Geometric Video Projector Auto-Calibration

Anonymous CVPR submission

Paper ID 610

Abstract

In this paper we address the problem of geometric calibration of video projectors. Like in most previous methods we also use a camera that observes the projection on a planar surface. Contrary to those previous methods, we neither require the camera to be calibrated nor the presence of a calibration grid or other metric information about the scene. We thus speak of geometric auto-calibration of projectors (GAP). The fact that camera calibration is not needed increases the usability of the method and at the same time eliminates one potential source of inaccuracy, since errors in the camera calibration would otherwise inevitably propagate through to the projector calibration. Our method enjoys a good stability and gives good results when compared against existing methods as depicted by our experiments.

1. Introduction

With the recent advances in projection display, video projectors are becoming the devices of choice for active reconstruction systems and 3D measurement. Such systems like Structured Light [**?**] and also Photometric Stereo [**?**, **?**] use video projectors to alleviate the difficult task of establishing point correspondences. However, even if active systems can solve the matching problem, calibrated video projectors are still required. In fact, a calibrated projector is required to triangulate points in a camera–projector structured light system, or to estimate the projector's orientation when the latter is used as an illumination device for a photometric stereo system.

The projection carried out by a video projector is usually modeled as the inverse projection of a pin-hole camera, and thus considered as a perspective projection.

In order to simplify the calibration process, a planar surface is often used as projection surface, onto which features or codified patterns are projected. The way patterns are codified and the projection surface orientation is estimated distinguishes most previous calibration methods from one another.

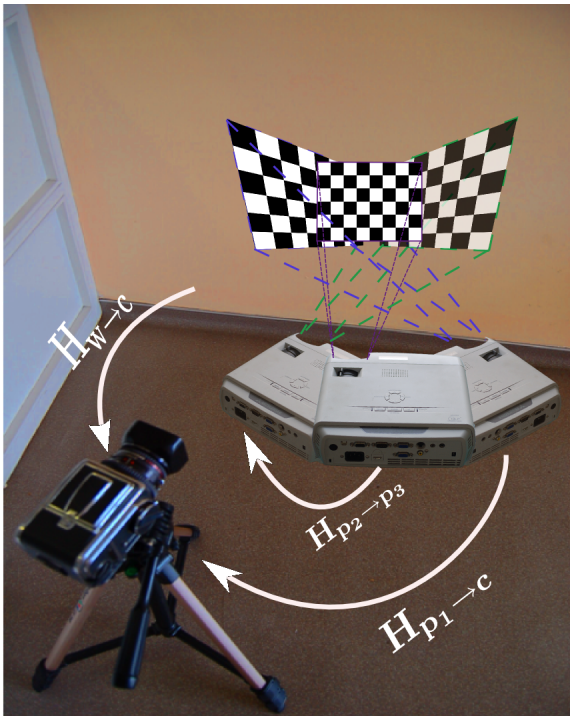

Figure 1. A Camera-Projector setup and its homographies (see text).

In [**?**, **?**], a video projector projects patterns on a plane mounted on a mechanically controlled platform. Thus, the orientation and position of the projection plane is known and is used to calibrate the structured light system using conventional camera calibration techniques.

For convenience and because the projection surface is usually planar, we will also refer to it as the *wall*.

In [**?**], a planar calibration grid is attached to the wall and observed by a calibrated camera. Due to the camera's calibration information and the metric information about the grid, the grid's and thus the wall's orientation and distance relative to the camera can be computed by classical pose estimation. After this, the 3D positions of features projected onto the wall by the video projector, can be easily

computed. If this is done for three or more positions of the video projector, a set of correspondences between the wall and the "projector images" can be obtained and then used to estimate the projector parameters with standard plane-based calibration methods [**?**, **?**]. We refer to this method as Direct Linear Calibration (DLC). Note that all this could actually be done without pre-calibrating the camera, purely based on plane homographies, as explained in section **??**. Further, to increase accuracy of the DLC, a printed planar target with circular markers is used in [**?**], to calibrate the camera as well as the projector.

In [**?**], a structured light system is calibrated without using a camera. This is made possible by embedding light sensors in the target surface (the wall). Gray-coded binary patterns are then projected to estimate the sensor locations and prewarp the image to accurately fit the physical features of the projection surface. The projector parameters are not explicitly estimated but the method could easily be extended for that purpose.

