Practical and precise projector-camera calibration

Liming Yang, Jean-Marie Normand, Guillaume Moreau

Practical and precise projector-camera calibration

Liming Yang* Jean-Marie Normand† Guillaume Moreau‡

Ecole Centrale de Nantes, UMR CNRS 1563 AAU, France

ABSTRACT

Projectors are important display devices for large scale augmented reality applications. However, precisely calibrating projectors with large focus distances implies a trade-off between practicality and accuracy. People either need a huge calibration board or a precise 3D model [12]. In this paper, we present a practical projector-camera calibration method to solve this problem. The user only needs a small calibration board to calibrate the system regardless of the focus distance of the projector. Results show that the root-mean-squared re-projection error (RMSE) for a 450cm projection distance is only about 4mm, even though it is calibrated using a small B4 (250×353mm) calibration board.

Index Terms: H.5.1 [INFORMATION INTERFACES AND PRESENTATION (e.g., HCI)]: Multimedia Information Systems—Artificial, augmented, and virtual realities;

1 INTRODUCTION

Projector-camera systems (referred to as ProCam systems in the following) are important vision systems that raise a lot of interest both for research and industry. They can indeed be used for many applications such as 3D measurement and augmented reality (AR). In AR applications, cameras are used to capture the environment and track its changes, and projectors are then used to project (i.e. to augment) information directly onto objects. This projection-based augmentation is commonly referred to as Spatial Augmented Reality (SAR) [2]. Unlike head-mounted or hand-held AR, one of the great advantages of SAR is that users are free of instruments. Users can interact with the environment more naturally, especially when they have complex tasks to carry out, such as in industrial productions and maintenance.

Calibration is one of the most essential elements of Projector-Camera systems, namely determining the intrinsic matrix and the distortion coefficients of cameras and projectors, as well as their relative position. A precise calibration will produce a precise mapping between the physical world and camera/projector images, which permits the augmented information to be projected to their desired positions in SAR. Although there exists very mature and simple methods for camera calibration [18], calibrating a projector remains difficult and unpractical, especially when projecting at large distances [12].

In this paper, we propose a method to calibrate a ProCam system consisting of one camera and one projector more practically, i.e. by using only a small B4 printed pattern attached to a rigid board. After a review of existing techniques, we present our method. Results are provided for both calibration and augmentation evaluations.

*e-mail: liming.yang@ec-nantes.fr

†e-mail: jean-marie.normand@ec-nantes.fr

‡e-mail: guillaume.moreau@ec-nantes.fr

1.1 Related works

There exist two families of methods for projector calibration: calibrating each light stripe [14] or each pixel [9] of the projector, and calibrating the projector as a whole using the pinhole model. The former is very time consuming since each light stripe or pixel needs to be calibrated individually. The latter mainly features two categories of techniques: two-views based methods and inverse camera methods.

As an example of "two-views based methods", Yamazaki et al. [15] use structured-light patterns to create dense point correspondences between the camera and the projector views to find the fundamental matrix by exploiting two-view geometry properties. The final intrinsic matrices of both the projector and the camera are calculated iteratively by assuming their initial values. This method has no error propagation but the matrices of the projector and of the camera have a coupling-effect. It is sensitive to initial values and only estimates 3 parameters of the intrinsic matrix. With the help of a known precise 3D model of a target object, Resch et al. [12] use bundle adjustment to iteratively correct the estimated matrices without dependence on initial values. However, a precise 3D model is not always available in the general case.

The "inverse camera" methods, treat projectors as inverse cameras, so that they can be calibrated using Zhang's method [18]. More precisely, a calibration board with a pre-defined pattern is used to calibrate the camera. In the meantime, the projector projects other patterns onto the calibration board. These projected patterns are used to find camera-projector point correspondences in order to calibrate the projector. Several methods exist for finding such correspondences. Structured-light methods [17, 8] employ series of coded gray-bars images. Even though methods for defocused projector exists [7], the board needs to be frozen at the same position by the user for several seconds, which is not convenient. Other methods capture fewer or even a single image for each board position, e.g. by projecting regular dot patterns [11], chessboards [5] or a matrix of ARToolKit markers [1]. Our method follows a similar approach.

In [5] and [1], the calibration board has to be positioned in the focus zone (where projected images are clear), or the pattern becomes too blurred to be recognized. Moreover, Zhang's method [18] requires the pattern to occupy a large part of the field of view of the projector in order to give a better result. Consequently if a wide-angle projector is focusing at about 2m, the user needs to manipulate at least an A0 size (841×1189mm) calibration board, which is a burden. The user can also choose to calibrate the projector at a short distance and then change the focus depending on the application, but this will induce a loss of precision, as showed later in Section 4. We also experimentally noticed that for Audet et al.'s method [1], projectors' brightness greatly influences markers recognition.

Ouellet et al. [11] need to take 3 images for each board position while the projector is successively projecting (1) a white image, (2) a regular pattern, (3) a calibration regular pattern. These images are highly related: (1) is subtracted from (2) to find the homography between the projector and the camera (H_{cp}). If the board moves or ambient lighting changes, detection may fail. (3) needs H_{cp} to interleave projected points between printed ones. If the board moves,

points may interfere. Moreover, a few missing points in the regular pattern can lead to detection failures. Thus the whole process is not very robust. At last, this method detects projected dot centers by back projecting them onto the board plane. When the board is out of focus, back projected dots may be heavily deformed, which can make the center detection difficult and introduce large errors.

