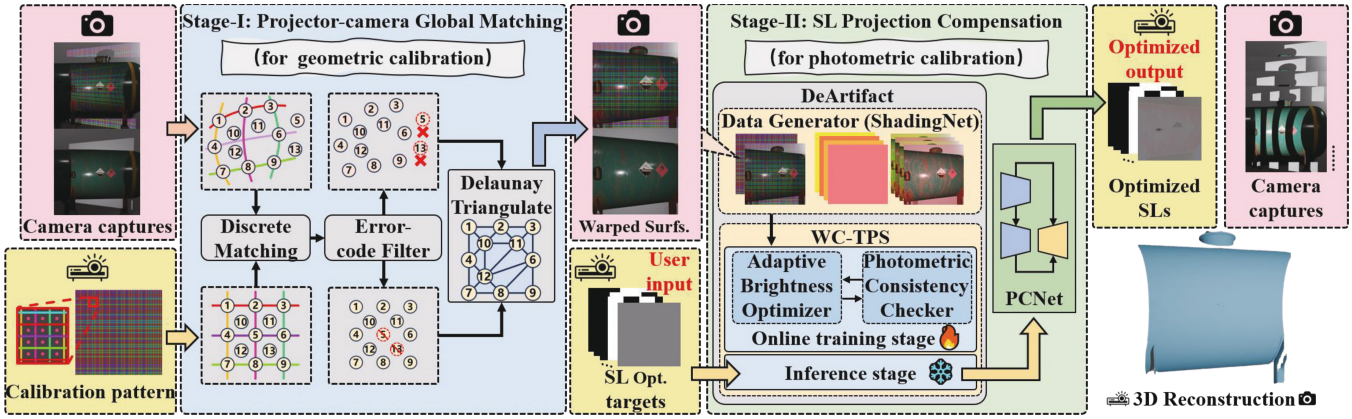

Figure 2: Model details of our GUSLO. Stage-I filters erroneous codes based on relative positions in the camera plane, then applies Delaunay triangulation and barycentric interpolation to extend discrete matches into continuous global correspondences between camera and projector pixels. In Stage-II, decoding artifacts are corrected via the DeArtifact Module, implemented using our WC-TPS method trained with ShadingNet data. The refined results are then processed by PCNet for photometric compensation, yielding optimized structured light patterns for high-precision 3D reconstruction.

optimization. We designed SL images using De Bruijn (Kawasaki et al. 2008) sequences for this purpose. In our method, the horizontal alphabet $A_{hor} = \{1, 3, 5, 7\}$ corresponds to red, lime, cyan, and purple, while the vertical alphabet $A_{ver} = \{2, 4, 6, 8\}$ represents yellow, green, blue, and magenta. With a window length of 3, no sequence of three consecutive lines repeats, allowing unique identification of stripe intersections. To increase the number of matching feature points, orange-red dots are placed at the center of the grid.

To avoid interference with stripe detection, the pattern background color is selected within a moderate RGB range, ensuring color separation from grid elements without causing clipping.

Discrete Matching and Error Code Filtering. Project the calibration pattern (shown in Fig. 2) onto the object surface, where the camera captures the deformed pattern (see Camera Captures in Fig. 2, top-left). Then use De Bruijn coding to match feature points between the projected and captured patterns. This results in discrete sets of projector-camera matching points, $P = \{p_i = (x_p^i, y_p^i),\ i = 1, 2, \ldots, n\}$ and $C = \{c_i = (x_c^i, y_c^i),\ i = 1, 2, \ldots, n\}$. The matching relationship between these points can be described as a bijection

$$M = \{(p_i, c_i) \mid p_i \in P, c_i \in C, i = 1, 2, \ldots, n\} \tag{1}$$

from the projector points P to the camera points C.

During the matching process, De Bruijn code losses in some regions can lead to decoding errors in the grid's central feature points. These errors are filtered out to avoid propagating mismatches.

We compute the horizontal and vertical relative position matrices, X_{map}^p and Y_{map}^p on the projector image plane, and X_{map}^c and Y_{map}^c on the camera image plane, by comparing the relative positions of the feature points (x_p^i, y_p^i) and (x_c^i, y_c^i) with their surrounding eight feature points. Eq. 2 presents the XOR of these matrices, indicating the similarity between the two sets of position encodings.

$$V_x = X_{map}^p \oplus X_{map}^c, \quad V_y = Y_{map}^p \oplus Y_{map}^c. \tag{2}$$

Finally, by performing row summation on V_x and V_y, we can obtain the voting scores for error codes in the x and y analyses, S_x and S_y, for each point. Let the elements in S_x and S_y be S_x^i and S_y^i, respectively. The set of filtered point indices, denoted by *indices*, is defined as:

$$\textit{indices} = \{i \mid i \in \{1, 2, ..., n\} \text{ and } \left(S_x^i \geq k \text{ or } S_y^i \geq k\right)\}. \tag{3}$$

The value of k is related to the number of points n involved in the error code check. Generally, k is set as Floor$(n/2)$, where Floor$(\cdot)$ denotes the floor function.