Kimura *et al*. [**?**] proposed a calibration method based on the camera-projector epipolar geometry. Again, the camera must be fully calibrated.

In this paper, a new projector calibration method is introduced. As opposed to most existing methods, the proposed method does not require a physical calibration grid nor any knowledge about the camera parameters. Indeed, our method imposes only two constraints on the calibration setup. Namely, the camera should remain static while the video projector displays patterns onto a planar surface and the user must put the projector once in a roughly fronto-parallel position relative to the wall. The latter constraint does not have to be exact and serves only as a starting point for a non-linear minimization as explained below.

The rest of the paper is organized as follows. In section **??**, our model for the geometric transformation associated with the video projector, is described. In section **??**, we explain the above mentioned DLC (direction linear calibration) approach, which serves as an introduction to the proposed auto-calibration method, described in section **??**. Experimental results are presented in section **??** and conclusions are drawn in section **??**.

2. Projector Model

Throughout this paper, the projector is assumed to have a perspective projection model like a pin-hole camera, with the slight difference that here the projection direction is reversed [**?**]. Based on this assumption, a 3D point $P = [X, Y, Z, 1]^{\mathsf{T}}$ is mapped to $p_p = [x, y, 1]^{\mathsf{T}}$ in the projector as:

$$p_p \sim \mathsf{K}_p \begin{pmatrix} \mathsf{R}_p & \mathbf{t}_p \end{pmatrix} P \tag{1}$$

where $\sim$ stands for equality up to scale between homogeneous coordinates. These 2D points p_p live in what we refer to by the "projector image".

The matrix R_p and the vector $\mathbf{t}_p$ represent the extrinsic parameters of the projector. The calibration matrix K_p is described by the sought internal parameters and is defined as follows:

$$\mathsf{K}_p = \begin{pmatrix} \rho f & 0 & u \\ 0 & f & v \\ 0 & 0 & 1 \end{pmatrix} \tag{2}$$

where f, ρ and (u, v) are respectively the focal length, the aspect ratio and the principal point coordinates.

Consider a camera imaging what is projected by the projector onto the wall. Since we assume the wall to be planar, it induces an homography $\mathsf{H}_{p\to c}$ between the projector and the camera image. Without loss of generality, we may assume that the world coordinate system is aligned with the wall, such that points on the wall have coordinates $Z = 0$. Then, the homography between projector and camera can be written as:

$$\mathsf{H}_{p\to c} \sim \underbrace{\mathsf{K}_c \begin{pmatrix} \bar{\mathsf{R}}_c & \mathbf{t}_c \end{pmatrix}}_{\mathsf{H}_{w\to c}} \underbrace{\left(\mathsf{K}_p \begin{pmatrix} \bar{\mathsf{R}}_p & \mathbf{t}_p \end{pmatrix}\right)^{-1}}_{\mathsf{H}_{p\to w}} \tag{3}$$

where $\bar{\mathsf{A}}$ refers to the first two columns of a 3×3 matrix A. K_c is the camera's calibration matrix and R_c and $\mathbf{t}_c$ represent its extrinsic parameters. The homography $\mathsf{H}_{p\to c}$ can also be seen as the product of the homography $\mathsf{H}_{p\to w}$ that maps the projector image plane to the wall with $\mathsf{H}_{w\to c}$, the homography that relates the wall to the camera image plane.

3. Direct Linear Calibration

In this section, we review the details of the Direct Linear Calibration for projectors. This method is used as a reference for our experiments. As opposed to [**?**], the variant presented here [**?**] is strictly based on homographies and does not require a calibrated camera.

A planar calibration grid is attached to the wall. This allows to estimate the homography $\mathsf{H}_{w\to c}$ between the wall and the camera, introduced above. It relates a point p_w on the wall to a point p_c in the camera image as follows:

$$p_c \sim \mathsf{H}_{w\to c} p_w \tag{4}$$

Once this homography is computed (details on homography estimation can be found in [**?**]), the video projector is used to project patterns while it is moved to various positions and orientations. For each projector pose i, correspondences are established between the camera and the video projector, leading to an homography $\mathsf{H}_{c\to p_i}$. A point p_c in the camera image is mapped into the projector at pose i as:

$$p_p^i \sim \mathsf{H}_{c\to p_i} p_c \tag{5}$$

Combining (**??**) and (**??**), a point p_w on the wall is mapped into the i^{th} projector as:

$$p_p^i \sim \underbrace{\mathsf{H}_{c\to p_i}\mathsf{H}_{w\to c}}_{\mathsf{H}_{w\to p_i}} p_w \tag{6}$$

We thus can compute the wall-to-projector homography for each pose i. It has the following form (see above):

$$\mathsf{H}_{w\to p_i} \sim \mathsf{K}_p \begin{pmatrix} \bar{\mathsf{R}}_p^i & \mathbf{t}_p^i \end{pmatrix} \tag{7}$$

It is now straightforward to apply classical plane-based calibration methods [**?**, **?**] to calibrate the projector and, if necessary, to compute its extrinsic parameters, from two or more poses.