1.2 Contribution

Compared to [1] and [11], our work has the following advantages which makes it much more practical and usable:

- it can work at large distances without having to manipulate unpractical huge calibration boards, unlike [1]
- it uses random dot patterns, which are robust to pattern interference and insensitive to lighting, unlike both [1] and [11]
- it only needs one image for each board position, thus has a faster recovery from detection failure, unlike [11]
- it gives more stable results for intrinsic matrices of cameras and projectors, compared to [1]

We believe this work can make projector calibration much easier both for academic researchers and spatial augmented reality users, especially when projecting at large distances.

2 Calibration Method

Our method relies on simple manipulations of a calibration board whatever the focus distance used on the projector: the user holds a small calibration board before the camera and the projector. If the board is still for a while ($\approx$ 1s), an image is automatically captured for calibration. Once a pre-defined number (K, cf. Section 3) of images are acquired, the system is calibrated (cf. Figure 1).

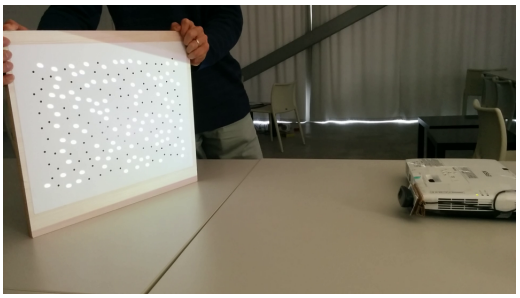

Figure 1: Calibration: only a small board is manipulated whatever the focus distance of the projector.

In the following, we first introduce some basic notations and the generic "inverse camera" method. Then we present our calibration patterns and how to find feature correspondences. Finally, the whole calibration procedure is presented.

2.1 Basic notations and inverse camera method

The ProCam system is described by the camera intrinsic matrix K_c, its distortion coefficients D_c, the projector intrinsic matrix K_p, its distortion coefficients D_p, the projector rotation in camera frame R and the position of the projector's optical center in camera frame t. More precisely, the intrinsic matrix is defined by [3]:

$$\mathbf{K}_\varepsilon = \begin{pmatrix} f_{\varepsilon x} & \gamma_\varepsilon & u_\varepsilon \\ 0 & f_{\varepsilon y} & v_\varepsilon \\ 0 & 0 & 1 \end{pmatrix} \quad (1)$$

where ε stands for c or p, $f_{\varepsilon x}$ and $f_{\varepsilon y}$ are focal lengths, $(u_\varepsilon, v_\varepsilon)$ is the position of the principal point, γ_ε is the skew.

The "inverse camera" method starts with two sets of correspondences for each view of the calibration board: $c_{cb} = \{(\mathbf{x}, \mathbf{x}^{(c)})\}$ and $c_{cp} = \{(\mathbf{y}^{(p)}, \mathbf{y}^{(c)})\}$, where $\mathbf{x}$ is an interest point (corner/center of markers, etc.) on the calibration board, $\mathbf{x}^{(c)}$ is its image in the camera; $\mathbf{y}^{(p)}$ is an interest point in the projected image, $\mathbf{y}^{(c)}$ is its image viewed by camera. With c_{cb} from $K (K \geq 3)$ views, the camera can be first calibrated by Zhang's method [18]. Lens distortion in $\mathbf{x}^{(c)}$ and $\mathbf{y}^{(c)}$ can then be removed.

For the projector, the Board-Camera homography H_{cb} is found with c_{cb}. The location on the calibration board of projected features $\mathbf{y}^{(p)}$ can then be expressed as $H_{cb}^{-1}\mathbf{y}^{(c)}$, thus a new correspondence set $c_{pb} = \{(\mathbf{y}^{(p)}, H_{cb}^{-1}\mathbf{y}^{(c)})\}$ can be established for each view. At last, the projector can be calibrated with c_{pb} from K views.

2.2 Calibration pattern

We choose to use randomly distributed circle points as both physical patterns and projected patterns, with the following advantages:

- Circle points have no internal structures unlike ARToolKit markers, they are thus less influenced by defocus or blur
- Different random dots-based markers can be distinguished, real-time tracked and accurately located even in the case of partial occlusion and over/under-detections, unlike regular patterns. This makes the method robust against pattern overlapping.
- Points occupy a small surface: the same area can contain more points than other geometries, so more correspondences can be established

In order to track patterns and find point correspondences, we use the Local Geometric Consensus (LGC) algorithm [16]. It is a general algorithm for matching/tracking random point patterns (2D/3D) undergoing various geometric transformations (e.g. affine, perspective) in real-time. LGC relies on local subset geometric distribution as characteristic features to generate rough hypotheses and uses neighboring subset geometric consensus to reject false matchings. It works in two phases: during the off-line registration phase, it splits model point patterns into different local patches and registers them using Geometric Hashing. During the on-line matching phase, local patches from scene point patterns are used to match with pre-registered model patches. Once a local match is found, LGC tries to find more geometrically coherent matchings in the neighborhood. If the matching contains enough correspondences, a model is claimed to be found. The method remains accurate with moderate point jitter, it is also robust to over/under detection and partial occlusion. So it is very suitable for calibration purposes.

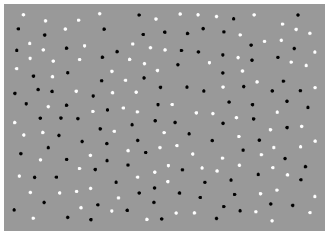

Figure 2: Calibration Pattern: P_b black dots printed on a piece of paper, P_p in white are projected. Both form the original pattern P_o.