Global Matching. We employ Delaunay triangulation to convert sparse matches into continuous mappings over the projector image plane. This extends M to every pixel on the projector image plane by triangulating the point set P, forming a continuous triangular mesh where each vertex p_1, p_2, p_3 corresponds to camera image coordinates c_1, c_2, c_3.

For any point $p = (x_p, y_p)$ within a triangular patch, its barycentric coordinates calculated using Eq. 4 to 7.

$$\det T = \left(y_p^2 - y_p^3\right)\left(x_p^1 - x_p^3\right) + \left(x_p^3 - x_p^2\right)\left(y_p^1 - y_p^3\right), \tag{4}$$

$$L_1 = \frac{\left(y_p^2 - y_p^3\right)\left(x_p - x_p^3\right) + \left(x_p^3 - x_p^2\right)\left(y_p - y_p^3\right)}{\det T}, \tag{5}$$

$$L_2 = \frac{\left(y_p^3 - y_p^1\right)\left(x_p - x_p^3\right) + \left(x_p^1 - x_p^3\right)\left(y_p - y_p^3\right)}{\det T}, \tag{6}$$

$$L_3 = 1 - L_1 - L_2. \tag{7}$$

If p lies within the triangle formed by p_1, p_2, p_3, then $L_1 \geq 0$, $L_2 \geq 0$, and $L_3 \geq 0$. The corresponding texture coordinates $c(x_c, y_c)$ can be obtained through barycentric interpolation.

$$\begin{bmatrix} x_c \\ y_c \end{bmatrix} = L_1 \begin{bmatrix} x_c^1 \\ y_c^1 \end{bmatrix} + L_2 \begin{bmatrix} x_c^2 \\ y_c^2 \end{bmatrix} + L_3 \begin{bmatrix} x_c^3 \\ y_c^3 \end{bmatrix} \tag{8}$$

By traversing each triangular patch on the projector image plane and performing barycentric interpolation for the points inside, corresponding camera coordinates are calculated, establishing a continuous global mapping between the projector and the camera.

Structured Light Projection Compensation

Compensation Process Modeling. To achieve adaptive adjustment of the projection input based on lighting and object reflectance, we need to model the projection process, which is represented by Eq. 9.

$$\boldsymbol{x}^* = \mathcal{F}^{\dagger}\left(\boldsymbol{x}; \tilde{\boldsymbol{s}}\right), \tag{9}$$

where $\tilde{\boldsymbol{s}}$ is the surface image of the object under global illumination, $\boldsymbol{x}^*$ is the SL projection pattern, and $\boldsymbol{x}$ is the projection pattern captured by the camera. To ensure that the projection pattern is unaffected by ambient light or object surface texture, we set the initial pattern as $\boldsymbol{x}_0$ and the ideal captured image as $\boldsymbol{x} = \boldsymbol{x}_0$. The optimized projection pattern is obtained using:

$$\widehat{\boldsymbol{x}} = \mathcal{F}^{\dagger}\left(\boldsymbol{x}_0; \tilde{\boldsymbol{s}}\right). \tag{10}$$

Optimizing the SL pattern requires solving the inverse projection function $\mathcal{F}^{\dagger}$. To generalize this solution, the photometric compensation network (PCNet), is trained to model $\mathcal{F}^{\dagger}$ using globally matched projector-camera image pairs $(\tilde{\boldsymbol{x}}; \boldsymbol{x})$ and corresponding surface images $\tilde{\boldsymbol{s}}$. PCNet consists of a siamese encoder and a decoder, as shown in Fig. 2. The encoders share weights, and by subtracting the features learned from $\tilde{\boldsymbol{s}}$ from those learned from $\tilde{\boldsymbol{x}}$, we can separate the varying parts of ambient light, object surface texture, and reflections from the overall photometric model. This allows the network to focus on the photometric effects caused by the projection, giving the photometric compensation network generalization capabilities.

However, many pixels in this process deviate from physical laws. For example, if blue light is projected onto a surface that absorbs it, adjusting the blue channel of $\widehat{\boldsymbol{x}}$ via Eq. 10 may not reproduce the effect of $\boldsymbol{x}_0$. This can cause brightness or chromaticity clipping, resulting in image artifacts. To address this, we introduce an additional module that minimally refines the optimization objective to achieve physically plausible results.