4. Projector Auto-Calibration

4.1. Basic Idea

The approach described in the previous section requires a calibration grid to be attached to the wall and, in the version of [**?**], the camera to be calibrated. In this section, we show that these requirements may be avoided and propose a true geometric video projector auto-calibration approach.

The key observation underlying the auto-calibration approach is as follows. It is "easy" to compute homographies between the projector image and the camera image, induced by the projection surface. There are indeed many possibilities to do so, the simplest ones consisting in projecting a single pattern such as a checkerboard and extracting and identifying corners in the camera image. More involved ones could make use of multiple patterns, sequentially projected from each considered projector pose, such as Gray codes, allowing for robust and dense matching. From the obtained matches, the computation of the homography is straightforward.

Consider now homographies associated with two poses of the projector, $\mathsf{H}_{c\to p_i}$ and $\mathsf{H}_{c\to p_j}$. From these we can compute an homography between the two projector images, induced by the planar projection surface:

$$\begin{aligned} \mathsf{H}_{p_i\to p_j} &\sim \mathsf{H}_{w\to p_j}\mathsf{H}_{w\to p_i}^{-1} \\ &\sim \mathsf{H}_{c\to p_j}\mathsf{H}_{w\to c}\left(\mathsf{H}_{c\to p_i}\mathsf{H}_{w\to c}\right)^{-1} \\ &\sim \mathsf{H}_{c\to p_j}\mathsf{H}_{c\to p_i}^{-1} \end{aligned}$$

We are now in the exact same situation as an uncalibrated perspective camera taking images of an unknown planar scene: from point matches, the associated plane homographies can be computed and it is well-known that camera auto-calibration is possible from these, as first shown by Triggs [**?**]. We may thus apply any existing plane-based auto-calibration method, e.g. [**?**, **?**, **?**] to calibrate the projector. Compared to auto-calibration of cameras, the case of projectors has an advantage; many and highly accurate point matches can be obtained since the scene texture is controlled, by projecting adequate patterns onto the wall.

Plane-based auto-calibration comes down to a non-linear optimization problem, even in the simplest case when only the focal length is unknown. To avoid convergence problems, we adopt an approach suggested in [**?**] that requires to take one image in a roughly fronto-parallel position relative to the scene plane. Here, this means of course by analogy that the projector should once be positioned in a roughly fronto-parallel position relative to the wall; subsequent poses can (and should) then be different. This allows for a closed-form initial solution to the auto-calibration problem, which may then be refined by a non-linear optimization (bundle adjustment). Note that the assumption of fronto-parallelism for one of the images is only required for the initialization; during optimization, this is then no longer enforced.

4.2. Initialization Procedure

We derive the initialization procedure in a different and simpler way compared to [**?**]. Let the fronto-parallel view correspond to pose 1; in the following we only consider homographies between that view and all the others. Consider first the wall-to-projector homography of the fronto-parallel view, $\mathsf{H}_{w\to p_1}$. So far, we have assumed that the world coordinate system is such that the wall is the plane $Z = 0$ (see section **??**). Without loss of generality, we may assume that the X and Y axes are aligned with those of the fronto-parallel view and that the optical center of that view is located at a distance equal to 1 from the wall. Note that these assumptions are not required to obtain the below results, but they simply make the formulae simpler. With these assumptions, the wall-to-projector homography for the fronto-parallel pose is simply:

$$\mathsf{H}_{w\to p_1} \sim \mathsf{K}_p$$

Consider now the homography between the fronto-parallel view and another view j:

$$\begin{aligned} \mathsf{H}_{p_1\to p_j} &\sim \mathsf{H}_{w\to p_j}\mathsf{H}_{w\to p_1}^{-1} \\ &\sim \mathsf{K}_p \begin{pmatrix} \bar{\mathsf{R}}_p^j & \mathbf{t}_p^j \end{pmatrix} \mathsf{K}_p^{-1} \end{aligned}$$