The calibration pattern (cf. Figure 2) is constructed as follows: $2N$ points are randomly generated in a rectangle region, called the

original pattern (P_o) afterwards. The first half of these points (P_b) are printed on a paper and attached to a rigid board, the second half (P_p) is pre-warped and projected onto the board in order to form P_o during the calibration procedure (cf. Section 2.3). We set a minimum inter-point distance for all $2N$ points to have a more homogeneous distribution, which improves camera calibration according to [13]. Since our objective is to calibrate the projector with a far focus by using a small calibration board, the board does not have to be close to the focus plane of the projector, so the projected points will be defocused and become larger. The minimum inter-point distance requirement makes the defocused projected points less likely to overlap with printed points.

2.3 Algorithm

We use LGC to find point correspondences, the two basic operations of this algorithm are defined as follows:

$$LGC.register(P) \tag{2}$$

$$[\tilde{P}, P_i, H] = LGC.match(I) \tag{3}$$

Operation (2) registers a model point pattern P into LGC, several models may be registered; Operation (3) matches detected black points in image I with all registered models. If a model P_i is found in the detected pattern, operation (3) returns P_i and their corresponding detected points $\tilde{P}$. An homography $H : \tilde{P} = H(P_i)$ is also returned.

The algorithm contains three parts: initialization, manipulation and improvements (cf. Algorithm 1). For initialization, both P_b and P_p points are registered into LGC as models. Initial value H_{pre} is defined so that $H_{pre}(P_p)$ exactly fits the projector resolution.

During manipulation, the algorithm matches P_b (i.e. calibration board) with LGC. Once the board is found, it starts to detect and match P_p on the calibration board in the same way. If P_p is found as well, an homography H_{pre} is computed which permits to warp the projected pattern to the right position for the next iteration (cf. Figure 3). Note that H_{pre} only warps points' positions. When both patterns are well *aligned* (meaning that they form well P_o) and the board is *steady* for $\approx$ 1s, image I is captured, both H_{bc} and H_{pc} are recorded. We choose $t_s = 3.0$ pixels and $t_a = 2.0$ pixels in Algorithm 1 empirically.

Once images from K views are obtained, an improvement step is executed. It is used to deal with the fact that the center of a circle is not perspective invariant. The approach is similar to [4]: each captured image I is rectified so that points are nearly circle. Then the OpenCV's MSER detector [10] is used to detect ellipse centers. The error introduced by this method in the defocused case is discussed in the following section.

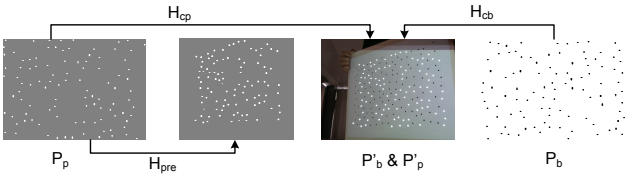

Figure 3: Point patterns used in Algorithm 1. From left to right: P_p, projected pattern $H_{pre}(P_p)$, P'_b and P'_p in camera view, board pattern P_b. $H_{cp} = H_{cb}$ if P'_b and P'_p are well aligned on board.

The initialization step takes $\approx$ 1s, the manipulation sequence is real time, the improvement step can process $2-3$ frames per second.

Algorithm 1 Algorithm

$LGC.register(P_b)$
$LGC.register(P_p)$
set H_{pre} such that $H_{pre}(P_p)$ exactly fits in projector resolution

while Less than K views are captured **do**
  Projector: Draw circles at $H_{pre}(P_p)$ and projecting
  Camera: Get new image I
  $[P'_b, P_b, H_{cb}] = LGC.match(I)$
  if P'_b found **then**
    $[P'_p, P_p, H_{cp}] = LGC.match(\tilde{I})$, $\tilde{I}$ is I with inverted color.
    if P'_p found **then**
      $steady = \forall p \in P'_b$ has moved less than t_s since last image
      $aligned = \forall p \in P_b, \left\| H_{cb}(p) - H_{cp}(p) \right\| < t_a$
      if $steady$ && $aligned$ **then**
        if last over 1s **then**
          Capture I, record correspondences (P'_b, P_b), $(P'_p, H_{pre}(P_p))$ and homographies H_{cb}, H_{cp}
        end if
      else
        $H_{pre} = H_{pre} H_{cp}^{-1} H_{cb}$
      end if
    end if
  end if
end while

Improve detection of P'_b in rectified views $I_b = H_{cb}^{-1} I$
Improve detection of P'_p in rectified views $I_p = H_{cp}^{-1} I$
Calibrate the system (cf. Section 2.1)

2.4 Defocusing error

For any projected circle, the center of its defocused projection is not equal to the projection of its center in the general case. This is what introduces the defocusing error.

Let $\mathbf{s}_o$ be a point source with center E_o on the projector plane. The projection of E_o on the calibration board is defined by the pinhole projector model where no defocus effect occurs. The relationship between points on the calibration board and points on the projector plane under pinhole model is governed by H_{bp}. So its projection on the calibration board is $H_{bp}(E_o)$. In the improvement step, projected dots centers are detected in the rectified view $I_p = H_{cp}^{-1} I$. Assume the projector works inversely, i.e. it captures world object and forms images on projector plane. Then, I_p can be regarded as an image "seen" by the pinhole inverse-projector. Let $\mathbf{s}_r$ be the spot on I_p with center E_r. $H_{cb}^{-1} H_{cp}(E_r) = H_{bp}(E_r)$ is then used as the world point for projector calibration. Thus, the error of detection due to projector's defocus on the calibration board can be expressed as $H_{bp}(E_r) - H_{bp}(E_o)$. If it is expressed on the projector plane, then $e = E_r - E_o$.

