MVA2013 IAPR International Conference on Machine Vision Applications, May 20-23, 2013, Kyoto, JAPAN

Calibration of the Projector with Fixed Pattern and Large Distortion Lens in a Structured Light System

Xingdou Fu[a], Zuofu Wang[a], Hiroshi Kawasaki[a], Ryusuke Sagawa[b], Ryo Furukawa[c]
[a]Kagoshima University, Kagoshima 890–8580, Japan Email: fuxd@ibe.kagoshima-u.ac.jp
[b]National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki 305–8568, Japan
[c]Hiroshima City University, Hiroshima 731–3166, Japan

Abstract

The most critical factor affects accuracy of a Structured Light System (SLS) is calibration. Camera calibration is easy to complete because of its extensive study. To simplify projector calibration, previous work models the projector as an inverse camera and tries to build similar 3D-2D mapping data for projector calibration. Achieved mapping data is directly fed to some classic two-step camera calibration methods. When projector comes with a large distortion lens, this kind of methods will fail because their first steps use closed-form solution to calculate initial guess for optimization in next steps. We proposed a new method to calibrate the projector by removing its distortion first. Because projector cannot "see" anything, not like camera case, constraints such as "straight lines remain straight" working just on 2D image is invalid for distortion estimation. With 3D-2D mapping data, the estimation will involve several extra unknowns into a non-linear optimization. We use partial mapping data whose 2D points in a "small central area" of projector pattern image to acquire an initial guess for those unknowns, and then use all mapping data to refine them and estimate distortion parameters. Experiments show our method can still calibrate the projector when classic methods fail.

1. Introduction

Nowadays Structured Light System (SLS) is extensively being used in various applications such as reverse engineering, augmented reality, medical examination, games, movies… A typical SLS is composed of a camera and a projector. Usually the projector projects contrasted pattern image(s) to encode 3D surface of an object, which can be easily decoded through perceived image by the camera. Because this technique significantly alleviates the matching problem at scanning phase, it comes with some inherent and unique features such as high accuracy and fast speed. SLS has already become one of the most important Non-contact 3D shape measurement methods. Especially single shot SLS is one of the dominant methods for fast moving object (or camera moving instead of object) scanning.

The most critical factor affects the accuracy of a SLS is calibration. The calibration estimates intrinsic parameters of the camera and the projector plus extrinsic ones between them. Because a projector cannot "see" anything as a camera does, it is not straightforward to build 3D-2D mapping data for its calibration, neither to estimate its lens distortion separately. Projector calibration happens to be the most difficult part of SLS calibration.

Since a lot of researchers have already done extensive studies on camera calibration[1][2], to simplify projector calibration, usually projection of a projector is modeled as an inverse procedure of a pin-hole model camera. Most of previous research [3-6] on projector calibration focused on how to set up 3D-2D mapping data and neglected the necessity of separate estimation of its lens distortion. The acquired 3D-2D mapping data then is directly fed to some classic two-step camera calibration methods[1]. However, classic two-step calibration methods themselves cannot deal with large distortion lens directly, because the first step of this kind of methods uses closed-form solution to calculate initial values for partial parameters. Then the second step is expected to estimate distortion coefficients and refine all parameters by non-linear optimization. With severely distorted 3D-2D mapping data, the classic method will fail to calculate initial values. **Without good initial guess, the search for the global minimum can be difficult and can be easily trapped in a local minimum due to the large number of unknowns and the ill conditioning of the problem** [3]. Solutions will be either good initial guess can be offered or the 3D-2D mapping data is undistorted in advance. Actually both of them inevitably need a separate estimation of lens distortion. Unfortunately, as mentioned above, estimation of lens distortion for projectors is not as straightforward as for cameras.

In the case of cameras, lens distortion can be simply considered has nothing to do with perspective model, which only occurs on 2D perceived image. Perspective invariants such as "straight lines remain straight lines" [7] can be used for constraints on this 2D image. Except distortion parameters, only a principal point or distortion center (some researchers distinguish them into two points, some consider them as one) is needed for a non-linear optimization. This fact much decreases the number of unknowns. **In the case of projectors**, the 2D image on projector image plane is the pattern to be projected out and it is not an observation on 3D entities (such as straight lines). Not like camera case, the 2D pattern image itself here leaves us nothing to be optimized. This is the essential difference for considering lens distortion between a camera and a projector.