In the following let us, for simplicity, drop all indices:

$$\mathsf{H} \sim \mathsf{K} \begin{pmatrix} \bar{\mathsf{R}} & \mathbf{t} \end{pmatrix} \mathsf{K}^{-1}$$

It follows that:

$$\mathsf{K}^{-1}\mathsf{H} \sim \begin{pmatrix} \bar{\mathsf{R}} & \mathbf{t} \end{pmatrix} \mathsf{K}^{-1}$$

Let us now multiple each side of the equation from the left with its own transpose:

$$\mathsf{H}^T\mathsf{K}^{-T}\mathsf{K}^{-1}\mathsf{H} \sim \mathsf{K}^{-T}\begin{pmatrix}\bar{\mathsf{R}} & \mathbf{t}\end{pmatrix}^T\begin{pmatrix}\bar{\mathsf{R}} & \mathbf{t}\end{pmatrix}\mathsf{K}^{-1}$$

Since $\bar{\mathsf{R}}$ consists of the first two columns of the rotation matrix R, we have $\bar{\mathsf{R}}^T\bar{\mathsf{R}} = \mathtt{I}$ and thus:

$$\mathsf{H}^T\mathsf{K}^{-T}\mathsf{K}^{-1}\mathsf{H} \sim \mathsf{K}^{-T}\begin{pmatrix}1 & 0 & \times\\ 0 & 1 & \times\\ \times & \times & \times\end{pmatrix}\mathsf{K}^{-1}$$

where entries marked as $\times$ depend on $\mathbf{t}$ and are irrelevant for the following. Due to the form of K, this becomes:

$$\mathsf{H}^T\mathsf{K}^{-T}\mathsf{K}^{-1}\mathsf{H} \sim \begin{pmatrix}1 & 0 & \times\\ 0 & \rho^2 & \times\\ \times & \times & \times\end{pmatrix} \tag{8}$$

Let us use the image of the absolute conic (IAC) to parameterize the projector's intrinsic parameters, defined as $\omega \sim \mathsf{K}^{-T}\mathsf{K}^{-1}$. From (**??**) we can now deduce the following two equations on the intrinsic parameters, which are similar to those of calibration based on a planar calibration grid [**?**, **?**]:

$$h_1^T\omega h_2 = 0 \tag{9}$$
$$\rho^2 h_1^T\omega h_1 - h_2^T\omega h_2 = 0 \tag{10}$$

where h_k denotes the kth column of H. Let us note that $\rho^2 = \omega_{11}/\omega_{22}$; hence, equation (**??**) can be written:

$$\omega_{11}h_1^T\omega h_1 - \omega_{22}h_2^T\omega h_2 = 0 \tag{11}$$

Equation (**??**) is linear in ω, whereas (**??**) is quadratic. There are different ways of using these equations to compute the IAC ω and from this, the intrinsic parameters. If the aspect ratio ρ is known beforehand, both equations are linear and thus easy to solve. If ρ is unknown, one can either use only the linear equation (**??**), which requires five views (the fronto-parallel one and four others), or compute ω from three views only. In the latter case, we have two linear and two quadratic equations and a "closed-form" solution in the form of a degree-4 polynomial in one of the unknowns, is straightforward to obtain.

4.3. Non-linear Optimization

Once an initial solution of the projector calibration is computed using the above approach, a non-linear optimization through bundle adjustment may be carried out. Let us briefly outline its peculiarities, compared to plane-based auto-calibration of a camera. Note that the only noisy observations in our scenario are features in the camera image: those in the projector "images" are perfectly known and noisefree! Hence, the cost function of the bundle adjustment should be based on the reprojection error in the camera image. The following formulation is one possible option:

$$\min_{\mathsf{H}_{w\to c},\mathsf{K}_p,\mathsf{R}_p^i,\mathbf{t}_p^i}\sum_{i,j} dist^2\left(p_c^{ij}, \mathsf{H}_{w\to c}\mathsf{H}_{p_i\to w}p_p^{ij}\right)$$

where i stands for projector poses and j for points. I.e. we optimize the wall-to-camera homography, the intrinsic projector parameters and its extrinsic parameters for all views, by minimizing the reprojection error when mapping from the projector images into the camera image (the $\mathsf{H}_{p_i\to w}$ are parameterized by K_p and the extrinsic projector parameters).