DeArtifact Module. We develop a Weighted-Constrained Thin-Plate Spline (WC-TPS) method to regulate PCNet's input within physically valid brightness clipping bounds. WC-TPS models compensation as a 3D TPS deformation governed by Eq. 11, enforcing spatial smoothness via spline regularization and clipping constraints through boundary conditions. This dual mechanism redistributes compensation gradients across non-clipping regions while preserving subpixel optical continuity, translating TPS smoothness into photometrically constrained adjustments.

$$\boldsymbol{x}^* = f\left(\boldsymbol{x}'\right) = \sum_{i=0}^{N-1} \omega_i \phi\left(\| \boldsymbol{x}' - \boldsymbol{x}'_i \|\right) + \mathbf{a}^T \boldsymbol{x}' + b, \tag{11}$$

where $\boldsymbol{x}^*$ is the compensation result, $\boldsymbol{x}'$ is the captured image, and ω_i, $\mathbf{a}$, and b are parameters determined by solving a linear system of equations. To compute the TPS projection compensation function, $N\,(N \geq 4)$ sets of projection compensation image pairs are needed. During global matching, a real projection compensation data pair is obtained, while the remaining $N-1$ simulated pairs need to be generated using synthetic data. To ensure that the generation process aligns with the current scenario, we know from 10 that the projection process is the inverse of the compensation process. Thus, we train a projection network (ShadingNet) suitable for the current scenario using a network architecture similar to PCNet, with the input and output positions swapped during training. By using uniformly distributed solid-color images in the color space as projection inputs, we can generate the corresponding projection results for the current scenario via ShadingNet, providing reliable synthetic data for TPS. During the parameter computation process, the real data pair $(\boldsymbol{x}^*_{\boldsymbol{real}}, \boldsymbol{x}'_{\boldsymbol{real}})$ has a higher weight:

$$\mathbf{w} = \left(\alpha w_1, \alpha w_2, ..., (1-\alpha)\, w_N\right)^T \quad \text{where} \quad \alpha = 0.35. \tag{12}$$

To minimize clipping, we adopt the optimization method proposed by (Grundhöfer and Iwai 2015) to smoothly adjust the input brightness.

$$\boldsymbol{x}_{0_{adapt}} = \boldsymbol{s} * \boldsymbol{x}_0, \tag{13}$$

First, we define the error functions that will be minimized in the optimization process. The error function includes the saturation error err_{sat}, which prevents color clipping, the gradient variation error err_{grad}, which ensures smooth image adjustments, and the intensity error err_{int}, which maintains overall brightness.

Next, the variable brightness scaling value S is adjusted non-linearly for each pixel. The optimal S is found by minimizing the following objective:

$$S_{opt} = \underset{S}{\operatorname{argmin}}\, err_{opt}(S). \tag{14}$$

The total error function is:

$$\begin{aligned} err_{opt}(S) =& \omega_{\text{sat}} \cdot err_{\text{sat}}(S) + \omega_{\text{grad}} \cdot err_{\text{grad}}(S) \\ &+ \omega_{\text{int}} \cdot err_{\text{int}}(S). \end{aligned} \tag{15}$$

Here, ω_{sat}, ω_{grad}, and ω_{int} are the weights for each error term.

To mitigate potential inaccuracies in TPS-based photometric compensation, we introduce a photometric inspector that uses PCNet-generated solid-color projections spanning the RGB space. Since TPS operates pixel-wise, it may produce color discontinuities, detected via gradient magnitude. To avoid misclassifying object texture edges as artifacts, regional structural similarity is also evaluated. Pixels failing either criterion are masked out, bypassing DeArtifact and reverting to direct PCNet optimization.

$$M(p) = \begin{cases} 1 & \text{if } G_p \geq \tau_g \text{ or } \text{SSIM}(I_{A(p)}, R_{A(p)}) \leq \tau_s \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

Here, G_p represents the color gradient magnitude at pixel p, and SSIM computes the structural similarity between the local image region $A(p)$ and its reference counterpart. The binary mask $M(p)$ is determined by comparing

these features against two optimized thresholds: τ_g for gradient sensitivity and τ_s for structural dissimilarity, where τ_g controls texture edge detection while τ_s governs allowable photometric deviations.

Open Structured Light Benchmarking System

We present an open-source benchmarking system for structured light research, addressing three key limitations of conventional methods: (1) prolonged calibration of complex hardware setups, (2) scene-specific constraints in real-world data collection, and (3) lack of reliable ground-truth for 3D reconstruction.

