An Energy Formulation for Establishing the Correspondence Used in Projector Calibration

Drouin, Marc-Antoine; Godin, Guy; Roy, S.

National Research Council Canada
Conseil national de recherches Canada

Institute for Information Technology
Institut de technologie de l'information

NRC · CNRC

An Energy Formulation for Establishing the Correspondence Used in Projector Calibration *

Drouin, M.-A., Godin, G., Roy, S.
June 2008

* published in The Proceedings of the Fourth International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT). Atlanta, Georgia, USA. June 18-20, 2008. NRC 50357.

Canada

An Energy Formulation for Establishing the Correspondence Used in Projector Calibration

Marc-Antoine Drouin[1&2] Guy Godin[1] Sébastien Roy[2]
[1]National Research Council Canada [2]Université de Montréal
{first name}.{last name}@nrc-cnrc.gc.ca royss@iro.umontreal.ca

Abstract

Projector-camera systems are used in many applications that need geometric calibration of the projectors. As is the case for cameras, calibration of a projector requires the establishment of the correspondence between known features of a calibration target and the projector pixels. However, a fundamental difference is that the method must rely on external sensors (i.e. cameras) for its calibration. This introduces additional sources of error that affect the accuracy of the projector calibration. As a solution, we propose a multi-camera energy-based approach for establishing the correspondence between the projector and a calibration target. Performing projector calibration using the correspondence obtained by our method allows a significant improvement in accuracy. We demonstrate the validity of our approach using off-the-shelf and low cost equipment, and validate the results using a laser tracker.

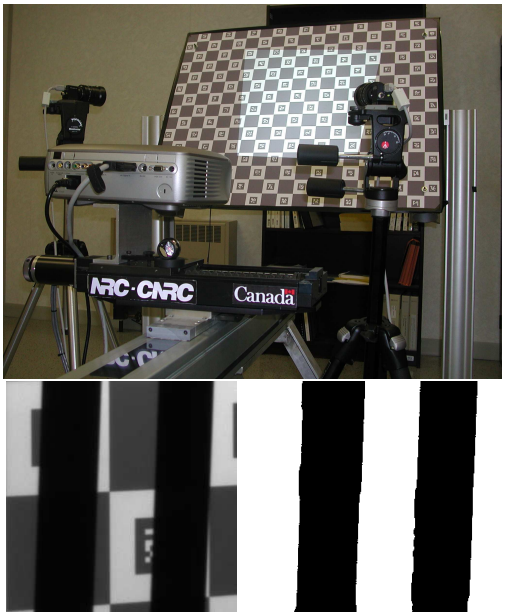

Figure 1. Top) Setup used for projector calibration. Bottom) Artifact present in decoded Gray code image with the standard approach. Left) Original image Right) Decoded Gray code pattern. Note that since the target is planar the transitions should form straight lines.

1 Introduction

In the last decades, projector-camera systems have been used in many applications such as 3D scanning, augmented reality and multi-projector visualization. For many applications, geometric calibration of the projector is a delicate and crucial step. By geometric calibration, we mean the calibration of the intrinsic and extrinsic (pose) parameters of the projector. It requires the establishment of the correspondence between known features of a calibration target and the projector pixels. When calibrating a camera, establishing this correspondence is straightforward. However, it is no longer the case for a projector, which is mathematically modeled in the same way as a camera, but one that cannot capture an image of the calibration target. The method must rely on external sensors (i.e. cameras) and structured light techniques for calibration. This introduces extra error sources that affect the accuracy of the projector calibration.