Another option would be to include camera intrinsics and extrinsics in the optimization instead of the "black-box" homography $\mathsf{H}_{w\to c}$, but since the camera is static in our case, at most two intrinsics can be estimated [**?**, **?**].

4.4. Estimation of Focal Length Changes

The above paragraphs constitute our auto-calibration approach. Here, we describe another method that allows to estimate the change of the projector's intrinsics caused by zooming. If the projector has been calibrated beforehand, this allows to update its calibration. We suppose that a zoom causes, besides the focal length, also the principal point to change (especially its vertical coordinates is likely to change in practice), but that the aspect ratio ρ remains constant.

We also suppose here that both the camera and the projector remain static. Let H be the projector-to-camera homography before zooming and H' the one afterwards. The inter-image homography between the two projector images is then given by:

$$\begin{aligned}\mathsf{M} &\sim (\mathsf{H}')^{-1}\mathsf{H}\\ &\sim \mathsf{K}'_p(\mathsf{K}_p)^{-1}\\ &\sim \begin{pmatrix}f' & 0 & u'f - uf'\\ 0 & f' & v'f - vf'\\ 0 & 0 & f\end{pmatrix}\end{aligned}$$

It is straightforward to compute the intrinsic parameters after zooming:

$$\begin{aligned}f' &= \frac{M_{11}}{M_{33}}f\\ u' &= \frac{M_{13} + uM_{11}}{M_{33}}\\ v' &= \frac{M_{23} + vM_{11}}{M_{33}}\end{aligned}$$

Note that M depends only on the three unknown intrinsic in K'_p and can thus be computed from two points matches

already. If the principal point can be assumed to remain constant, a single match is sufficient. A single match is also sufficient if only one coordinate of the principal point is supposed to change due to zooming (which is often the case for video projectors).

5. Experiments

The proposed algorithm has been tested on synthetic and real data. Both tests are detailed in the next two subsections.

5.1. Synthetic Data

We performed several tests of our algorithm using synthetic data to assess its sensitivity to noise, number of projector poses and fronto-parallelism inaccuracy. Throughout all the synthetic experiments, we used a camera panned at 30 degrees w.r.t the projection surface. The camera resolution was set to 1000×1000 and its calibration matrix defined as:

$$\mathsf{K}_c = \begin{pmatrix} 1000 & 0 & 500 \\ 0 & 1000 & 500 \\ 0 & 0 & 1 \end{pmatrix} \qquad (12)$$

The projector parameters are identical to the camera parameters.

Sensitivity to noise level. For this test, we used 20 inter-image homographies computed by orienting the projector at random. The range of the orientations was ±20 degrees w.r.t the projection surface. Projector points were then imaged by the camera, and a gaussian noise with mean 0 and increasing standard deviation was added to the image points. The standard deviation σ varied from 0.1 to 1.5. As in [**?**], we performed 100 independent runs for each noise level and computed the average errors for both the focal length and the principal point. As we can see from Fig. **??** and Fig. **??** the error increases almost linearly for both the focal length and the principal point. For a noise level of $\sigma = 0.5$ the error in the focal length is about 0.6% and the error in the coordinates of the principal point is less than 3 pixels which represents, or less than 0.7% relative error.

Sensitivity to the number of projector poses. We set the amount of noise to $\sigma = 1$ and we varied the number of projector poses from 2 to 20 in a range of ±20 degrees w.r.t the projection surface. The average errors (from 100 independent runs) for both the focal length and the principal point are reported in Fig. **??** and Fig. **??** . We notice that, as may be expected, the results gain stability when the number of projector poses is increased.

