A Single-shot-per-pose Camera-Projector Calibration System For Imperfect Planar Targets

Bingyao Huang *
Temple Univeristy

Samed Ozdemir †
Rowan Univeristy

Ying Tang ‡
Rowan Univeristy

Chunyuan Liao §
HiScene Info. Technologies

Haibin Ling ¶
Temple Univeristy

ABSTRACT

Existing camera-projector calibration methods typically warp feature points from a camera image to a projector image using estimated homographies, and often suffer from errors in camera parameters and noise due to imperfect planarity of the calibration target. In this paper we propose a simple yet robust solution that explicitly deals with these challenges. Following the structured light (SL) camera-project calibration framework, a carefully designed correspondence algorithm is built on top of the De Bruijn patterns. Such correspondence is then used for initial camera-projector calibration. Then, to gain more robustness against noises, especially those from an imperfect planar calibration board, a bundle adjustment algorithm is developed to jointly optimize the estimated camera and projector models. Aside from the robustness, our solution requires only one shot of SL pattern for each calibration board pose, which is much more convenient than multi-shot solutions in practice. Data validations are conducted on both synthetic and real datasets, and our method shows clear advantages over existing methods in all experiments.

Index Terms: Computing methodologies—Camera calibration; Computing methodologies—3D imaging; Computing methodologies—Reconstruction

1 INTRODUCTION

Camera-projector systems are popular in 3D surface reconstruction and projected augmented reality, where in most cases, structured light (SL) is applied due to its ease of use and accuracy. Compared with passive feature point based 3D reconstruction methods, such as stereo vision and structure from motion (SfM), SL is able to reconstruct a denser and more precise surface. Moreover, SL works for texture-less or repetitively textured objects.

A typical SL system consists of a calibrated camera and projector pair placed at a fixed distance and orientation as shown in Fig. 1. Firstly, the projector projects known encoded patterns onto the target object, then the projected patterns are deformed according to the surface shape of the target object. Once the camera captures the deformed patterns, pixel correspondences between camera and projector can be established by matching the captured and projected patterns. In the end, the 3D coordinates of the deformed pattern pixels are triangulated, given the camera-projector parameters and pixel correspondences.

Despite the simplicity, the 3D reconstruction precision of an SL system is highly dependent on the joint camera-projector pair calibration. Unlike a binocular stereo vision system, in an SL system, the projector is unable to capture images. So most SL calibration

---

*e-mail: bingyao.huang@temple.edu

†e-mail: ozdemi63@students.rowan.edu

‡e-mail: tang@rowan.edu

§e-mail: liaocy@hiscene.com
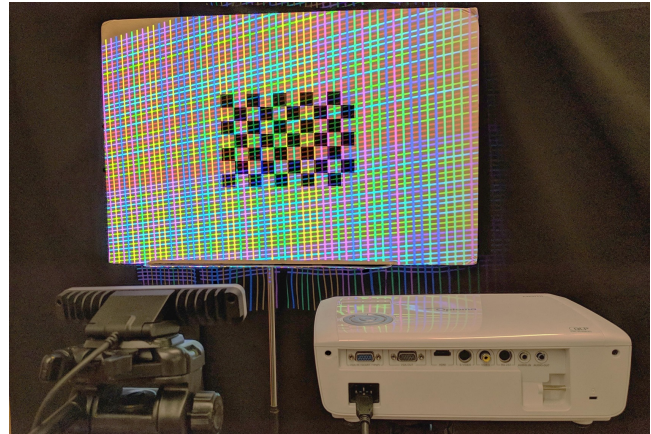
¶e-mail: hbling@temple.edu

Figure 1: System setup: a projector on the bottom-right, a camera on the bottom-left and calibration board with a checkerboard pattern attached to it.

systems model the projector as an inverse camera that can "see" the calibration target [4, 20, 37]. Then it can be simply calibrated like a camera using Zhang's method [39].

The above camera-projector framework requires correspondences between the projector image plane and a reference plane, which is usually approximated by a planar calibration board with a printed checkerboard, dots or circles pattern. The camera can be calibrated using these patterns. Afterwards, the projector projects encoded SL patterns onto the calibration board, and these patterns are then captured by the camera for calibrating the projector.

In practice, the calibration board and printed pattern are hardly perfect planar due to manufacturing and/or glue. As pointed out by [1, 13, 29, 30, 39], an imperfect calibration target may significantly impact the accuracy of Zhang's method. As most existing calibration methods rely on camera parameters to warp printed feature points to the projector image space, the camera error may be propagated to the projector image plane. This adds to the error of projector calibration that is again done by Zhang's method.

To address this issue, we propose a novel additional step to jointly rectify the camera and projector models. Specifically, after calibrating the camera and projector using the traditional method, we put them into a bundle adjustment (BA) framework [10] for rectification, together with a scale regularization for further improvement. Another key component in our system is the reliable correspondence construction process. By using an efficient De Bruijn pattern [12, 36] and a carefully designed keypoint extraction algorithm, our system provides reliable keypoint correspondence for the calibration algorithms. Moreover, being a single-shot per pose[1] solution, our system brings practical convenience over systems that require multiple shots of SL patterns for a single calibration board pose. This is particularly important for applications that require frequent re-calibrations, e.g., with the camera/projector moving around.

---

[1]Following [11], we call it *single-shot* for conciseness in the rest of the paper.

To summarize, our system brings the following contributions:

- Our system explicitly deals with the noise in target planarity with a novel BA solution. This is the first such system for joint camera-projector calibration, to the best of our knowledge.

- Unlike many existing methods, to calibrate the system, we apply points from an SL pattern rather than from a checkerboard. This strategy boosts both the number of feature points and their spatial distribution, and hence improves calibration robustness.

- The proposed method performs camera-projector pair calibration with only a single-shot per pose, making it practically convenient in many applications. It can provide a flexible and accurate results even when the board is handheld.