The limited depth of field, the lens aberration, the anti-aliasing filter and the small fill factor are some of the error sources originating from the cameras. Moreover, when using color cameras, artifacts are introduced by Bayer color processing and the Bayer matrix itself. The number and position of the cameras imaging the calibration target are also important. Photogrammetry or metrology grade camera and lens could be used to reduce the impact of those error sources, but this would dramatically increase the cost of a calibration bench. Futhermore, those error sources are added to the ones already present in traditional calibration, that is, the fitness of the calibration model, the quality of the calibration target and the characteristics of the device that must be calibrated. For projectors, the limited depth of field, the lens aberration and the color wheel affect the de-

tectability of intensity-edge of the projected image. Also, error is introduced by dithering and signal synchronization. Figure 1 shows a calibration bench and an example of artifacts that impact the accuracy of the correspondence and ultimately the accuracy of the recovered calibration parameters.

We propose a novel and robust energy framework for establishing the correspondence between the target and the projector. We show that our approach allows a significant improvement in accuracy when calibrating a projector using planar calibration[23, 18]. In our approach, we use many low resolution cameras with overlapping views of the calibration target. Whilst projector calibration methods usually rely on sparse correspondence between target and projector [11, 22], ours uses dense correspondence. Our method improves the robustness to noise and compensate for the different error sources. Also, the acquisition time is reduced since low resolution cameras usually have a higher frame rate than high resolution ones and are moreover cheaper. This makes our approaches very suitable for 3D computer vision applications. We demonstrate the validity of our approach using off-the-shelf and low cost equipment, and our results are validated using a laser tracker. The remainder of this paper is divided as follows: in Section 2, previous work is presented; Section 3 presents our method for establishing the correspondence between projector and calibration target; Section 4 presents the minimization framework; experimental results are discussed in Section 5.

2 Previous work

Since the establishment of the correspondence for projector calibration relies on structured light techniques, we will first briefly review the state of the art in structured light systems before presenting the previous work more directly related to our approach.

2.1 Structured light

In a structured light system used for projector calibration, the projector "applies" textures to the calibration target and the images (or patterns) are designed such that it is possible to determine which projector pixel illuminates the part of the target viewed by a camera pixel. A recent survey by Salvi *et al.* [16] classified structured light techniques in 3 categories: time-multiplexing, neighborhood codification and direct codification. Codes from the direct codification category use a single pattern that contains all the information needed to establish the correspondence. However, it is sensitive to noise. Codes based on neighborhood codification establish a sparse correspondence using a single pattern. The spatial resolution is reduced in order to gain robustness to noise. In this paper, we limit ourselves to the last category that uses many patterns in order to obtain a dense correspondence. The Gray code is probably the best known time-multiplexing code[8]. It uses a sequence of patterns composed of white and black stripes. For this reason, it is also known as a stripe-based code. The width of the stripes vary with the patterns. Whilst we use Gray code patterns, our approach is applicable to other codes. Energy-based methods for decoding structured light pattern have been proposed [2, 21, 9, 10, 19]. Some methods use minimization to remove ambiguity in their code[2, 21, 9]. In contrast, we use it to remove ambiguities introduced by the different error sources. In [10] an energy minimization *phase shift* method that adapts the patterns to the scene geometry is proposed. In [19] an off-line ICM-based method is presented.

2.2 Correspondence for projector calibration

Zhang *et al.* propose to project *phase shift* patterns in order to obtain a one-to-one mapping between pixels of the camera and of the projector [22] (similar remapping was also proposed for Gray code [20]). Then, fiducial markers on the calibration target are localized in the camera image and transferred to the projector image using the camera-projector mapping. This allows to match fiducial markers with projector pixels. Note that it is also possible to use this mapping to compute the homography between camera and projector that is used to transfer the fiducial markers' camera coordinates into projector coordinates. Legarda-Sáenz *et al.* have proposed a method where it is not necessary to have the complete one-to-one mapping [11]. Fiducial markers are localized within the image and a *phase shift* method is used to obtain the projector pixel corresponding to the camera points viewing fiducial markers[11]. All those methods allow to find which projector pixel matches a fiducial marker but do not allow to know which fraction of the pixel. We propose to use the *frontier* between projector pixels as features and for each *frontier* which we call transition we find the corresponding point on the calibration target. The correspondence problem is thus reversed: we no longer search for the position of a fiducial marker in the projector, but we search for the position of the projector pixels on the calibration target. This allows to obtain a dense correspondence rather than a sparse one as in [22, 11]. The fiducial markers are by nature strong intensity variations on the calibration target that make decoding of structured light patterns more prone to error. In the proposed approach, the frontiers between projector pixels are usually located on an untextured area of the calibration target, thus reducing structured light decoding error. As will be shown, our approach allows a significant improvement of accuracy.