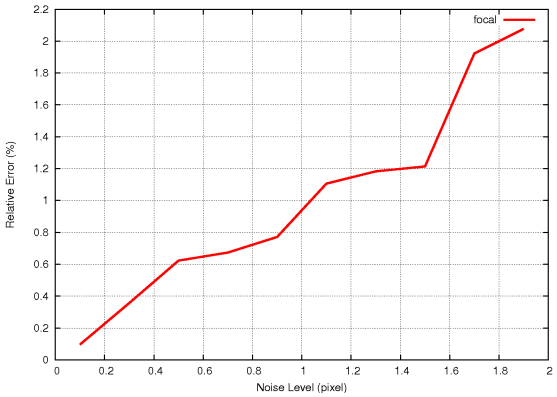

Figure 2. Focal length error vs. noise level

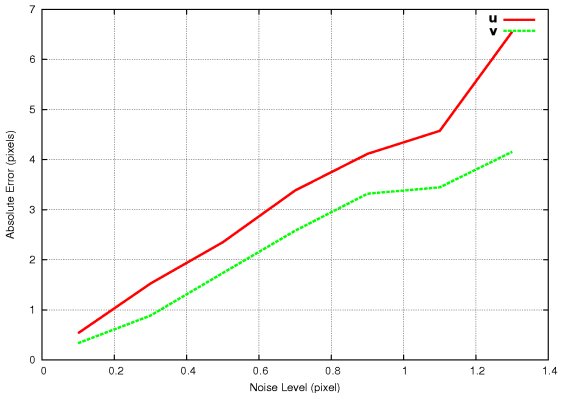

Figure 3. Principal point error vs. noise level

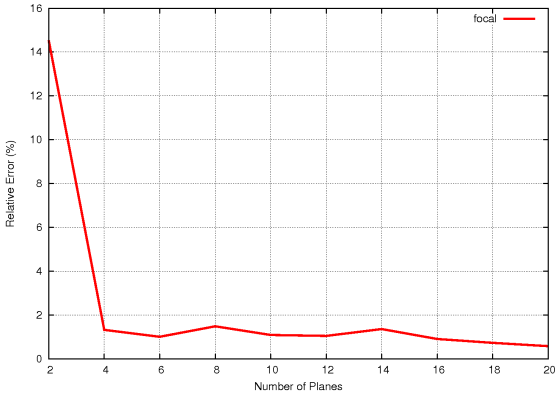

Figure 4. Focal length error vs. nb poses ($\sigma = 1$).

Sensitivity to fronto-parallelism inaccuracy. We conclude these synthetic experiments by assessing the sensitivity of our algorithm to the fronto-parallelism assumed in one of the images. The standard deviation of the noise added to the point coordinates was 0.5. We altered the orientation of the projector fronto-parallel to the projection surface. The resulting errors on the focal length and the principal point are reported in Fig. **??** and Fig. **??**

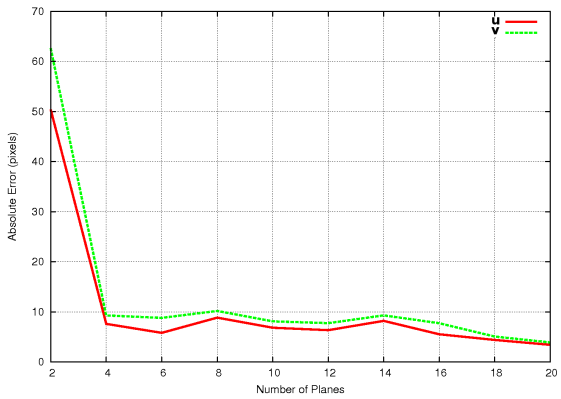

Figure 5. Principal point errors vs. nb poses ($\sigma = 1$).

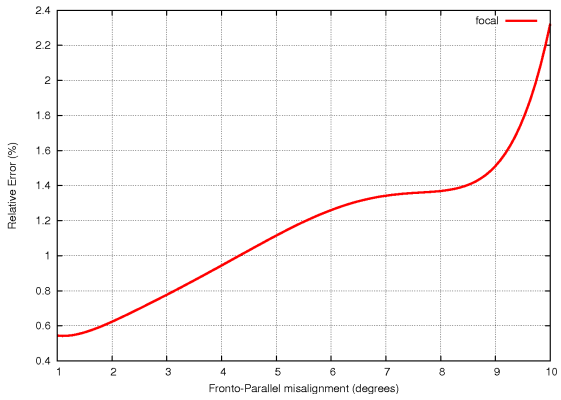

Figure 6. Focal length error vs. fronto-parallel misalignment.

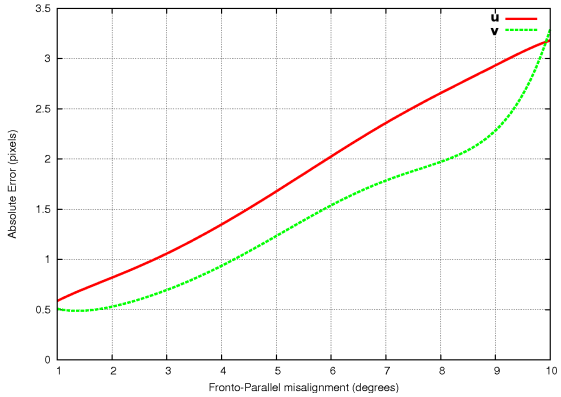

Figure 7. Principal point error vs. fronto-parallel misalignment.