We show in the Appendix that the upper bound of this error can be roughly expressed as:

$$e_{max} = \frac{0.52 f d^2}{b^2} \approx 0.29 px \tag{4}$$

In our experiments, we used the following values: effective focal length $f \approx$ 2000px, diameter of the lens d=8.35mm, distance between the calibration board and the projector b =50cm.

Note that Eq. (4) can also be used to estimate the smallest admissible board distance for different projectors $b_{min} = d\sqrt{\frac{f}{2e_{max}}}$. This indicates the drawback of our method for large lens projectors.

3 CALIBRATION RESULTS

Our experimental system consists of a LogiTech C270 webcam (working resolution 640×480), an Epson EB-1771W projector (resolution 1920×1080, lens diameter 8.35mm), a laptop (Intel i7-4510U@2.00GHz CPU, 8GB of RAM) and calibration boards. We use our method to calibrate the system with the projector focusing at various distances, and compare the results with Audet et al. [1]. This method is chosen as a reference since it presents competitive results compared to other ones, source code is available online and it is easy to use.

Audet et al. [1] propose to use a B4 paper (250×353mm) attached to a rigid board as calibration pattern. In order to ease comparison, we use a calibration board of the same size with $2N = 200$ points. We found that with this size, using a 2mm radius for points with 16mm minimum inter-point distance works well. The radius of projected points is $r_o = 6$px.

Figure 4: Comparison of a $2 \times A0$ calibration board used by Audet's method [1] and a $B4$ calibration board used by our method (bottom-right) for large focus distances.

For the projector we use, a B4 board covers almost all the projection view only at about 50cm from the projector. For other focus distances, three different sizes of pattern (A2, A1, $2 \times A0$) are printed. Marker centers are used for Audet et al.'s method as they report to have a better result. For our method, we only use the B4 calibration board to demonstrate that it is much easier to use (cf. Figure 4). At each focus distance, we calibrate 5 times for each method and average the result. For each calibration, $K = 10$ images are captured, following [1]. Both methods call OpenCV's *calibrateCamera* method for final calibration, with the same set of options and parameters. To prevent projector's optical properties changing with temperature, the projector is pre-warmed 30 minutes before experiments.

We first show the results of calibration re-projection root mean square errors (RMSE) in Figure 5. RMSE is the most commonly used error measurement for camera and projector calibrations. Compared to Audet et al.'s method, our average RMSE is smaller but our maximum RMSE is larger. This is reasonable since the more correspondences a method uses, the more likely extreme values appear.

Both methods rely on Zhang [18] to find the projector's intrinsic matrix and distortions coefficients by minimizing the sum of re-projection errors of all correspondences. However, as pointed out by [6, 13], minimizing re-projection errors on calibration data cannot guarantee that the estimated internal parameters are the best ones. Ideally, when no modification is applied to the ProCam system, separated consecutive calibrations should lead to repeatability, i.e. stable intrinsic matrix and lens distortions for both devices.

Figure 6 shows the results on focal lengths estimation at different focus distances. Focal lengths estimated by our method are very stable, with $t(5) = 2.69$ for standard deviation (SD) of f_{px} and $t(5) = 2.49$ for SD of f_{py}, $p < 0.05$. One can clearly observe the effective focal length changes due to focus change. In Figure 7, our principal point positions seem more stable as well, but the difference is not significant in t-test ($t(5) = 1.35$ for SD of u and $t(5) = 1.92$ for SD of v, $p > 0.05$). For the camera, our method gives stable results $f_c = 812.4 \pm 0.9$ against Audet's $f_c = 808.1 \pm 4.2$, (f_{cx}, f_{cy} are averaged together).

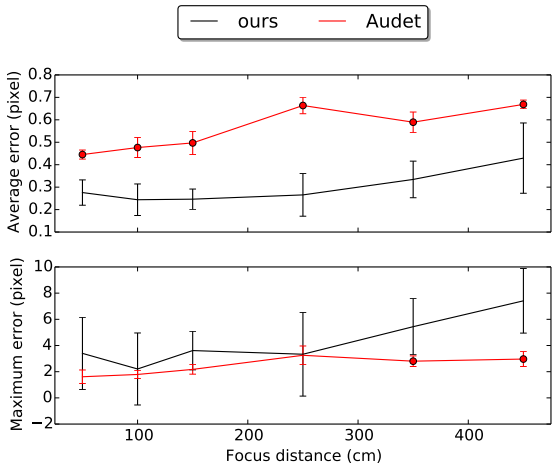

Figure 5: Average and maximum re-projection errors (RMSE): A red point on the curve indicates a significant difference between Audet's method and ours, at $p < 0.05$ with t-test.

To make a short summary, our method gives a smaller RMSE on calibration data, and more stable intrinsic estimates, despite using a much smaller calibration board.

4 AUGMENTATION EVALUATION

In the SAR community, people do not care about the true value of intrinsic matrix, nor about the RMSE of calibration data, since real augmentation will hardly lie on these calibration points. They care more about how precise some information can be projected in the focus zone of the projector. In this section, we use calibration-independent data to evaluate this effect.