Our Blender-based (Blender Online Community 2025) virtual calibration system emulates projector-camera setups for geometry-consistent simulation. It supports flexible radiometric configurations (e.g., BRDFs, dynamic lighting) and geometric distortion modeling (e.g., specularities, complex textures), with automated multi-parameter optimization via scripting to generate large-scale radiometric and geometric datasets.

The system defines standardized SL evaluation protocols, including programmable encoding tests, intrinsic/extrinsic calibration using virtual targets, and multi-level quantitative metrics for texture fidelity and 3D accuracy. By decoupling evaluation from physical hardware, it enables reproducible testing and cross-platform comparison. We will release the full codebase and preset scenes to support standardized, verifiable SL research.

Experiments

We adopt a two-stage training strategy to enhance the generalization of the photometric compensation network. PCNet and ShadingNet are first pre-trained on 150 synthetically generated datasets via our procedural pipeline, then fine-tuned on 40 real-world measurements under diverse illumination, combining 20 samples from (Huang and Ling 2019) and 20 captured by our multi-projector array.

Each dataset contains 700 projection images per object surface, with 500 for training and 200 for testing. During training, object surface images and their corresponding projections are randomly paired to disentangle global illumination from surface reflections. The calibration pattern's background color is set to $\mathbf{v} = (0.25, 0.25, 0.25)$. For Eq. 16, we use $\tau_g = 20$ and $\tau_s = 0.8$.

We validate our global matching algorithm against Gray code and deep learning baselines, focusing on continuous projector-camera mapping. Then, we evaluate optimization performance across multiple SL patterns, establishing the first unified optimization framework across SL types. Decoding accuracy is used as the evaluation metric. All experiments are conducted using a Hikvision MV-CU013-A0GC camera (1280×1024) and a Sony VPL-EX570 projector (1024×768).

Comparison of Global Matching Methods

We compare our global matching method (GMM) with traditional Gray code (GC) and deep learning-based WarpingNet (Huang, Sun, and Ling 2021). GC requires 42 patterns to encode each projector pixel, but it produces only discrete correspondences and fails to establish continuous projector-camera mappings. This results in decoding artifacts, especially at codeword transitions and in regions with resolution mismatch.

Our method addresses these limitations using a single projection, applying De Bruijn-based sparse matching, error-code filtering, and Delaunay-based interpolation to produce smooth, continuous mappings. As shown in Fig. 3, our results demonstrate better texture continuity and fewer artifacts.

We also compare with WarpingNet, which is trained on 500 projected images. While effective for single-object scenes, its performance drops significantly on multi-object or textured surfaces. Fig. 3 shows that WarpingNet suffers from interpolation blur and cross-object mismatch, whereas GMM preserves finer details.

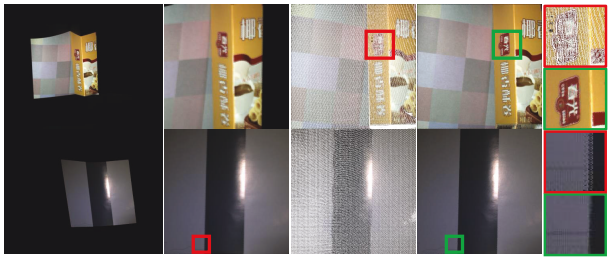

Figure 3: Comparison with the method of global encoding using Gray code (GC) and WarpingNet.

To ensure a fair comparison, we apply linear interpolation and our GMM to the 42-projection GC outputs for continuous mapping. Table **??** presents the results: our GMM reduces error rate by 60.31% compared to GC, and achieves only 3.67% higher error than GC+linear. The addition of our EC filter further improves accuracy by 14.01%.

Method	Continuous Surface		Discontinuous Surface	
	SSIM↑	Err. (%)↓	SSIM↑	Err. (%)↓
WarpingNet	0.813	4.61	N/A	N/A
GC	0.361	5.42	0.362	5.63
GMM	0.912	2.15	0.889	3.38
GC+Linear	0.935	2.09	0.927	3.35
GC+GMM	**0.957**	**2.07**	**0.952**	**3.31**

Table 1: Our method (GMM) achieves reliable matching results using only a single image.

Optimization Effects of Different Structured Light Patterns

We evaluate our method on three representative SL patterns—binary stripes (e.g., Gray code), speckle, and color-coded—under diverse lighting, texture, and reflectance conditions (see Fig. 4). Captured images are mapped into projector space to enable pixel-wise decoding error analysis before and after optimization. Additional results, including variations in exposure and pattern parameters, are provided in the supplementary material.