5.2. Real Images

We tested our algorithm on a Mitsubishi Pocket Projector and compared it to our variant of the DLC method, described in section **??**. The projector has a native resolution of 800×600 and a fixed focal length. The acquisition device was a Nikon D50 camera. A $50mm$ lens was used on the camera and the resolution was set to 1500×1000.

We acquired 20 images of projected patterns while the projector underwent several orientations. Some images of the projected chessboard along with detected features are depicted on Figure.**??**.

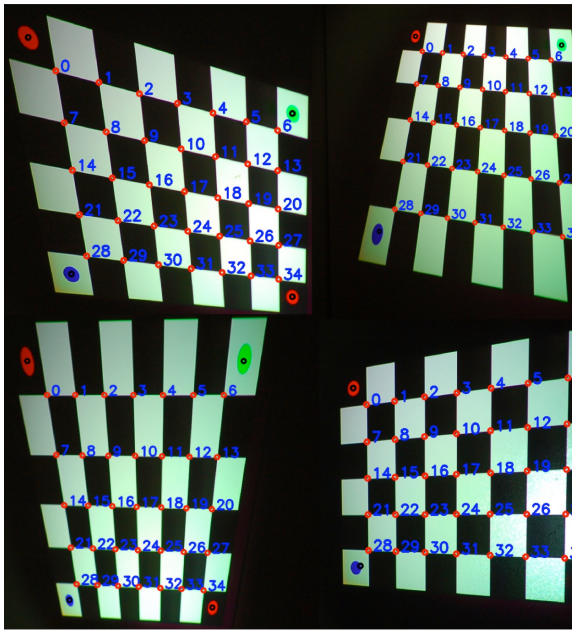

Figure 8. Images of projected patterns and detected features. The numbers and small red dots are added for illustration only. The large dots in the 4 corners are part of the projected pattern.

We calibrated the projector with the proposed method and with our implementation of the DLC. The result of this benchmark is outlined in Table **??**.

The table provides the estimated parameters and the reprojection error in pixels. Because our method was initialized with several fronto-parallel images we reported the range of reprojection error instead of an error average.

Table 1. Projector calibration benchmark: Direct method and the proposed Auto-Calibration method.

Method	$\mathbf{f_{proj}}$	ρ	**u**	**v**	**Error**
DLC	1320.13	1.002	382.1	448	0.46
Auto-Calib	1312.27	1.007	370.28	466	$0.42 - 0.27$

We performed a second calibration test on a video projector (Mitsubishi XD430U) with a zooming capability and a native resolution of 1024×768. For this test, we estimated the intrinsic parameters with two different zoom settings and the results were compared to the predictions obtained using the method introduced in section **??**.

We observed that both methods are consistent as reported

in Table **??**.

Table 2. Calibration results with varying parameters.

Method	$\mathbf{f_{proj}}$	ρ	**u**	**v**
Zoom 1	2292.29	1.045	584.42	969.36
Zoom 2 (pred)	1885.7	1.045	587.64	949.55
Zoom 2 (est)	1873.14	1.045	590.9	944

6. Conclusion

In this paper we presented a new video projector auto-calibration method. It does not require a physical calibration grid or other metric information on the scene. Also, the camera used together with the projector, does not need to be calibrated; it is indeed merely used to get plane homographies between "images" of the projector associated with different poses. We believe that this aspect of our method increases its stability, otherwise the error of the camera calibration would affect the accuracy of the projector calibration [**?**]. Of course, as usual with auto-calibration methods, a certain number of poses, and especially a sufficient variety of poses (especially orientation), are required to get good results. In our synthetic experiments, results are very good with 4 poses or more.

Very simple to implement, the proposed method is fast, gives good results and is completely linear if one uses common assumptions regarding the projector aspect ratio. In the near future we will implement and test the bundle adjustment procedure outlined in the paper. This is straightforward and is expected to further improve our results.

More generally, we believe that our method will enable to handle large projector-camera systems that were previously impossible to calibrate due to cumbersome calibration chessboards required by previous methods.