When relying on Audet et al.'s method to augment information at large distances, the user may choose to calibrate either at a short focus distance with a small board (B4 size) with a loss of precision, or to calibrate at a larger distance with more burden (due to the use of potentially very big calibration boards). As mentioned before, our method offers the advantage to be able to calibrate the ProCam system with only a B4 size calibration board. In order to draw a fair comparison between our method and Audet et al.'s method, three different results are compared here: our method and Audet's calibrating at Correct Focus Distances with different size of boards (i.e. focus distances of the projector during calibration and evaluation are the same, denoted in the figures as Audet-CFD), and Audet method calibrating with a focus set at 50cm with a B4 board (denoted as Audet-50cm).

We use two asymmetric circle patterns (cf. Figure 8) to investigate projection errors in the focus zone of the projector. Q_b contains the coordinates of all black points while Q_w contains the coordinates of the white ones. According to the distance between the ProCam system and the calibration board, different sizes of these patterns are generated. Q_b are printed at exact positions and with known sizes so that we know their true physical positions on

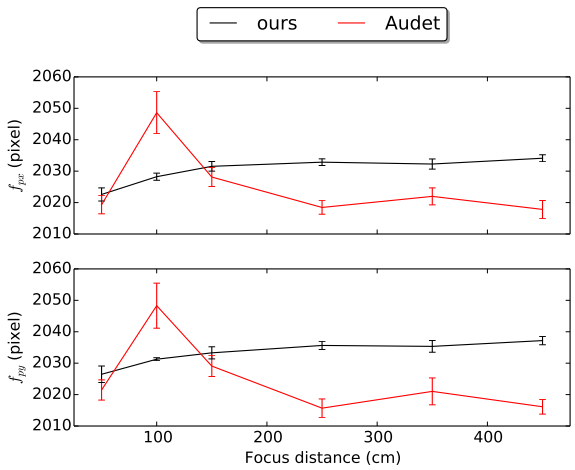

Figure 6: Focal length results: our method gives a significantly more stable estimation at $p < 0.05$ with t-test. It shows the trend of focal length variation.

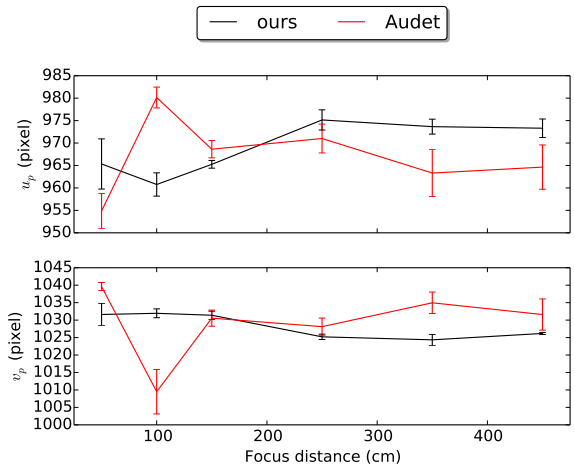

Figure 7: Principal point position results: there is no significant difference between our method and Audet's one, despite the size difference between calibration patterns.

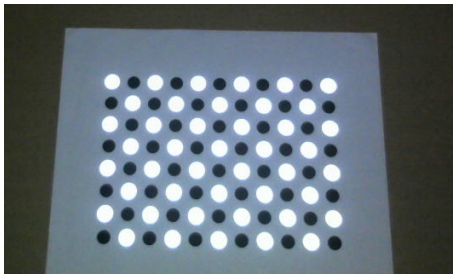

Figure 8: Evaluation pattern: (printed) black points are used for localization, (projected) white points are used to measure projection errors.

the paper. Q_w contains the ground-truth positions at which white points should be projected. They are drawn in an image $I(Q_p)$ to be warped and projected later. After projection, the real positions of the projected circles are measured by the camera and compared to their ground truth values to compute projection errors. A detailed version of this evaluation algorithm is presented in Algorithm 2.

Algorithm 2 Evaluation

Input: K_c, D_c, K_p, D_p, R, t to be evaluated, Q_b and Q_w
Output: List of re-projection errors $\{\vec{e}_r\}$
$I \leftarrow undistortCameraView(K_c, D_c)$
$Q'_b \leftarrow detectBlackDots(I)*$
$H_{cb} \leftarrow computeHomography(Q'_b, Q_b)$ with $Q'_b = H_{cb}(Q_b)$
$R_{cb}, t_{cb} \leftarrow computePatternPosition(K_c, H_{cb})$
$R_{pb}, t_{pb} \leftarrow coordinateTransformation(R, t, R_{cb}, t_{cb})$
$H_{pb} \leftarrow constructHomography(R_{pb}, t_{pb}, K_p)$
$I_{dp} \leftarrow projectDistortedImage(I(Q_p), D_p, H_{pb})$
$Q'_w \leftarrow detectWhiteDots(I_{dp})*$
$\{\vec{e}_r\} \leftarrow H_{cb}^{-1}(q'_w) - q_w$, for all $(q'_w, q_w) \in (Q'_w, Q_w)$
(*)Perspective and lens' distortions are removed from images. This allows us to find real circle centers.

Before showing the results, we need to address an issue: how precise can this evaluation algorithm be? The precision depends on both camera and projector intrinsics. Let us take an example with the camera (projector would follow the same reasoning): assuming a board parallel to the camera's image plane is positioned at a distance z from the camera's origin (cf. Figure 9). A segment of length L is measured by the camera as being l pixels long, so we have $l = f_c z^{-1} L$. Furthermore, if the angle between the board and the camera optical plane is φ, $l = f_c z^{-1} L \cos(\varphi)$.