Unfortunately, in many circumstances customized projectors with high distortion lens are utilized in a SLS such as endoscope scanners. The reason is lens design has to compromise to the extremely small size limit of endoscope scanners. Another common requirement for micro scanners is to use a projector with fixed physical pattern instead of a LCD projector. Given above reasons, **this paper will focus on the calibration of the projector with fixed pattern and high distortion lens in a SLS.**

2. Related work

Because a projector cannot "see" anything, its calibration must be completed with the help from a camera, especially during the stage to build 3D-2D mapping data for projector calibration. Depending on different ways making use of the camera, previous work can be roughly classified into two categories (1) **based on uncalibrated camera**; (2) **based on calibrated camera**. The first group methods use cameras only for capturing images. Usage here has nothing to do with its perspective parameters. One advantage is that accuracy of projector calibration does not depend on the accuracy of camera calibration. Methods in this category can be further divided into two sub categories (a) **One to One Pixel Mapping** [3]; (b) **Active Feature Point Matching** [4]. The second group methods use camera parameters to help with acquiring 3D-2D mapping data for projector calibration. They leaves projector calibration accuracy depending on camera but they are easier and faster to implement. Methods in this category can also be further divided into two sub categories (a) **Back Projection** [5] ; (b) **Non-mapping Method** [6].

Legarda-Sáenz [3] describes an absolute phase measurement method, which uses phase shifting patterns to encode every pixel on projector pattern image. On camera side, captured pattern image can be decoded inversely. Thus, a one to one mapping at pixel level from camera image to projector pattern image is available. Since any 3D points captured by camera can be mapped to projector image plane, it is easy to achieve 3D to 2D mapping data for projector calibration. Major disadvantage of this method is luminance nonlinearity of projectors can cause inaccurate camera-projector mapping. Martynov [4] uses an active way to achieve the mapping between camera image and projector image. By observing from camera side, they iteratively adjust position of each projected features until all of them coincide with printed features on calibration board. Then every printed 3D feature has a 2D corresponding feature on projector image plane. The challenge for this kind of methods is the accuracy of feature coincidence. Ouellet [5] calibrates the projector with the calibrated camera and back projection. Methods based on back projection use plan-based method to calibrate the camera first, and then calculate 3D coordinates of projected feature points on checker board with calibrated camera parameters. These 3D points and 2D feature points on projector pattern give 3D-2D mapping data for projector calibration. Kimura [6] calibrates projector with calibrated camera but does not try to find 3D-2D mapping data for projector calibration. Instead, it utilizes the idea of Structure from Motion. Because fundamental matrix needs to be induced by homographies between the camera and the projector as its first step, it is inevitable for lens distortion from both camera and projector to contaminate the data.

All previous work mentioned here use LCD projector with non-fixed digital pattern and **none of them considers a separate projector lens distortion to fit the case projector lens comes with large distortion.**

3. Our method

To calibrate the projector with fixed pattern and large distortion lens in a SLS, our method first use back projection to build 3D-2D mapping data for projector calibration since **this is the only feasible method for fixed physical pattern**; and then estimate the projector lens distortion separately with 3D-2D mapping data; finally feed undistorted 3D-2D mapping data to a revised classic two -step calibration method (utilizes distortion-free pin-hole model by setting and fixing all distortion parameters to zeros). The reason why last step does not use a closed-form solution is because even if the actual camera is distortion free, nonlinear optimization can still improve the closed-form solution [1]. The difficulty here is if the estimation of projector lens distortion has to be based on 3D-2D mapping data, planar perspective transformation or pin-hole model will be involved, which could introduce up to 8 or 9 extra unknowns into a non-linear optimization.