The effectiveness of the proposed solution over existing ones is clearly demonstrated in our experiments on both synthetic and real data, especially when the calibration board is imperfect planar. In addition, the source code is publicly available at `https://github.com/BingyaoHuang/single-shot-pro-cam-calib`.

In the rest of the paper, we summarize related work in Sec. 2 and introduce our camera-projector calibration method in Sec. 3. Then, we report experiments in Sec. 4, and conclude this paper in Sec. 5.

2 RELATED WORK

Most existing camera-projector pair calibration methods apply Zhang's method [39], where the 3D-2D correspondences between the points on the calibration board and the projector image are computed by some transformations. Regardless of a multi-shot or single-shot method, their transformations fall into one of the following methods: global homography [2, 7–9, 14, 16, 21, 22, 35], local homography [19, 20], direct pixel-to-pixel transformation [37] and incremental projector image pre-warp [3, 6, 28, 34, 38].

It is worth noting that a global homography-based method usually ignores both projector lens distortions and imperfect planarity of the calibration board. While the other three types of calibration methods can model projector lens distortion [6, 19, 20, 37, 38], they are highly dependent on camera calibration accuracy. In addition, the imperfect planarity of calibration board is ignored in all the reviewed methods above, and such imperfectness can cause errors as pointed out in [1, 13, 29, 30, 39] and illustrated in Sec. 4.

Other than using Zhang's method, self-calibration algorithms [18, 32, 33, 35] are capable of calibrating intrinsics and extrinsics of the camera-projector pair without a known planar target, instead a fundamental matrix or its variant is estimated using camera and projector image correspondences. With some priori of the intrinsics, e.g., unit aspect ratio and the principle point is assumed to be at the center of the image. However, these two assumptions are often violated, especially for projectors [20].

Additional cameras can also be included to either reduce calibration board poses [5] or allow arbitrary shapes as calibration targets [24]. However, these methods are even inflexible since they either require additional hardware or precise 3D measurement of a complex object, let alone multiple shots.

Multi-shot and single-shot methods According to [11], SL-based camera-projector calibration methods can be categorized into two types: multi-shot [5, 6, 20, 23, 32–34, 37] and single-shot [2, 4, 17, 35, 36]. Specifically, multi-shot and single-shot indicate the number of SL pattern shots for *each* calibration board pose, rather than the total number of shots for the whole process.

Multi-shot methods project a sequence of patterns onto the calibration board, the patterns are encoded in Gray/binary code [20, 23, 32] or multiple phase shifting [37], leading to a pixel-wise or even sub-pixel resolution. However, a disadvantage is that it is slow and computationally expensive due to multiple shots, e.g., [20] requires about 20 shots and captures for each pose. Incremental methods [3, 6, 19, 28, 34, 38] also belong to multi-shot, since the
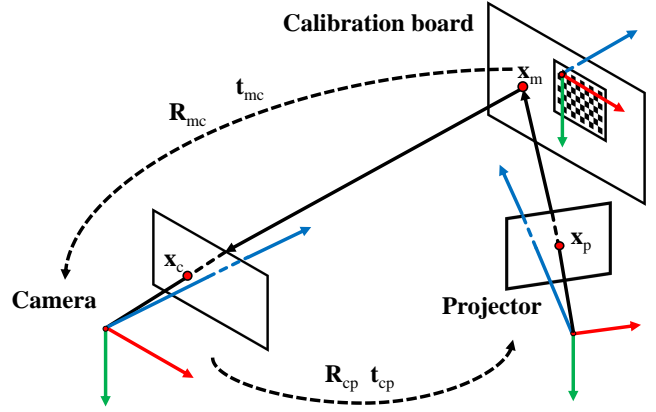

Figure 2: Coordinate system. The world origin is at the camera optical center. Red, green and blue axes represent X, Y and Z directions, respectively.

projected pattern is incrementally adjusted to fit the printed pattern until a perfect superimposition is achieved, which requires at least two shots per pose.

Despite the correspondence accuracy, multi-shot calibration methods are both computationally expensive and memory inefficient compared with single-shot ones. Moreover, multi-shot methods are sensitive to motion; even a little shift or jitter between two consecutive captures can produce huge SL decoding errors due to pattern misalignment. For example, when a user holds the calibration board or a mobile camera-projector pair, it is very hard to ensure absolute stillness of the target between consecutive shots.

Single-shot methods only require one shot per pose and adopt spatial multiplexed patterns including the De Bruijn sequence [12, 15, 25, 27, 36], M-array [4, 35], checkerboard [2] and phase shifting fringes. The feature point correspondences are uniquely encoded in a single SL pattern. Consequently, single-shot SL allows for faster and more flexible camera-projector calibration than multi-shot SL.

Compared with previous studies, our camera-projector calibration is simple and fast, and requires only one shot per calibration board pose. Moreover, it refines imperfectly calibrated camera and projector parameters due to imperfect planar calibration board using a bundle adjustment method. The experiments show that our method outperforms the other counterparts on both synthetic and real data.

3 METHOD

Notations. Throughout the paper, we use the mathematical notations as shown in Tab. 1. In addition we use subscripts c, p and m for camera, projector and calibration board model space, respectively. Thus, the subscript cp (or mp) stands for the transformation from camera (or calibration board) coordinate system to projector coordinate system (Fig. 2).