An algorithm for auto-calibration of the pose between

the camera and projector of a reconfigurable structured-light scanner has been proposed[12]. Okatani and Deguchi[13] have proposed an auto-calibration method of projector-camera system where internal parameters of the projector must also be known. Our method for establishing the correspondence is complementary to those algorithms. Projector calibration is also discussed in [17] and [15].

3 Establishing the correspondence

Our algorithm for establishing the correspondence between a projector and a calibration target takes as input two sets of images $\mathcal{G}$ and $\mathcal{R}$. The first one contains the images $G^c_{h,b}$ of a planar calibration target acquired by camera c while the horizontal Gray code pattern representing bit b is projected. The set also contains the images $G^c_{v,b}$ that are similarly defined for the vertical Gray code patterns. The images $R^c_{h,b}$ and $R^c_{v,b}$ of the other set are the images $G^c_{h,b}$ and $G^c_{v,b}$ respectively where white and black stripes of the Gray code are reversed. Using both $\mathcal{G}$ and $\mathcal{R}$ increases the robustness of the technique. There are $2B$ images in each set $\mathcal{G}$ and $\mathcal{R}$, that is, one for each pattern (bit) and orientation of the Gray Code. Note that the cameras do not need to be fully calibrated, we only need the homographies between the calibration target and the cameras and possibly to pre-warp the images in order to compensate for radial (and possibly tangential) distortions. The set $\mathcal{P}$ contains points on the calibration plane. Those are chosen so that they form a regular lattice on the calibration plane at a specified sampled rate λ. Discussion on the determination of λ is postponed until Section 3.1. These points are not the same as the fiducial markers used to establish the correspondence between the cameras and calibration target. Also, there is in general no intensity variation on the target at the location of the points. This makes the decoding of structured light patterns less error prone. The coordinate system of the lattice is aligned with that of the fiducial markers. It is thus possible to know the position of a lattice point on the calibration target even if it is not visible in an image. Figure 2 shows a representation of the calibration target with both fiducial markers and points. Since the homographies between the cameras and calibration target are known (using fiducial markers from the the ARTag system [5]), it is possible to warp the images from $\mathcal{G}$ and $\mathcal{R}$ onto the lattice of points. Figure 2 also shows the warped images. Section 4 will present an energy formulation that combines the information from all the cameras and allows for each corresponding pair of images in $\mathcal{G}$ and $\mathcal{R}$ to determine whether each point of the lattice is on a white or black stripe of the projector image (see fig. 3 A and B). Once all the vertical patterns are processed, it is possible to decode the Gray code and identify the column of the projector that illuminates each point on the lattice. The same can be done for all the rows of the projector using the horizontal stripe patterns. It is thus possible to construct a pixel map that associates a projector's pixel to each point of the lattice (see fig. 3-C). The *transitions* in the pixel map are used as feature points in order to compute the homography between the projector and calibration target which is then used to perform the calibration of the projector. The localization error of the transition found using our discrete labeling is expected to shrink as the sampling rate of the lattice increases. The next sub-section discusses the selection of the sampling rate of the lattice of points $\mathcal{P}$. Note that some points of the lattice may not be visible from the projector, those are simply removed from the set $\mathcal{P}$ (see Fig. 3). In our experiments, we project a white frame and a black frame and points of $\mathcal{P}$ that do not exhibit a significant change of illumination are removed from $\mathcal{P}$.