3.1. Distortion model and projector model

We use a sophisticated polynomial distortion model which only includes radial distortion, as seen in formula (1), since usually tangential distortion is small and ignored.

$$\begin{aligned} x_d &= (1 + K_1 r^2 + K_2 r^4) x_u \\ y_d &= (1 + K_1 r^2 + K_2 r^4) y_u \end{aligned} \qquad (1)$$

(x_u, y_u) is a 2D point on image plane; (x_d, y_d) is its distorted coordinate; K_1, K_2 are distortion coefficients; r is the distance between (x_u, y_u) and distortion center. Usually, second order is enough for the distortion model to achieve decent accuracy. Higher order may cause instability instead of further accuracy. Undistortion compensation can be solved with Taylor expansion or iterative method. For projection, we use classic pin-hole model to simulate an inverse camera. Intrinsic and extrinsic parameters are same as those in the plane-based camera calibration [8].

3.2. Distortion estimation with partial 3D-2D mapping data

The distortion estimation method relies on 3D-2D (also 2D-2D since it is plane based, we will use the word "3D" to represent 2D points on plane also) mapping data as input and utilizes a fact that central image area has less distortion, which was ever mentioned in some papers such as [9]. Considering one set 3D-2D mapping points (3D points on the calibration board at one orientation), by abuse of notation, if without distortion considered, they can be represented with formula:

$$x_i = HX_i$$

x_i is a 2D point, X_i is the corresponding 3D point on checker board;

With distortion considered, they can be represented with formula:

$$x_i = F_d(HX_i)$$

Function $F_d(x)$ is the distortion function based on the model as described in formula (1).

Because of noise, $F_d(HX_i)$ will not equal to x_i exactly. By minimizing formula (2) with non-linear optimization (Levenberg-Marquardt method) ,we can get estimated distortion coefficients, distortion center and homography matrix. The initial value(s) for distortion coefficients are set to zeros, for distortion center is set to the center of projector pattern image; initial value of homography matrix can be calculated with partial mapping data whose 2D points in a "small central area" of projector pattern image. How to decide the "small central area" will be explained later in the paper.

$$\{\hat{K}_1,\hat{K}_2,\hat{c},\hat{H}\} = \mathrm{argmin}_{K_1,K_2,c,H}\{\textstyle\sum_{i=1}^{n}|x_i - F_d(HX_i)|\} \quad (2)$$

c is the distortion center, n is the total number of 2D points in one set 3D-2D mapping points.

3.3. Distortion estimation with full 3D-2D mapping data

The back projection calibration method usually utilizes multiple sets of 3D-2D mapping points due to several observations on multiple orientations of checker board. Apply the minimization in formula (2) (suppose we call it atomic optimization) to all sets mapping data (suppose we call it one "round"), we will have different sets of estimated variables as well. **Among estimated variables, distortion coefficients, distortion center are supposed to be always same for each set of 3D-2D mapping points.** To improve the accuracy of optimization result, multiple "rounds" optimizations can be performed as described in formula (3). At the end of each round, select the "best" estimated distortion coefficients and distortion center in current round, whose atomic optimization has the minimal Mean Square Error (MSE). At the beginning of next round, set initial values of distortion coefficients, distortion center for all atomic optimizations to the "best" estimated values from previous round.

$$\begin{aligned}&\{\hat{K}_{1j},\hat{K}_{2j},\hat{c}_j,\ \hat{H}_j\} = \mathrm{argmin}_{K_{1j}K_{2j},\,c_j,\,H_j}\{\textstyle\sum_{i=1}^{n}|x_{ij} - F_d(H_jX_{ij})|\}\\ &\{K_{1j+1},\ K_{2j+1}, c_{j+1}\} = \mathrm{argmin}_{\hat{K}_{1j},\hat{K}_{2j},\hat{c}_j}\{MSE_j\},\ \ j\epsilon\{0,\ 1,\ \ldots m\}\end{aligned} \quad (3)$$

j represents the index of different set 3D-2D mapping data, and *m* is the total number.

The reason for multiple rounds optimization is that noise may impose different effect on each case checker board at different orientation. For example, same checker board may cause back projection accuracy worse when the board is less vertical to camera optical axis. There is a chance to get more accurate optimized variables from current round for next round initialization. Using all 3D-2D mapping data in only one optimization means optimization needs compromise to all data.