The smaller the focal length is, the smaller l is. Assuming $l = 0.5$ pixel (used for illustrative purpose as an approximation of the MSER detector precision), the minimum *on board* difference we can measure is:

$$d_{min} = \frac{z}{2min(f_c, f_p)\cos\varphi} \tag{5}$$

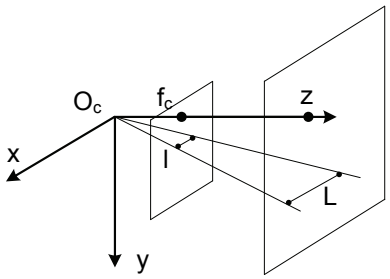

Figure 9: Projection of a segment: O_c is the camera's origin. World segment L is on a board parallel to camera image plane at z, its image l is on the image plane. We have $l/L = f_c/z$.

RMSE is calcuated from $\{\vec{e}_r\}$ in each evaluation image. Considering two series of RMSE $\mathbf{e}_1$ and $\mathbf{e}_2$, if the difference of the average value $\sum \mathbf{e}_1/n_1 - \sum \mathbf{e}_2/n_2$ is less than d_{min}, we consider it to be numerical noise. This effect is represented in Figures 10, 13, 14 as a gray region englobing our curve, which means that other RMSE inside this region shows no difference from our result due to the measurement limit. Otherwise, an independent significant test (t-test with different variations) is performed against the result of our method:

$$H_0 : \sum \mathbf{e}_1/n_1 = \sum \mathbf{e}_2/n_2 \quad (6)$$

We choose $p = 0.05$, which means the probability that H_0 is wrongly rejected is less than 5%. In our system, $f_c < f_p$, so we have $min(f_c, f_p) = f_c \approx 810$ (cf. Section 3).

4.1 Focus distance

We first evaluate RMS re-projection errors at different focus distances (cf. Figure 10). The circle pattern is placed in the center of the focus zone to have the least lens distortions. Even only calibrated at 50cm, Audet-50cm method works well for short focus distances ($<$ 250cm), but the RMSE grows rapidly for large focus distances. Except for 250cm, there is no significant performance difference between our method and Audet-CFD method although we have to remind the results of Audet-CFD's method are obtained with huge cumbersome calibration boards. For our method, the RMSE for a 450cm focus/projection distance is only about 4mm. To have a clear visual difference, we choose 450cm focus to show the re-projection difference (cf. Figure 11).

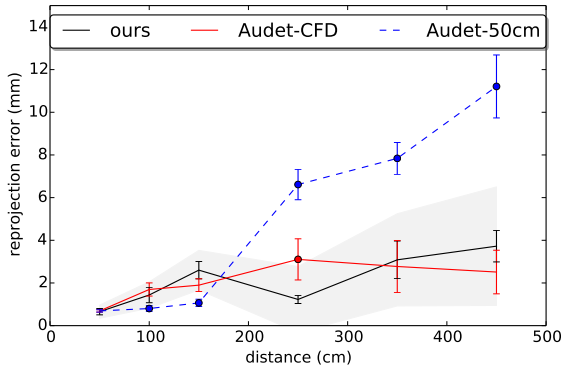

Figure 10: RMSE with different focus distances. Color dots mean that RMSE significantly differs from our result (H_0 is rejected).

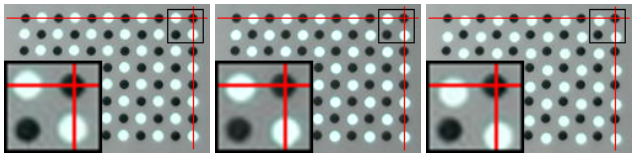

Figure 11: Re-projection error in rectified views (focusing at 450cm). From left to right: ours, Audet-CFD, Audet-50cm. Red lines are drawn to show points alignment. Circles diameter is 48mm.

4.2 Error distribution

Figure 12 shows error distribution in a front view at 250cm. The circle board is placed at four different places to cover the whole projector view. Both our method and Audet-CFD method have small errors while Audet-50cm gives large errors especially for small y values. We can notice that Audet-CFD generally has smaller errors than ours in Figure 12 while our method shows smaller RMSE at 250cm in Figure 10. However, this is not contradictory, since the error difference reaches the limits of the measurement system.

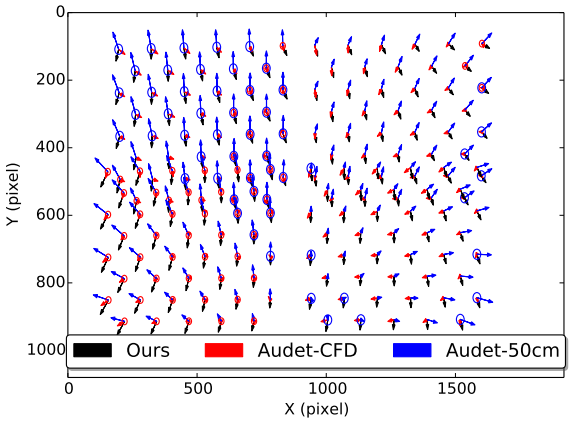

Figure 12: Error distribution in projector's view (focusing at 250cm). Re-projection errors (instead of RMSE) are used in the t-test: Color circles indicate significant differences from our result.