Table 1: Notations

Notation	Example	Meaning
italic	$a, A, \ldots$	scalars
lower-case boldface	$\mathbf{a}, \mathbf{b}, \ldots$	vectors
boldface capital	$\mathbf{A}, \mathbf{B}, \ldots$	matrices
calligraphic	$\mathscr{A}, \mathscr{B}, \ldots$	sets
index range	$\mathbf{a}^{1:N}$	$\mathbf{a}^1, \mathbf{a}^2, \ldots, \mathbf{a}^N$
dot	$\dot{\mathbf{a}}, \dot{\mathbf{A}}, \ldots$	initial guess
hat	$\hat{\mathbf{a}}, \hat{\mathbf{A}}, \ldots$	estimation
bar	$\bar{\mathbf{a}}, \bar{\mathbf{A}}, \ldots$	homogeneous coordinates

3.1 System Overview

Our camera-projector calibration system (Fig. 1) consists of an RGB camera, a projector and a white board with a printed checkerboard
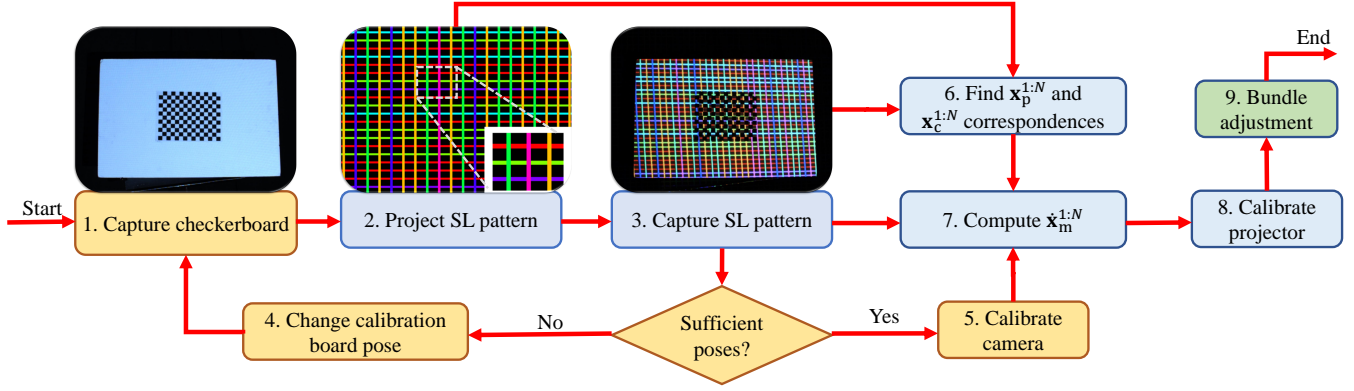
Figure 3: System flowchart. We divide the calibration algorithm into three procedures (Alg. 1): **Yellow blocks**: camera initial calibration. **Blue blocks**: projector initial calibration and **Green block**: bundle adjustment. Best viewed in color.

pattern attached to it. As summarized in Alg. 1, it contains three stages: (1) initial camera calibration using checkerboard images, (2) initial projector calibration using projected SL patterns, and (3) joint refinement of camera and projector parameters using bundle adjustment (BA).

As shown in the system flowchart in Fig. 3, we start by capturing an image of the calibration board, then we project a color-encoded SL pattern to the calibration board and take an image of the superimposed SL pattern. We change the pose of the calibration board manually and repeat the steps above to get sufficient (at least three) pose samples. Then, we first calibrate the camera using Zhang's method to get the initial camera model, including camera intrinsics and rotations and translations of each calibration board pose relative to the camera.

We then undistort the captured SL images. After that, we decode the SL patterns in the camera image plane and find their correspondences to the original SL pattern in the projector image plane. Following that, we transform the SL points to the calibration board model space using rotations and translations obtained in last step. Note the SL points in the calibration board model space may also be erroneous due to inaccurate camera calibration. With these correspondences, we apply Zhang's method again to calibrate the projector. The relative rotation and translation between camera and projector are estimated using stereo calibration. Similar to camera parameters, the projector parameters obtained are also initial guesses and subject to propagated errors from camera calibration and imperfect planarity of the calibration board.

Finally, we gather camera and projector parameters along with the SL points to perform a BA refinement. This last step largely reduces errors in initial calibration (Subsec. 3.4).

3.2 Structured Light Pattern

To allow for single-shot calibration, we employ a spatial multiplexed SL technique and use only a single color-encoded pattern (Fig. 3 steps 2-3). The SL pattern is a variant of [26] composed of vertical and horizontal colored stripes with a De Bruijn sequence encoding. A De Bruijn sequence of order n over an alphabet of k color symbols is a cyclic sequence of length k^n with a so-called window property that each subsequence of length n appears exactly once [31].

Let $\mathscr{C} = \{1, 2, ..., 8\}$ be the set of encoding color primitives, each number represents a different color. In particular, red (1), lime (3), cyan (5) and purple (7) are used for the horizontal stripes, while yellow (2), green (4), blue (6) and magenta (8) for vertical ones. In the inset of Fig. 3 step 2, the vertical color stripes are $(4, 8, 2)$ from left to right, the horizontal stripes are $(1, 3, 7)$ from top to bottom.

We employ De Bruijn encoding to both vertical and horizontal stripes, and construct a color grid with $m \times m$ intersections, where $m = k^n + 2$, in our case $k = 4, n = 3$. More importantly, a unique k-color horizontal sequence overlain atop a unique k-color vertical sequence only occurs once in the grid. As shown in inset of Fig. 3

step 2, this 3×3 subset color grid appears only once in the whole pattern. We represent the color-coded pattern using an undirected graph $\mathscr{G} = (\mathscr{V}, \mathscr{E})$, in which $\mathscr{V} = \{\mathbf{v}_1, \mathbf{v}_2, \dots \mathbf{v}_{m \times m}\}$ is a set of graph nodes, which represent color stripes intersections, where

$$\mathbf{v}_i = \{\mathbf{x}_c(i), \mathbf{x}_p(i), \mathbf{x}_m(i)\}, \quad (1)$$

such that $\mathbf{x}_c(i) = [u_c(i), v_c(i)]^T$, $\mathbf{x}_p(i) = [u_p(i), v_p(i)]^T$, and $\mathbf{x}_m(i) = [x_m(i), y_m(i), z_m(i)]^T$ represent the coordinates of the ith node in, respectively, the camera image space, the projector image space and the calibration board model space. $\mathscr{E} = \{\mathbf{e}_{00}, \mathbf{e}_{01}, \dots, \mathbf{e}_{ij}, \dots, \mathbf{e}_{m^2 \times m^2}\}$ is the set of all edges representing color stripe segments. We have $\mathbf{e}_{ij} = \{\mathscr{L}, \tau\}$, where $\mathscr{L}$ is a list of pixels belonging to this edge and $\tau \in \mathscr{C} \cup \{0\}$ is the color label of the edge (if the link exists) or 0 (otherwise).