3.4. Define "small central area"

The initial value of homography matrix H for optimization should be close enough to ideal value, but not necessarily very accurate, because H will be refined by further optimization as a variable instead of being used as a constant.

Consider the 2D points in one set 3D-2D mapping data: based on distortion center, a blank area expands starting from the nearest to furthest 2D points. When each point is included, by calculating the homography between 2D points currently inside the area and their corresponding 3D points, we can get a value of residual. Figure 1 shows an example of the relation between MSE and number of points. Set MSE to some a threshold such as 1.0, we can get the number of points (178 in the sample) around the distortion center. The selection of the threshold can be affected by many factors such as feature points density in the pattern and etc., therefore for general cases threshold setting should be subject to the specific experiment configuration.

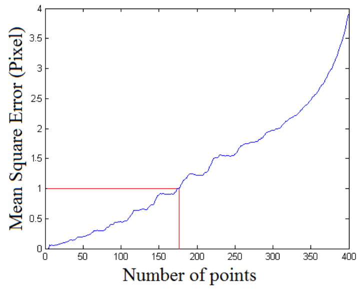

Figure 1. MSE with different number of points

3.5. Summarization of our method

Our method has couple different aspects from the methods [9][10] using similar ideas. The method in paper [10] does not make use of points in a "small central area" to calculate a homography matrix. They directly use identity matrix as initial value for perspective transformation in their optimization. The method in paper [9] uses calculated homography matrix for optimization but as a constant. Experiments in our paper show that optimizing the matrix and multiple "rounds" optimization can improves the accuracy. And multiple "rounds" optimization can recover wrongly-set distortion center to some extent. Another difference from the method in paper [9] is the way to decide "small central area". They use a greedy algorithm trying to make H as accurate as possible at initialization phase, which could be very time-consuming with dense pattern.

4. Experiments and conclusion

To evaluate estimated parameters, experiments **I** to **IV** are conducted on simulated data since it is straightforward to get ground truth. Our method is based on back projection and plane-based camera calibration, so the simulation data is generated as following steps (the scenario is as seen in figure 2; An example of the feature point images on camera image plane, calibration board and projector image plane is shown in figure 3):

- **(1)** 400 feature points (wave line pattern) on projector image plane are projected to 3D space through projector optical center respectively;

(2) These projected 3D rays from last step intersect with 9 calibration boards at different orientations (represent one calibration board at 9 different orientations) and then generate 9 groups of 3D points;

(3) All 3D points from last step are captured by the camera and leave 9 groups of 2D points on camera image plane.

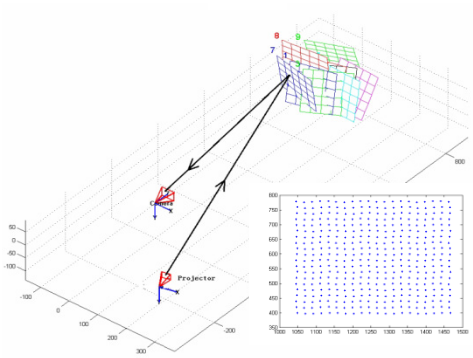

Figure 2. Simulated scenario and wave line pattern

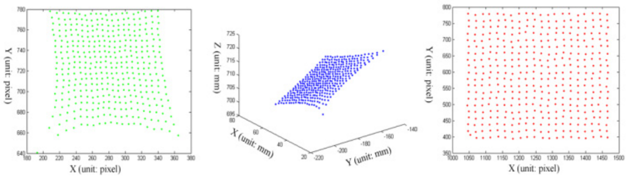

Figure 3. Images of feature points on (from left to right) camera image plane, calibration board and projector image plane

Simulated input for our method will be 9 groups (9*400) 2D points on camera image plane and 400 feature points (wave line pattern) on projector image plane. White noise [-0.1, 0.1] (unit: pixel) is added to input data to simulate camera reprojection error and feature extraction error.