4.3 Perspective and depth

At last, we study the influence of perspective angle and depth in the projector's focus zone. The projector is always focusing at 250cm. Figure 13 shows that all methods perform worse with large oblique angles. Audet-50cm method is significantly worse in many cases while our method and Audet-CFD give almost the same result, yet again at the expense of the size of the calibration board. Figure 14 shows that when further away from the projector, Audet-50cm method has a poor performance.

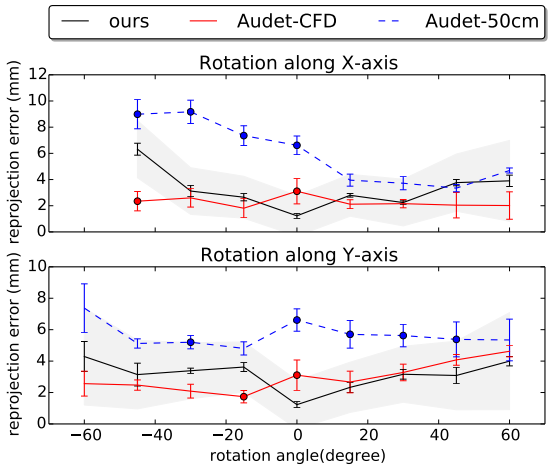

Figure 13: RMSE with various rotations (focusing at 250cm). Color dots mean RMSE significantly differs from our result (H_0 is rejected).

5 CONCLUSION

We have presented a practical method to calibrate a projector-camera system. The method is projector focus independent, user-friendly and interacts with the user in real-time. We evaluated the calibration results with calibration-independent data. The re-projection error of the system is competitive with regard to the state-of-art, but the method only demands a small calibration board

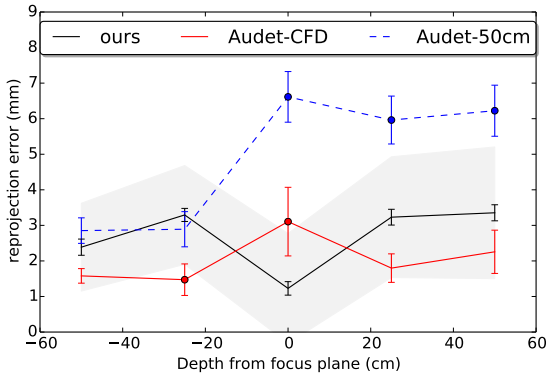

Figure 14: RMSE with various depths (focusing at 250cm). Color dots mean RMSE significantly differs from our result. (H_0 is rejected).

and simple manipulations even at large focus distances (more than 2.5m). Results also show that the calibrated intrinsic matrix of the projector is more stable than other state-of-the-art methods [1]. The drawback of our method is that it potentially has larger system error for projector with large lens. Nevertheless this can be solved by calibrating the system at a slightly bigger distance, which is still closer than state-of-the-art methods [1].

APPENDIX

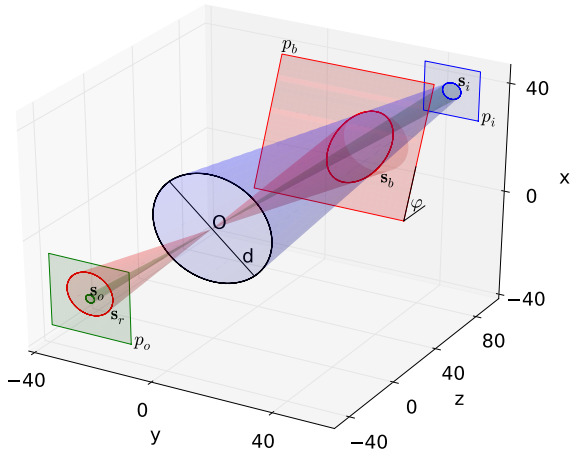

Figure 15: Projection geometry for defocusing error: O is projector optical center, p_o is the projector plane, p_i is the plane of focus, p_b is the board plane. $\mathbf{s}_b$ is the defocused spot on the board, $\mathbf{s}_r$ is the defocused spot "seen" by an inverse pinhole projector.

In this section, we compute e from Eq. (4) in a general 3D case.

In Figure 15, the projector lens is located at O with diameter d. A circle $\mathbf{s}_o(x_o, y_o, -u; r_o)$ is located on the projector's plane $p_o(z = -u)$. It is projected and forms a clear image $\mathbf{s}_i(x_i, y_i, v; r_i)$ on the plane of focus $p_i(z = v)$. All light coming from $\mathbf{s}_o$ and transmitted by the lens forms a blue cone C_b, which follows a standard lens projection model.

C_b intersects the board plane p_b and forms the light spot $\mathbf{s}_b$ on the board. The ellipse $\mathbf{s}_b$ is "seen" by the inverse-projector (cf. Section 2.4) with pinhole model (represented by a red elliptic cone C_r), and forms $\mathbf{s}_r$ on the rectified view I_p, which coincides with the projector plane. As discussed in Section 2.4, the error e in the projector's plane is the difference between $\mathbf{s}_o$ and $\mathbf{s}_b$.

To derive the error e, the coordinate system is constructed so that the angle between p_b and the projector's z-axis is φ. Then the equations of C_b and the board plane p_b are:

$$C_b : \begin{cases} x(t,\alpha) &= \frac{x_o}{z_o} t + \frac{d}{2}\left[1-(d-2r_i)\frac{t}{dv}\right]\cos\alpha \\ y(t,\alpha) &= \frac{y_o}{z_o} t + \frac{d}{2}\left[1-(d-2r_i)\frac{t}{dv}\right]\sin\alpha \\ z(t,\alpha) &= t \end{cases} \tag{7}$$

$$l_b : z - x\cot\varphi - b = 0 \tag{8}$$

where t, α are two parameters for C_b.