The correspondences between the camera captured image and the projected SL pattern is built by decoding the color codeword of the SL pattern. Since finding SL correspondences is not the main focus of this paper, we provide the details in the supplementary material.

Once we have the camera and projector coordinates of all the nodes, we apply the homography $\mathbf{H}^j$ to transform node points from the camera image plane to the calibration board model space, where $\mathbf{H}^j$ is the transformation for the j-th pose estimated by initial calibration (Subsec. 3.3).

3.3 Initial Calibration

The camera and projector view spaces and calibration board model space follow a right hand coordinate system as shown in Fig. 2. The world origin is at the camera optical center.

Camera and Projector Model We employ the pin-hole model for both camera and projector calibration, with intrinsic matrices denoted by $\mathbf{K}_c$ and $\mathbf{K}_p$, respectively:

$$\mathbf{K}_c = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{K}_p = \begin{bmatrix} f'_x & 0 & c'_x \\ 0 & f'_y & c'_y \\ 0 & 0 & 1 \end{bmatrix}, \quad (4)$$

where f_x, f'_x and f_y, f'_y represent camera and projector focal lengths in x and y directions. (c_x, c_y) and (c'_x, c'_y) represent camera and projector image principle point coordinates. The camera and projector distortion coefficients are given by:

$$\mathbf{d}_c = [k_1, k_2, p_1, p_2], \quad \mathbf{d}_p = [k'_1, k'_2, p'_1, p'_2], \quad (5)$$

where k_1, k'_1 and k_2, k'_2 are radial distortion factors; p_1, p'_1 and p_2, p'_2 are tangential distortion factors. In addition, we model extrinsics parameters, i.e., relative rotation and translation of the camera with respect to the projector as:

$$\mathbf{r}_{cp} = (r_x, r_y, r_z)^T, \quad \mathbf{t}_{cp} = (t_x, t_y, t_z)^T. \quad (6)$$

Note that $\mathbf{r}_{cp} \in \mathfrak{so}(3)$ is a rotation vector, i.e., the associated Lie algebra of rotation matrix $\mathbf{R}_{cp} \in \mathrm{SO}(3)$.

Camera Calibration We first calibrate the camera using Zhang's method [39], with all the checkerboard corner correspondences from

Algorithm 1 The proposed calibration algorithm

1: Input: camera captured images $\mathscr{I}^{1:N}$
2: Output: camera-projector pair parameters $\hat{\Psi}$
3: // *Stage 1. Initial Camera Calibration*
4: $\mathbf{K}_c, \mathbf{d}_c, \mathbf{R}_{mc}^{1:N}, \mathbf{t}_{mc}^{1:N} \leftarrow \text{ZhangCalib}(\mathscr{I}^{1:N})$
5: **for** $j \leftarrow 1$ to N **do**
6: $\qquad \mathbf{H}^j = \mathbf{K}_c * [\mathbf{r1}_{mc}^j, \mathbf{r2}_{mc}^j, \mathbf{t}_{mc}^j]$ (2)
7: **end for**
8: // *Stage 2. Initial Projector Calibration*
9: **for** $j \leftarrow 1$ to N **do**
10: $\qquad \bar{\mathbf{x}}_c^j \leftarrow \text{undistort}(\mathbf{x}_c^j, \mathbf{d}_c)$
11: $\qquad$ **for** $i \leftarrow 1$ to M^j **do**
12: $\qquad\qquad \dot{\mathbf{x}}_m^j(i) = \text{inv}(\mathbf{H}^j) * \bar{\mathbf{x}}_c^j(i)$ (3)
13: $\qquad$ **end for**
14: **end for**
15: $\mathbf{K}_p, \mathbf{d}_p, \mathbf{R}_{mp}^{1:N}, \mathbf{t}_{mp}^{1:N} \leftarrow \text{ZhangCalib}(\dot{\mathbf{x}}_m^{1:N}, \bar{\mathbf{x}}_p^{1:N})$
16: $\mathbf{R}_{cp} = \text{median}(\mathbf{R}_{mp}^{1:N} * (\mathbf{R}_{mc}^{1:N})^T)$
17: $\mathbf{t}_{cp} = \text{median}(\mathbf{t}_{mp}^{1:N} - \mathbf{R}_{mp}^{1:N} * (\mathbf{R}_{mc}^{1:N})^T * \mathbf{t}_{mc}^{1:N})$
18: // *Stage 3. Bundle Adjustment*
19: $\Psi = [\mathbf{K}_c, \mathbf{d}_c, \mathbf{R}_{mc}^{1:N}, \mathbf{t}_{mc}^{1:N}, \mathbf{K}_p, \mathbf{d}_p, \mathbf{R}_{cp}, \mathbf{t}_{cp}, \dot{\mathbf{x}}_m^{1:N}]$
20: $\hat{\Psi} \leftarrow \text{bundleAdjust}(\Psi)$
21: **return** $\hat{\Psi}$

---

camera images $\{\mathscr{I}^1, \mathscr{I}^2, \dots \mathscr{I}^N\}$ to the calibration board model space. We obtain initial guess of camera intrinsics $\mathbf{K}_c$ and $\mathbf{d}_c$, as well as relative rotation $\mathbf{R}_{mc}^j$ and translation $\mathbf{t}_{mc}^j$ between the jth calibration board pose and the camera view space. A homography $\mathbf{H}^j$ between the calibration board and the camera image plane can then be calculated by Eq. 2, where $\mathbf{r1}_{mc}^j$ and $\mathbf{r2}_{mc}^j$ are the 1st and 2nd columns of $\mathbf{R}_{mc}^j$ of the jth pose.