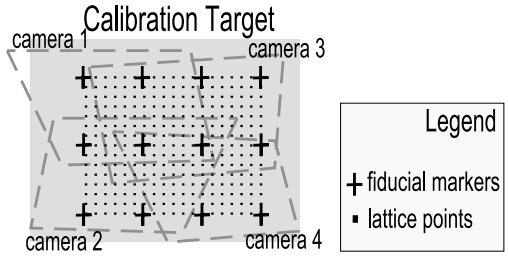

Figure 2. Representation of the calibration target.

3.1 Sampling the calibration target

Rather than directly specifying λ, it is much more intuitive to specify the desired mean magnification factor ω between the pixels in the different images and the points in $\mathcal{P}$ (i.e. ω controls the amount of sub-sampling that occurs in the images). In order to compute λ from the user-defined ω, we need the relation between a point $(x, y)^T$ in $\mathcal{P}$ and a pixel $(x^c, y^c)^T$ belonging to camera c which is

$$x = \lambda \frac{\mathbf{h}_1^{cw} \cdot (x^c, y^c, 1)}{\mathbf{h}_3^{cw} \cdot (x^c, y^c, 1)} \quad \text{and} \quad y = \lambda \frac{\mathbf{h}_2^{cw} \cdot (x^c, y^c, 1)}{\mathbf{h}_3^{cw} \cdot (x^c, y^c, 1)} \tag{1}$$

where $\mathbf{h}_i^{cw}$ is the i^{th} row of the matrix representing the homography $\mathbf{h}^{cw}$ that maps a point from the camera c to the calibration plane. When this transformation preserves the area, the determinant of the Jacobian

$$\frac{\partial(x, y)}{\partial(x^c, y^c)} = \begin{pmatrix} \frac{\partial x}{\partial x^c} & \frac{\partial y}{\partial x^c} \\ \frac{\partial x}{\partial y^c} & \frac{\partial y}{\partial y^c} \end{pmatrix} \tag{2}$$

is equal to one [1]. Similarly to the rectification algorithm of [6], the sampling rate λ is obtained by minimizing

$$\sum_{c<C} \sum_{X^c} (det \frac{\partial(x, y)}{\partial(x^c, y^c)} - \omega)^2 \tag{3}$$

which is linear in λ^2 and where C is the number of cameras (indexed from 0 to $C-1$). The choice of ω depends on the quality of the cameras used. For example, it will be large with a high-resolution photogrammetry-grade camera and small with a consumer webcam.

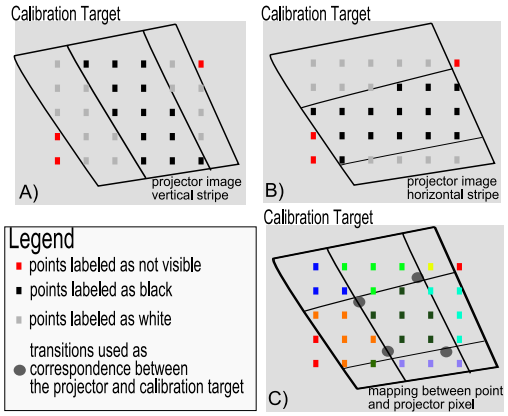

Figure 3. Process of establishing the correspondence between projector and calibration target: A) pixels are labeled as being lit by white or black vertical stripe of the projector; B) *Idem* for horizontal stripe; C) combining all the horizontal and vertical stripes allows to obtain the mapping between lattice points and projector pixels.

3.2 Target and camera position

In our experiments, the size of the calibration target was approximately 76×98 cm, printed on paper that was laminated on fiberboard, and was thus produced at very low cost. However, this process resulted in a calibration plane that is not perfectly planar. Table 4 contains planarity error metrics measured with a laser tracker of various calibration targets that we used to produce the results shown in Section 5.3. Also, another source of error is the non-uniform stretching of the paper during the making of the target; however, we did not quantify this error.