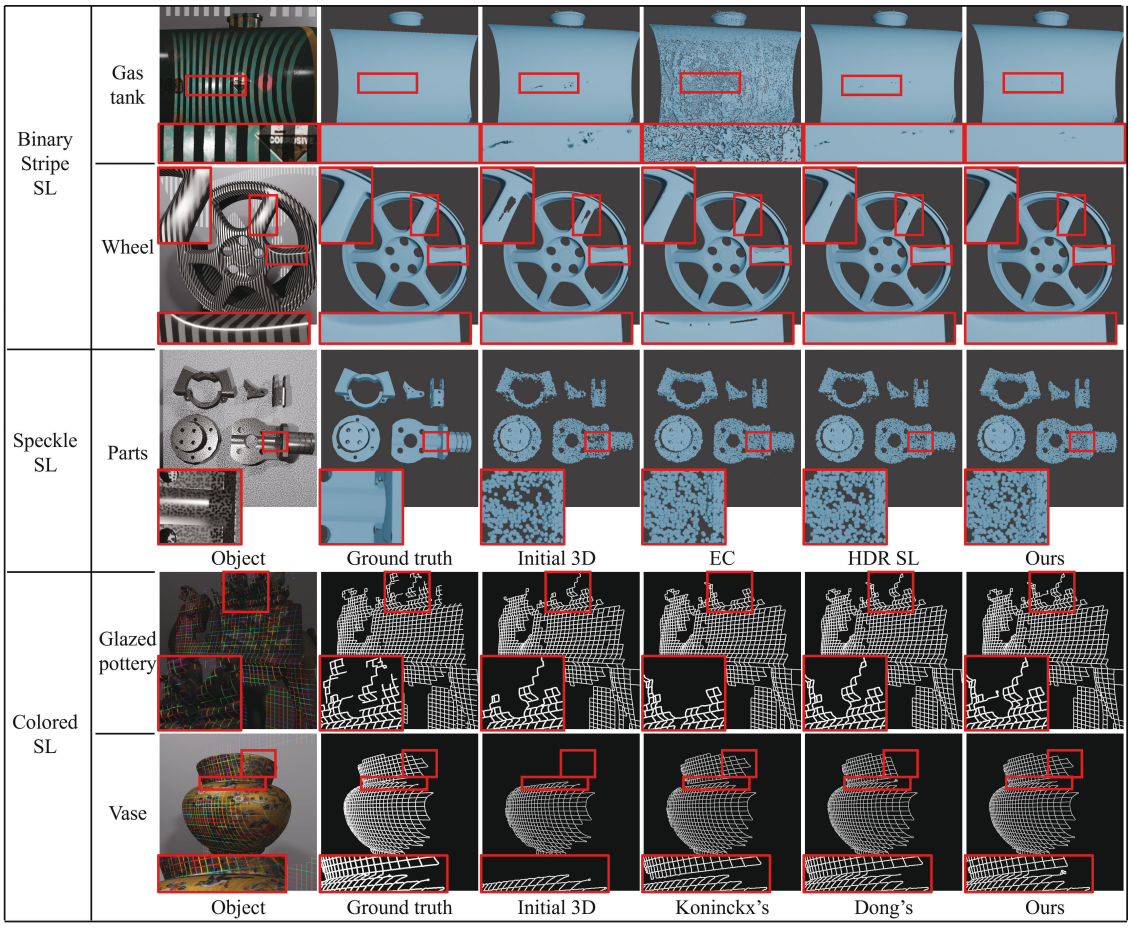

Figure 4: Visual comparison of optimized structured light patterns under different conditions: darkroom (rows 1 & 4), multi-lighting (rows 2, 3, 5), Lambertian (rows 2 & 3), non-Lambertian (rows 1, 4, 5), textured (rows 1, 4, 5), and textureless surfaces (rows 2 & 3). Our optimized patterns yield more complete and accurate geometry across all scenarios.

Binary Stripe Structured Light Pattern Gray code is widely used in binary encoding. We compare our method with HDR SL (Fu et al. 2023), which adjusts light intensity to mitigate overexposure, and Exposure Correction (EC) (Zou et al. 2024), which post-processes captured images.

To ensure fairness, all methods use the same projection intensity and decoding pipeline. HDR SL is grayscale-only and lacks spatial adaptability; EC often overcompensates for dark textures, mistaking them for underexposure. In contrast, our method adapts to both surface texture and illumination, achieving a 3.05% average absolute error reduction and a 6.34% relative improvement over HDR SL. It also outperforms EC by at least 5.98% across all eight scenarios. Fig. 4 (rows 1–2) shows more complete depth reconstruction.

Speckle Structured Light Pattern Speckle SL is widely used in stereo-based matching but is sensitive to reflectance and exposure variations. We compare our method with HDR SL and EC under consistent speckle generation and decoding parameters, isolating the impact of photometric optimization.