Projector Calibration After initial camera calibration, we transform the SL pattern nodes from camera image space to calibration board model space by Eq. 3 in Alg. 1, where $\bar{\mathbf{x}}_c^j(i)$ is the undistorted homogeneous coordinates of node $\mathbf{v}_i$ in the camera image space, imaged at the jth pose. To be clear, we do not use checkerboard corners for projector calibration, instead we employ the SL nodes since they provide more robust and accurate initial guess.

Once we obtained the node point pairs $(\dot{\mathbf{x}}_m, \bar{\mathbf{x}}_p)$, Zhang's method is applied to calibrate the projector parameters, as shown in line 15 of Alg. 1. The relative translation and rotation between camera and projector are computed as shown in lines 16-17 of Alg. 1.

3.4 Bundle Adjustment

The imperfect planarity of the calibration board can bring errors to the initial calibration. Attacking this problem, given the initial camera and projector calibration, we propose a bundle adjustment (BA) algorithm (Stage 3 of Alg. 1) on the initial intrinsics and extrinsics, as well as node point coordinates $\mathbf{x}_m^{1:N}$ subject to reprojection errors.

Specifically, we set the world origin at camera optical center and let camera and projector parameters be:

$$\Psi_c = (\mathbf{K}_c, \mathbf{d}_c, \mathbf{r}_{mc}^{1:N}, \mathbf{t}_{mc}^{1:N}) \qquad (7)$$
$$\Psi_p = (\mathbf{K}_p, \mathbf{d}_p, \mathbf{r}_{cp}, \mathbf{t}_{cp}), \qquad (8)$$

where $\mathbf{r}_{mc}^{1:N}$ and $\mathbf{t}_{mc}^{1:N}$ are relative rotation and translation vectors of the calibration board with respect to the camera; $\mathbf{r}_{cp}, \mathbf{t}_{cp}$ are relative rotation and translation of the camera with respect to the projector.

Now, the camera-projector calibration problem can be formulated

as minimizing the following BA cost:

$$\{\hat{\Psi}_c, \hat{\Psi}_p, \hat{\mathbf{x}}_m^{1:N}\} = \underset{\Psi_c, \Psi_p, \mathbf{x}_m^{1:N}}{\arg\min} \left(\text{cost}(\Psi_c, \Psi_p, \mathbf{x}_m^{1:N}) \right), \qquad (9)$$

More specifically, suppose the jth calibration board pose has M^j nodes imaged on the calibration board, and denote $n_p = \sum_{j=1}^N M^j$. The objective function is formulated as:

$$\text{cost} = \sum_{j=1}^N \sum_{i=1}^{M^j} \left(\delta_c^j(i) + \delta_p^j(i) + \lambda\, \delta_m^j(i) \right). \qquad (10)$$

The first two terms represent reprojection errors of the node $\mathbf{v}_i$ in camera and projector image space:

$$\delta_c^j(i) = \|\mathbf{x}_c^j(i) - f(\Psi_c; \mathbf{x}_m^j(i))\|^2 \qquad (11)$$
$$\delta_p^j(i) = \|\mathbf{x}_p^j(i) - f(\Psi_p, \mathbf{r}_{mc}^{1:N}, \mathbf{t}_{mc}^{1:N}; \mathbf{x}_m^j(i))\|^2, \qquad (12)$$

where $f : \mathbb{R}^3 \mapsto \mathbb{R}^2$ projects a node coordinate $\mathbf{x}_m^j(i)$ from the calibration board model space to the camera/projector image space using camera/projector parameters.

In addition, we add a scale constraint that bounds the scale of model point coordinates during bundle adjustment:

$$\delta_m^j(i) = \|\hat{\mathbf{x}}_m^j(i) - \dot{\mathbf{x}}_m^j(i)\|^2 \qquad (13)$$

This term is necessary since the model point coordinates are coupled with extrinsic parameters $\mathbf{r}_{cp}$ and $\mathbf{t}_{cp}$. The original model point coordinates $\dot{\mathbf{x}}_m^j(i)$ are computed by Eq. 3. We introduce a weight λ to control the weight of the scale constraint, and empirically set $\lambda = \exp(-\delta_m^j(i))$.

We apply the trust-region-reflective algorithm to solve for Eq. 9. Since we introduce n_p node model coordinates $\mathbf{x}_m$ to bundle adjustment, leading $3 \times n_p$ extra parameters to optimize, a sparse Jacobian matrix pattern is designed to speed up numerical finite derivative computation.

4 EXPERIMENTS AND RESULTS

Our camera-projector pair consists of an Intel RealSense F200 RGB-D camera with image resolution of 640×480, and an Optima 66HD DLP projector set to the resolution of 800×600, as shown in Fig. 1. Note we only use RGB camera for calibration and reconstruction, the depth camera is employed only for generating ground truth. The distance between the camera and the projector is 1500 mm and all the calibration board poses are around 700 mm to 3000 mm in front of the camera-projector pair.

4.1 Evaluated Baselines

We compare our method with three other methods: a generalized global homography method, a multi-shot local homography method [20], and a degenerated baseline of the proposed method.

To compare with other camera-projector calibration methods (e.g., [3]), even if we replicate their calibration patterns (e.g., the ARTags pattern) and configurations, the obtained calibration points are different. It is hard to make a fair comparison this way. Instead, we generalize a method named **Global homography** to represent global homography-based methods in our experimental configuration and therefore we are able to use the same calibration points. For local homography-based method, we employ the popular system by **Moreno & Taubin** [20] with default parameters.