Default parameters include projector pattern image resolution (2600*1300), MSE threshold (1 pixel) for defining "small central area". Other ground truth is shown in each table. Because we generate simulation data based on tools in [8]. Ground truth of distortion coefficients in simulation data (K_1=-8, K_2=2) are converted already to the format in our method (K_1=-8.888e-07, K_2=2.469e-14).

The experiment I is to show the result of distortion estimation (in table 1)and projector calibration (in table 2). **Calibration toolbox in [8] stops working under this distortion.** **The experiment II** (in table 1) is to show slightly different size of "small central area" does not change the final estimation too much. MSE threshold for defining "small central area" is set to 2 pixels here. **The experiment III** (in table 1) is to show putting homography matrix into optimization as a constant instead of a variable achieves less accurate result. In this experiment homography matrix is used as a constant. **The experiment IV** (in table 3) is to show multiple rounds optimi-optimization can deal with the situation that distortion center is not around plane image center. The initial distortion center is set to (1400,750). In above experiments, we can see K_2 is very sensitive to noise, and other parameters are more meaningful to the evaluation.

Table 1. Distortion estimation (different methods)

Variables	***K1***	***K2***	**Distortion center**	**Minumal MSE**
Ground truth	-8.888e-07	2.469e-14	(1285,640)	/
Experiment I	-8.874e-07	5.193e-14	(1284.597, 640.251)	0.0797
Experiment II	-8.966e-07	9.246e-14	(1285.001, 640.727)	0.0867
Experiment III	-6.258e-07	-1.761e-12	(1283.727, 639.715)	0.3492

Table 2. Projector calibration

Variables	**Focal length**	**Principle point**	**Reprojection error (SD)**
Ground truth	(3000,3000)	(1285,640)	/
Experiment I	(2999.812, 3000.419)	(1274.182, 641.682)	(0.252, 0.272)

Table 3. Distortion estimation (different rounds)

Variables	***K1***	***K2***	**Distortion center**	**Minumal MSE**
Ground truth	-8.888e-07	2.469e-14	(1285,640)	/
Experiment IV (Round 1)	-4.636e-07	-2.828e-14	(1344.517, 727.918)	2.1680
Experiment IV (Round 5)	-8.911e-07	1.191e-14	(1344.516, 639.192)	0.0824
Experiment IV (Round 10)	-8.895e-07	9.390e-15	(1284.442, 638.432)	0.0825

Acknowledgment

This work was supported in part by SCOPE No.101710002 and NEXT program No.LR030 in Japan.

References

[1] Juyang Weng. Camera Calibration with Distortion Models and Accuracy Evaluation. IEEE Trans. Pattern Anal. Mach. Intell., 14(10):965-980. Oct 1992.

[2] Zhengyou Zhang. Flexible Camera Calibration By Viewing a Plane From Unknown Orientations. ICCV, Pages 666-673, 1999.

[3] Ricardo Legarda-Sáenz. Accurate procedure for the calibration of a structured light system. Opt. Eng., 43(2):464-471, Feb 2004.

[4] Ivan Martynov. Projector Calibration by "Inverse Camera Calibration". SCIA, Pages 536-544, 2011.

[5] Jean-Nicolas Ouellet. Geometric Calibration of a Structured Light System Using Circular Control Points. 3DPVT, Jun 2008.

[6] M. Kimura. Projector Calibration using Arbitrary Planes and Calibrated Camera. CVPR, Pages 1-2, Jun 2007.

[7] Frédréric Devernay. Straight lines have to be straight: automatic calibration and removal of distortion from scenes of structured enviroments. MVA, 13(1):14-24, Aug 2001.

[8] Jean-Yves Bouguet. Camera Calibration Toolbox for Matlab, http://www.vision.caltech.edu/bouguetj/calib_doc/.July2000

[9] A. Nowakowski. Homography of central points for optical distortion compensation. Opto-Electronics Review, 15(4):202-209, Dec 2007.

[10] Leonardo Romero. Correcting Radial Distortion of Cameras with Wide Angle Lens Using Point Correspondences. Scene Reconstruction, Pose Estimation and Tracking, Page 530, Jun 2007.