As shown in Fig. 4, EC struggles with highlight and low-contrast regions, while our method reduces artifacts effectively. It yields an 8.70% average absolute error reduction, with 13.84% and 54.41% relative improvements over HDR SL and EC, respectively. We further evaluate performance under varying exposure times and speckle densities/sizes (see the supplementary material). Our method consistently delivers higher decoding accuracy.

Colored Structured Light Pattern Color-coded SL allows high-density single-shot reconstruction. We optimize De Bruijn stripe patterns and compare with Koninckx et al. (Koninckx et al. 2005), which calibrates projector-camera response curves, and Dong et al. (Dong, Ling, and Huang 2023), which performs unsupervised color detection.

For fairness, we pre-calibrate response curves for Koninckx's method and use the same De Bruijn patterns and resolution across methods. Koninckx achieves only localized correction in overexposed regions ($\leq$12.7%), with a 5.55% average error reduction. Dong's method achieves 14.16% via uniform color segmentation but sacrifices spatial discriminability. Our method integrates physical modeling with

adaptive chromatic mapping, achieving a 29.39% average error reduction and artifact-free decoding (see Fig. 4).

Ablation Study

To systematically validate our proposed technical components, we conduct ablation studies focusing on two critical modules: (1) the global matching module with error code filtering, and (2) the DeArtifact module for SL projection compensation. Both components address fundamental challenges in structured light reconstruction through distinct technical approaches.

Global Matching with Error Code Filtering: In the projector-camera global matching module, we employ error code filtering to reduce feature point matching errors during global matching. To demonstrate the effectiveness of this module, we conducted an ablation experiment by removing the error code filter and performing triangulation and global mapping on the original matching pairs. The comparative results are shown in Fig. 5.

When there is continuous loss in the surrounding De Bruijn encoding, the orange-red dots may experience decoding errors, leading to incorrect placement. This can result in erroneous triangle color mapping, as shown in the figure. After filtering out these errors, the region is correctly matched through re-triangulation and color mapping within the triangles. Quantitative evaluation of the EC filter's effectiveness, presented in Table **??**.

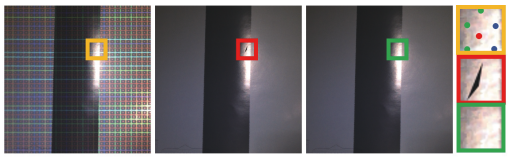

Figure 5: Global matching results with and without the error code (EC) filter. The fourth column is an enlarged view of the erroneous matches, with red dots indicating incorrect matches, green dots indicating correct matches, and blue dots indicating unmatched points.

Model	SSIM↑	PSNR↑	RMSE↓	Avg. Err.↓	Max Err.↓
w/o EC filter	0.871	21.75	13.93	3.53	8.19
w/ EC filter	**0.889**	**22.39**	**13.38**	**3.38**	**5.18**

Table 2: The EC filter significantly reduces the maximum decoding error and improves optimization stability

DeArtifact Projection Optimization: We propose the DeArtifact module to reduce artifacts that traditional projection compensation networks cannot avoid. A pre-trained ShadingNet serves as the projection network to provide WC-TPS with foundational training data. We further incorporate a real SL image pair into TPS training, adjusting data weights to improve photometric realism.

To evaluate the contributions of synthetic and real data, we conduct ablation experiments comparing the accuracy of resulting TPS photometric models. Without ShadingNet-generated data, we simulate projections using a multiplication blend (MB) mode processed by PCNet. As shown in Fig. 6, this leads to significant color deviations or data loss. Adding real SL data improves results but still shows deviations. ShadingNet enables closer alignment with actual projection behavior, and combining it with real data allows fine-tuning toward the ground truth.

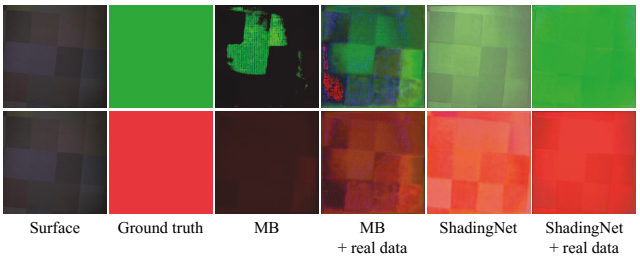

Figure 6: Contribution of different modules to photometric representation. ShadingNet offers a robust initial estimate, while real calibration data refines color-related biases.

Our experiments are configured using the matched ground truth as input, thereby avoiding the influence introduced by Stage I. Table 3 compares artifact reduction effectiveness via SSIM, PSNR, and RMSE. In DeArtifact, we first verify the TPS photometric mapping, then exclude pixels with unreliable information from optimization. Our experiments show three improvements: (1) smoother gradients in color transitions, (2) better spectral fidelity via optimized color coding, and (3) removal of grid discontinuities.