In addition, we generate a degraded version of the proposed algorithm by excluding the BA stage, named as **Proposed w/o BA**. In particular, this baseline only includes $\mathbf{r}_{mc}^{1:N}$ and $\mathbf{t}_{mc}^{1:N}$ in the nonlinear optimization. Unlike **Global homography** that uses points from a checkerboard, the degraded method applies points from an SL pattern. This strategy boosts both the number of feature points and their spatial distribution, and hence is more robust and accurate.

The three methods are tested together with the proposed one using both synthetic and real data. Root mean square (RMS) reprojection
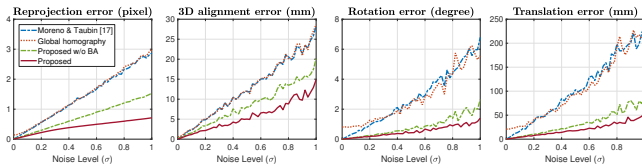
Figure 4: Synthetic data. Reprojection, 3D alignment, rotation and translation errors when noise level $\sigma = 0 \to 1$.

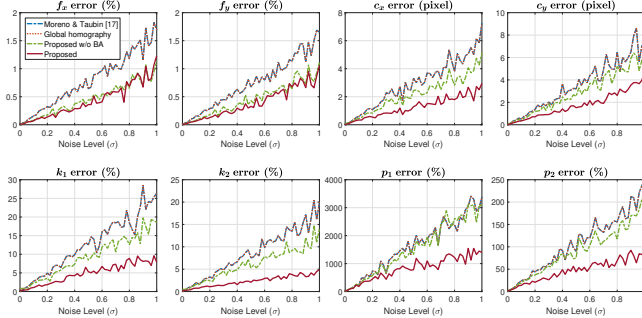Figure 5: Synthetic data. Errors in **camera intrinsics** for different noise levels $\sigma = 0 \to 1$. f_x and f_y are camera focal lengths in two directions, (c_x, c_y) is the camera image principle point, and k_1, k_2 and p_1, p_2 are radial and tangential distortion factors, respectively.

errors, 3D alignment errors, intrinsics and extrinsics errors are used as criteria. The synthetic data results (Fig. 4-Fig. 6) and real data results (Fig. 7) show clearly the benefits of the proposed method.

4.2 Synthetic Data

To compare the proposed method with baseline methods statistically, we first use synthetic data as benchmarks, where the camera and projector intrinsics and extrinsics, checkerboard corners and calibration board geometry are known. In particular, synthetic data provides absolute ground truth and accurate error measurement.

We start by generating the data by projecting the projector SL patterns to the world space. Each pair of node coordinate in projector image space $\mathbf{x}_p$ and projector optical center forms a ray that intersects with a set of predefined calibration boards, those intersections represent node's coordinates in the calibration board model space $\mathbf{x}_m$. Next we project $\mathbf{x}_m$ to the camera image space using pre-defined camera intrinsics and extrinsics, obtaining node's coordinates in camera image space $\mathbf{x}_c$. Finally, we follow the steps in Alg. 1 to calibrate our camera-projector pair.

We add Gaussian white noise with zero mean and standard deviation of σ to both camera and projector images. It is worth noting that, to simulate imperfect planarity, we also add Gaussian white noise to checkerboard and SL nodes in calibration board model space, whereas the noise units are in millimeters (mm). We generate the statistical benchmarks by inspecting the RMS reprojection errors, 3D alignment errors, intrinsics errors and extrinsics errors at each noise level $\sigma = 0 \to 1$. The 3D alignment errors are discrepancies between a synthetic 3D geometry and reconstructed 3D geometry.

The experiments are performed 100 times for each noise level σ and we plot the median of the errors as shown in Fig. 4 to Fig. 6. The proposed method clearly outperforms the other three methods. Moreover, **Global homography**'s reprojection error, rotation error, translation error and some projector intrinsics errors are nonzero even when the noise level $\sigma = 0$ due to its inability to model projector lens distortions (Fig. 4). In Fig. 5, **Global homography**'s and **Moreno & Taubin**'s curves overlap because they apply the same camera calibration method to the same set of checkerboard points.

4.3 Real Data

We evaluate our calibration using an imperfect planar white board with a printed checkerboard pattern glued to it. As shown in Tab. 2 column 2, our method is able to refine imperfect planar points and
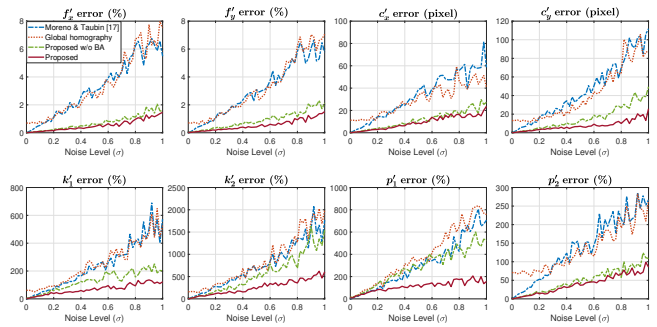

Figure 6: Synthetic data. Errors in **projector intrinsics** for different noise levels $\sigma = 0 \to 1$. f_x' and f_y' are projector focal lengths in two directions. (c_x', c_y') is the projector image principle point, and k_1', k_2' and p_1', p_2' are radial and tangential distortion factors, respectively.
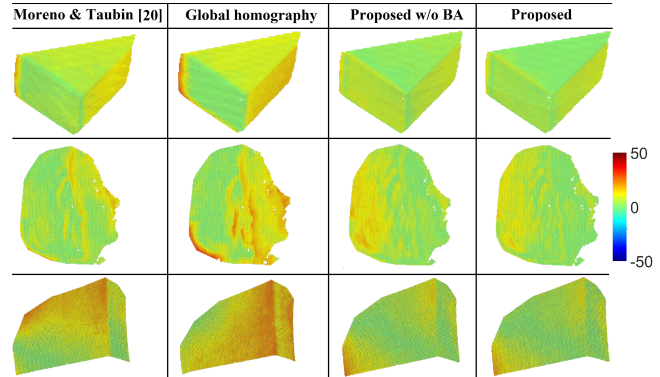