By using the fact that $\mathbf{s}_b$ is the intersection of C_b and l_b; both $\mathbf{s}_b$ and $\mathbf{s}_r$ lines on the surface of C_r; C_r is an elliptic cone passing through O; and $\mathbf{s}_r$ is the intersection between C_r and p_o. We can prove that $\mathbf{s}_r$ is an ellipse centered at (x_r, y_r), with:

$$\begin{cases} x_r &= x_o + \frac{ud^2\cot\varphi}{4b^2 - d^2\cot^2\varphi}\left(1 - \frac{x_o}{y_o}\cot\varphi - \frac{b}{v} + \frac{2br_o}{ud}\right) \\ y_r &= y_o \end{cases} \tag{9}$$

Thus the error e is

$$e = \frac{ud^2\cot\varphi}{4b^2 - d^2\cot^2\varphi}\left(1 - \cot\varphi\frac{x_o}{u} - \frac{b}{v} + \frac{2br_o}{ud}\right) \tag{10}$$

where u is the object distance; v is the focus distance; d is the lens diameter; φ is the angle between z-axis and the board (cf. Fig. 15), usually in $[45°, 135°]$ for calibration; p_b intersects z-axis at $(0, 0, b)$; x_o measures the distance between $\mathbf{s}_o$'s center and the principle point in the projector plane along x-axis.

We can see from Eq. (10) that, if the board is vertical (i.e. $\varphi = 90°$), or with pinhole projector (i.e. $d = 0$), the error will be 0. This corresponds to our intuitive knowledge.

In our experiment, the projector's width $W = 1920$ px, its height $H = 1080$ px, the effective focal length (i.e. object plane distance) $f = u \approx 2000$ px, $b = 50$cm, $r_o = 6$ px. The maximum error e_{max} is obtained when $v \to \infty$, $\varphi = \varphi_{max} = 135°$ and $x_o = \sqrt{W^2/4 + H^2}$. It can be approximated by:

$$e_{max} \approx \frac{0.52fd^2}{b^2} \approx 0.29px \tag{11}$$

REFERENCES

[1] S. Audet and M. Okutomi. A user-friendly method to geometrically calibrate projector-camera systems. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 47–54, 2009.

[2] O. Bimber and R. Raskar. *Spatial Augmented Reality: Merging Real and Virtual Worlds*. CRC Press, 2005.

[3] G. Bradski and A. Kaehler. *Learning OpenCV*. O'Reilly, 2008.

[4] A. Datta, J.-S. Kim, and T. Kanade. Accurate Camera Calibration using Iterative Refinement of Control Points. In *IEEE 12th International Conference on Computer Vision Workshops*, pages 1201–1208, 2009.

[5] W. Gao, L. Wang, and Z. Hu. Flexible method for structured light system calibration. *Optical Engineering*, 47(8):083602–083602, 2008.

[6] J. I. González, J. C. Gámez, C. G. Artal, and A. N. Cabrera. Stability Study of Camera Calibration Methods. In *CI Workshop en Agentes Físicos*, WAF, 2005.

[7] B. Li, N. Karpinsky, and S. Zhang. Novel calibration method for structured-light system with an out-of-focus projector. *Applied optics*, 53(16):3415–3426, 2014.

[8] Z. Li, Y. Shi, C. Wang, and Y. Wang. Accurate calibration method for a structured light system. *Optical Engineering*, 47(5):053604–053604, 2008.

[9] H. Luo, J. Xu, N. H. Binh, S. Liu, C. Zhang, and K. Chen. A simple calibration procedure for structured light system. *Optics and Lasers in Engineering*, 57:6–12, 2014.

[10] D. Nistér and H. Stewénius. Linear Time Maximally Stable Extremal Regions. In *10th European Conference on Computer Vision*, ECCV, pages 183–196, 2008.

[11] J.-N. Ouellet, F. Rochette, and P. Hébert. Geometric Calibration of a Structured Light System using Circular Control Points. In *3D Data Processing, Visualization and Transmission*, 3DPVT, pages 183–190, 2008.

[12] 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 Transactions on Visualization and Computer Graphics*, 21(11):1211–1220, 2015.

[13] S.-W. Shih, Y.-P. Hung, and W.-S. Lin. Accuracy Analysis on the Estimation of Camera Parameters for Active Vision Systems. In *13th International Conference on Pattern Recognition*, volume 1 of *ICPR*, pages 930–935, 1996.

[14] K. Yamauchi, H. Saito, and Y. Sato. Calibration of a Structured Light System by Observing Planar Object from Unknown Viewpoints. In *19th International Conference on Pattern Recognition*, ICPR, pages 1–4, 2008.

[15] S. Yamazaki, M. Mochimaru, and T. Kanade. Simultaneous self-calibration of a projector and a camera using structured light. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 60–67, 2011.

[16] L. Yang, J.-M. Normand, and G. Moreau. Local Geometric Consensus: A General Purpose Point Pattern-Based Tracking Algorithm. *IEEE Transactions on Visualization and Computer Graphics*, 21(11):1299–1308, 2015.

[17] S. Zhang and P. S. Huang. Novel method for structured light system calibration. *Optical Engineering*, 45(8):083601–083601, 2006.

[18] Z. Zhang. A Flexible New Technique for Camera Calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.