WC-TPS	ShadingNet	Rea	SSIM↑	PSNR↑	RMSE↓	Err. (%)↓
			0.760	19.99	24.62	6.31
✓			0.763	20.01	24.33	3.75
✓	✓		0.823	21.77	22.36	1.22
✓	✓	✓	**0.849**	**23.24**	**18.76**	**0.98**

Table 3: Contribution of different modules in Stage II to the photometric representation, while also demonstrating their optimization capability on the final decoding performance within binary coding at the highest encoding density.

Conclusion

This work addresses two key challenges in structured light: scene-specific calibration inefficiency and pattern-dependent optimization. We propose GUSLO, a general and unified framework that performs single-shot geometric calibration and artifact-aware photometric adaptation via physics-guided transfer. By integrating geometric interpolation with adaptive color mapping, GUSLO enables robust optimization across binary, speckle, and color-coded SL patterns without multi-pattern calibration. Experiments show consistent improvements under challenging lighting and surface textures, demonstrating its practicality for industrial inspection and cultural heritage digitization. Its compatibility with diverse encoding schemes offers a scalable foundation for future multi-device reconstruction systems.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (Grant No. GG-2024-01-02), the National Natural Science Foundation of China (Grant No. 62277011), and the Project of Chongqing Municipal Economy and Information Technology Commission (Grant No. YJX-2025001001009). Additional support was provided by the Open Research Fund of the Guangdong Laboratory of Artificial Intelligence and Digital Economy (Shenzhen) under Grant No. GML-KF-24-18. The authors gratefully acknowledge Beijing DZMatrix Technology Co., Ltd. for their valuable collaboration and technical support. We also extend our sincere thanks to Dr. Xuancheng Zhang and Donglin Di for their insightful guidance and invaluable assistance in the experimental work.

References

Bak, D. 2003. Rapid prototyping or rapid production? 3D printing processes move industry towards the latter. *Assembly Automation*.

Blender Online Community. 2025. Blender – The Free and Open Source 3D Creation Suite. https://www.blender.org. Accessed: 2025-04-05.

Ding, Q.-K.; and Liang, H.-E. 2024. Digital restoration and reconstruction of heritage clothing: a review. *Heritage Science*, 12(1): 225.

Dodda, V. C.; Kuruguntla, L.; Elumalai, K.; Chinnadurai, S.; Sheridan, J. T.; and Muniraj, I. 2023. A denoising framework for 3D and 2D imaging techniques based on photon detection statistics. *Scientific Reports*, 13(1): 1365.

Dong, X.; Ling, H.; and Huang, B. 2023. Adaptive Color Structured Light for Calibration and Shape Reconstruction. In *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 1240–1249. IEEE.

Fu, Y.; Fan, J.; Jing, F.; and Tan, M. 2023. High dynamic range structured light 3-D measurement based on region adaptive fringe brightness. *IEEE Transactions on Industrial Electronics*.

Fu, Y.; Huang, Y.; Xiao, W.; Li, F.; Li, Y.; and Zuo, P. 2024. Deep learning-based binocular composite color fringe projection profilometry for fast 3D measurements. *Optics and Lasers in Engineering*, 172: 107866.

Fujii, K.; Grossberg, M. D.; and Nayar, S. K. 2005. A projector-camera system with real-time photometric adaptation for dynamic environments. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, 814–821. IEEE.

Grundhöfer, A.; and Iwai, D. 2015. Robust, error-tolerant photometric projector compensation. *IEEE Transactions on Image Processing*, 24(12): 5086–5099.

Gupta, P. 2007. *Gray code composite pattern structured light illumination*. Ph.D. thesis, University of Kentucky Libraries.

Holl, K.; Pallas, L.; and Bellendorf, P. 2021. Structured light scanning as a monitoring method to investigate dimen-sional changes due to climatic changes on cultural heritage. In *Proceedings of the International Conference on Cultural Heritage and New Technologies*, volume 26.

Huang, B.; and Ling, H. 2019. End-to-end projector photometric compensation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 6810–6819.

Huang, B.; Ozdemir, S.; Tang, Y.; Liao, C.; and Ling, H. 2018. A single-shot-per-pose camera-projector calibration system for imperfect planar targets. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, 15–20. IEEE.

Huang, B.; Sun, T.; and Ling, H. 2021. End-to-end full projector compensation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(6): 2953–2967.

Jia, T.; Li, X.; Yang, X.; Lin, S.; Liu, Y.; and Chen, D. 2024. AdaptiveStereo: Depth estimation from adaptive structured light. *Optics & Laser Technology*, 169: 110076.