Figure 7: Real data. Reconstructed paper box (1st row), plaster bust (2nd row) and folded paper board (3rd row) using a camera-projector pair calibrated by the four calibration methods. Reconstruction errors (mm) are shown in pseudocolor.

inaccurate camera parameters using BA, thus leading to lower projector and stereo RMS reprojection errors than its counterparts. Note the stereo RMS reprojection error is the RMS of camera and projector reprojection errors.

It is worth noting that the projector RMS reprojection errors of the first two methods are high for two reasons: (1) they use Zhang's method to calibrate the camera-projector pair and thus suffer from imperfectness in the planarity of the calibration board. (2) The errors of extrinsics, i.e., $\mathbf{R}_{cp}$ and $\mathbf{t}_{cp}$, propagate to the projector (see Eq. 8 and Eq. 12), leading to enlarged high RMS reprojection errors.

One may notice that our camera reprojection error is a bit higher than the other two methods. This is because SL nodes are used for camera-projector calibration, while the reprojection errors are based on nodes rather than checkerboard points. Namely, reprojection errors solely are not sufficient to represent calibration accuracy if different set of points are employed.

Thus, we evaluate reconstruction errors by comparing the reconstructed point cloud with the ground truth. We first employ the calibration data from the four methods to reconstruct a point cloud using SL. Then the reconstruction errors are calculated as the RMS discrepancies between the SL reconstructed point cloud and the RGB-D camera captured point cloud. As shown in Fig. 7, a paper box, a plaster bust and a folded paper board are reconstructed using the calibration data of the four evaluated methods. The statistics of reconstruction errors are given in Tab. 2, columns 3-5, the **proposed w/o BA** method outperforms **Global homography** and **Moreno & Taubin** [20], since it applies SL nodes to calibration. The **proposed** method outperforms the degraded version, proving that BA is able to compensate for imperfect nonplanarity.

Our method outperforms both global and local homography-based methods on projector and stereo RMS reprojection errors and recon-

Table 2: Calibration RMS reprojection errors (pixels) (column 2) and reconstruction errors (mm) of real objects (columns 3-5).

Method	Reproj. errors (pixels)			Paper box (mm)			Plaster bust (mm)			Folded paper board (mm)		
	Cam.	Pro.	Stereo	Mean	Median	Std.	Mean	Median	Std.	Mean	Median	Std.
Moreno & Taubin [20]	**0.12**	1.59	1.13	8.47	7.08	5.93	5.60	4.72	3.93	9.82	9.69	5.72
Global homography	**0.12**	5.79	4.09	11.88	11.94	9.99	9.81	9.86	4.85	18.42	19.91	9.41
Proposed w/o BA	0.42	0.71	0.58	6.78	6.88	**4.10**	6.10	5.28	4.16	5.68	4.86	4.09
Proposed	0.35	**0.64**	**0.51**	**5.60**	**4.59**	4.70	**4.82**	**4.12**	**3.50**	**5.09**	**4.46**	**3.53**

struction errors. In practice, our single-shot method also overcomes the drawbacks of requiring many shots per pose, whereas **Moreno & Taubin** [20] needs 20 shots per pose, and one local homography per checkerboard corner. Additionally, Audet *et al.* [3] and Yang *et al.* [34] need at least two shots per pose for prewarp and additional time for incremental adjustment.

5 CONCLUSIONS

In this paper we present a flexible single-shot camera-projector calibration method. Compared with existing calibration systems, our method has two advantages: (1) Both synthetic and real data demonstrate that our method can refine imperfectly calibrated camera/projector parameters and imperfect planar calibration board points, thus leading to higher calibration accuracy and robustness against noises in planarity. (2) Requiring only a single shot of SL pattern per pose, our system enables fast and efficient calibration, especially in applications that need frequent re-calibration. Furthermore, the one-shot calibration provides a flexible and accurate results even when the board is handheld.

REFERENCES

[1] A. Albarelli, E. Rodolà, and A. Torsello. Robust Camera Calibration using Inaccurate Targets. In *BMVC*, pp. 16.1–16.10, 2010.

[2] H. Anwar, I. Din, and K. Park. Projector calibration for 3D scanning using virtual target images. *IJPEM*, 13:125–131, 2012.

[3] S. Audet and M. Okutomi. A user-friendly method to geometrically calibrate projector-camera systems. In *CVPRW*, pp. 47–54, 2009.

[4] A. Ben-Hamadou, C. Soussen, C. Daul, W. Blondel, and D. Wolf. Flexible calibration of structured-light systems projecting point patterns. *CVIU*, 117:1468–1481, 2013.

[5] N. Bird and N. Papanikolopoulos. Optimal Image-Based Euclidean Calibration of Structured Light Systems in General Scenes. *IEEE TASE*, 8:815–823, 2011.

[6] C.-Y. Chen and H.-J. Chien. An Incremental Target-Adapted Strategy for Active Geometric Calibration of Projector-Camera Systems. *Sensors*, 13:2664–2681, 2013.

[7] D. S. Dhillon and V. M. Govindu. Geometric and radiometric estimation in a structured-light 3D scanner. *MVA*, 26:339–352, 2015.

[8] J. Drareni, S. Roy, and P. Sturm. Geometric video projector auto-calibration. In *CVPRW*, pp. 39–46, 2009.