Kawasaki, H.; Furukawa, R.; Sagawa, R.; and Yagi, Y. 2008. Dynamic scene shape reconstruction using a single structured light pattern. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1–8. IEEE.

Koninckx, T. P.; Peers, P.; Dutré, P.; and Van Gool, L. 2005. Scene-adapted structured light. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, 611–618. IEEE.

Li, Y.; Qian, J.; Feng, S.; Chen, Q.; and Zuo, C. 2022. Deep-learning-enabled dual-frequency composite fringe projection profilometry for single-shot absolute 3D shape measurement. *Opto-Electronic Advances*, 5(5): 210021–1.

Li, Y.; Yin, W.; Li, J.; and Xie, X. 2023. Physics-based efficient full projector compensation using only natural images. *IEEE Transactions on Visualization and Computer Graphics*.

Montusiewicz, J.; Miłosz, M.; Kesik, J.; and Żyła, K. 2021. Structured-light 3D scanning of exhibited historical clothing—a first-ever methodical trial and its results. *Heritage Science*, 9(1): 74.

Nguyen, A.-H.; Ly, K. L.; Qiong Li, C.; and Wang, Z. 2022. Single-shot 3D shape acquisition using a learning-based structured-light technique. *Applied Optics*, 61(29): 8589–8599.

Pages, J.; Salvi, J.; Garcia, R.; and Matabosch, C. 2003. Overview of coded light projection techniques for automatic 3D profiling. In *IEEE International Conference on Robotics and Automation*, volume 1, 133–138 vol.1.

Park, J.; Jung, D.; and Moon, B. 2022. Projector compensation framework using differentiable rendering. *IEEE Access*, 10: 44461–44470.

Qian, J.; Feng, S.; Tao, T.; Hu, Y.; Liu, K.; Wu, S.; Chen, Q.; and Zuo, C. 2019. High-resolution real-time 360 3d model reconstruction of a handheld object with fringe projection profilometry. *Optics Letters*, 44(23): 5751–5754.

Qian, J.; Feng, S.; Xu, M.; Tao, T.; Shang, Y.; Chen, Q.; and Zuo, C. 2021. High-resolution real-time 360° 3D surface defect inspection with fringe projection profilometry. *Optics and Lasers in Engineering*, 137: 106382.

Raskar, R.; van Baar, J.; and Chai, J. X. 2002. A low-cost projector mosaic with fast registration. In *Asian Conference on Computer Vision (ACCV)*, volume 3.

Sansoni, G.; Trebeschi, M.; and Docchio, F. 2009. State-of-the-art and applications of 3D imaging sensors in industry, cultural heritage, medicine, and criminal investigation. *Sensors*, 9(1): 568–601.

Shih, K.-T.; Liu, J.-S.; Shyu, F.; and Chen, H. H. 2020. Enhancement and speedup of photometric compensation for projectors by reducing inter-pixel coupling and calibration patterns. *IEEE Transactions on Image Processing*, 30: 418–430.

Stelzner, J.; Stelzner, I.; Martinez-Garcia, J.; Gwerder, D.; Wittköpper, M.; Muskalla, W.; Cramer, A.; Heinz, G.; Egg, M.; and Schuetz, P. 2022. Stabilisation of waterlogged archaeological wood: the application of structured-light 3D scanning and micro computed tomography for analysing dimensional changes. *Heritage Science*, 10(1): 60.

Sugimoto, M.; Iwai, D.; Ishida, K.; Punpongsanon, P.; and Sato, K. 2021. Directionally decomposing structured light for projector calibration. *IEEE Transactions on Visualization and Computer Graphics*, 27(11): 4161–4170.

Wang, K.; Song, L.; Wang, C.; Ren, Z.; Zhao, G.; Dou, J.; Di, J.; Barbastathis, G.; Zhou, R.; Zhao, J.; et al. 2024. On the use of deep learning for phase recovery. *Light: Science & Applications*, 13(1): 4.

Xu, X.; Lin, Y.; Zhou, H.; Zeng, C.; Yu, Y.; Zhou, K.; and Wu, H. 2023. A unified spatial-angular structured light for single-view acquisition of shape and reflectance. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 206–215.

Yamada, Y.; Takatori, S.; Adachi, M.; Hasegawa, K.; Li, L.; Pan, J.; Thufail, F. I.; Yamaguchi, H.; and Tanaka, S. 2024. High-Visibility Edge-Highlighting Visualization of 3D Scanned Point Clouds Based on Dual 3D Edge Extraction. *Remote Sensing*, 16(19): 3750.

Zou, Z.; Yu, W.; Huang, J.; and Zhao, F. 2024. Semantic Pre-supplement for Exposure Correction. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 5961–5970.