[9] M. Fiala. Automatic Projector Calibration Using Self-Identifying Patterns. In *CVPRW*, vol. 3, pp. 113–113, 2005.

[10] R. Furuakwa, K. Inose, and H. Kawasaki. Multi-view reconstruction for projector camera systems based on bundle adjustment. In *CVPRW*, pp. 69–76, jun 2009.

[11] J. Geng. Structured-light 3D surface imaging: a tutorial. *AOP*, 3:128, 2011.

[12] B. Huang and Y. Tang. Fast 3D reconstruction using one-shot spatial structured light. In *IEEE SMC*, pp. 531–536, 2014.

[13] L. Huang, Q. Zhang, and A. Asundi. Flexible camera calibration using not-measured imperfect target. *Applied Optics*, 52:6278, 2013.

[14] Z. Huang, J. Xi, Y. Yu, and Q. Guo. Accurate projector calibration based on a new point-to-point mapping relationship between the camera and projector images. *Applied Optics*, 54:347, 2015.

[15] H. Kawasaki, R. Furukawa, R. Sagawa, and Y. Yagi. Dynamic scene shape reconstruction using a single structured light pattern. In *CVPR*, pp. 1–8, 2008.

[16] M. Kimura, M. Mochimaru, and T. Kanade. Projector Calibration using Arbitrary Planes and Calibrated Camera. In *CVPR*, pp. 1–2, 2007.

[17] K. H. Lee, C. Je, and S. W. Lee. Color-stripe structured light robust to surface color and discontinuity. In *ACCV*, pp. 507–516, 2007.

[18] F. Li, H. Sekkati, J. Deglint, C. Scharfenberger, M. Lamm, D. Clausi, J. Zelek, and A. Wong. Simultaneous Projector-Camera Self-Calibration for Three-Dimensional Reconstruction and Projection Mapping. *IEEE TCI*, 3:74–83, 2017.

[19] T. T. Li, H. Y. Zhang, and J. Geng. Geometric calibration of a camera-projector 3D imaging system. *IVCNZ*, 2010.

[20] D. Moreno and G. Taubin. Simple, Accurate, and Robust Projector-Camera Calibration. In *3DIMPVT*, pp. 464–471, 2012.

[21] A. R. Orghidan, C. M. Gordan, D. A. Vlaicu, and B. J. Salvi. Projector-camera calibration for 3D reconstruction using vanishing points. In *IC3D*, pp. 1–6, 2012.

[22] J.-N. Ouellet, F. Rochette, and P. Hbert. Geometric calibration of a structured light system using control points circular. In *3DPVT*, pp. 183–190, 2008.

[23] T. Petkovic, T. Pribanic, and M. Donlic. Single-Shot Dense 3D Reconstruction Using Self-Equalizing De Bruijn Sequence. *IEEE TIP*, 25:5131–5144, 2016.

[24] C. Resch, H. Naik, P. Keitler, S. Benkhardt, and G. Klinker. On-Site Semi-Automatic Calibration and Registration of a Projector-Camera System Using Arbitrary Objects with Known Geometry. *IEEE TVCG*, 21:1211–1220, 2015.

[25] R. Sagawa, Y. Ota, Y. Yagi, R. Furukawa, N. Asada, and H. Kawasaki. Dense 3D reconstruction method using a single pattern for fast moving object. In *ICCV*, pp. 1779–1786, sep 2009.

[26] J. Salvi, J. Batlle, and E. Mouaddib. A robust-coded pattern projection for dynamic 3D scene measurement. *Patt. Recog. Letters*, 19:1055–1065, 1998.

[27] J. Salvi, J. Pagès, and J. Batlle. Pattern codification strategies in structured light systems. *Patt. Recog.*, 37(4):827–849, apr 2004.

[28] M. Shahpaski, L. Ricardo Sapaico, G. Chevassus, and S. Susstrunk. Simultaneous geometric and radiometric calibration of a projector-camera pair. In *CVPR*, 2017.

[29] K. H. Strobl and G. Hirzinger. More accurate pinhole camera calibration with imperfect planar target. In *ICCVW*, pp. 1068–1075, 2011.

[30] W. Sun and J. R. Cooperstock. An empirical evaluation of factors influencing camera calibration accuracy using three publicly available techniques. *MVA*, 17:51–67, 2006.

[31] T. van Aardenne-Ehrenfest and N. G. de Bruijn. Circuits and trees in oriented linear graphs. In *Classic papers in combinatorics*, pp. 149–163. 2009.

[32] S. Willi and A. Grundhöfer. Robust geometric self-calibration of generic multi-projector camera systems. In *ISMAR*, pp. 42–51, Oct 2017.

[33] S. Yamazaki, M. Mochimaru, and T. Kanade. Simultaneous self-calibration of a projector and a camera using structured light. In *CVPRW*, pp. 60–67, 2011.

[34] L. Yang, J. M. Normand, and G. Moreau. Practical and precise projector-camera calibration. In *ISMAR*, pp. 63–70, Sept 2016.

[35] B. Zhang, Y. Li, and Y. Wu. Self-recalibration of a structured light system via plane-based homography. *Patt. Recog.*, 40:1368–1377, 2007.

[36] L. Zhang, B. Curless, and S. M. Seitz. Rapid shape acquisition using color structured light and multi-pass dynamic programming. In *3DPVT*, pp. 24–37, 2002.

[37] S. Zhang and P. S. Huang. Novel method for structured light system calibration. *Optical Engineering*, 45:45 – 45 – 8, 2006.

[38] X. Zhang, Z. Zhang, and W. Cheng. Iterative projector calibration using multi-frequency phase-shifting method. In *ICCIS and RAM*, pp. 1–6, 2015.

[39] Z. Zhang. A flexible new technique for camera calibration. *IEEE TPAMI*, 22:1330–1334, 2000.