The diagram in Figure 4 shows two cameras imaging a perfect and a distorted calibration target. For camera 1, the variation in z (Δz) of the target induces a large variation of localization of the feature in the camera (Δx^1). This would introduce significant error when warping the image from the camera to the target and would impact the accuracy of the calibration. This is not the case for camera 2 since Δx^2 is small. Consequently, we should attempt to place the cameras such that the impact of an imperfectly planar target on the accuracy of the calibration is minimized. To do so, the perspective effect in $\mathbf{h}^{cw}$ must ideally be eliminated. This is achieved when the first two elements of $\mathbf{h}_3^{cw}$ are zero. As presented in [23], the homography $\mathbf{h}^{cw}$ is defined by the superior triangular matrix A^c containing the internal parameters of camera c, the rotation matrix R^c and translation vector T^c both representing external parameters of the camera c. Explicitly,

$$\mathbf{h}^{cw} = \left(A^c \begin{pmatrix} r_{11}^c & r_{12}^c & t_1^c \\ r_{21}^c & r_{22}^c & t_2^c \\ r_{31}^c & r_{32}^c & t_3^c \end{pmatrix} \right)^{-1} \tag{4}$$

where r_{ij}^c is the element of the i^{th} row and j^{th} column of the rotation matrix R^c, t_i^c is the i element of the translation vector T^c. The first two elements of $\mathbf{h}_3^{cw}$ are zero when r_{31}^c and r_{32}^c are zero, which corresponds to having the camera focal plane parallel to the calibration target. In our experiments, the camera position had a significant impact on the accuracy of the calibration. We position the cameras such that the image plane is as parallel as possible to the calibration target without being in the field of view of the projector.

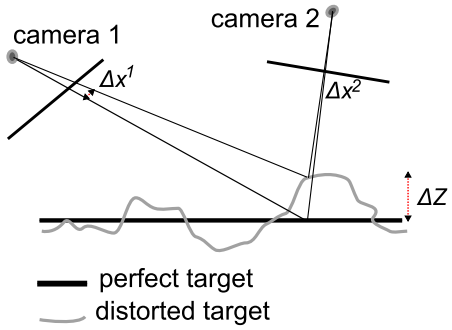

Figure 4. The planarity error of the calibration target has a bigger impact on camera 1 than on camera 2.

4 The energy function

Once λ is computed, the set of points $\mathcal{P}$ can be constructed and, for each pattern of the Gray code, every point can be labeled as illuminated by either a black or a white stripe of the projector. We define the set $\mathcal{L}$ as $\{0, 1\}$ where 0 and 1 represent a black and white stripe respectively. Explicitly, a $\mathcal{B}$-configuration $f : \mathcal{P} \mapsto \mathcal{L}$ associates a label to every point. The energy function for the pattern representing the bit b of the vertical Gray code is $E_{v,b}(f) =$

$$\sum_{\mathbf{p}\in\mathcal{P}} \Big(\underbrace{e_{v,b}(\mathbf{p}, f(\mathbf{p}))}_{\text{likelihood}} + \underbrace{\sum_{\mathbf{q}\in\mathcal{N}_\mathbf{p}} \gamma\delta(f(\mathbf{p}), f(\mathbf{q}))}_{\text{smoothing}} \Big) \tag{5}$$

where δ is 0 at 0 and 1 elsewhere and γ is a user-defined parameter. Further discussion about γ is postponed until Section 5. The neighborhood of point $\mathbf{p}$ is $\mathcal{N}_\mathbf{p}$. Since we

use many cameras that acquire close-up views of the calibration target, we expect that the lattice of points will be large. As an example, the lattice can reach 10000×10000 points while the resolution of common projectors is usually limited to 1024×768 or 1280×1024. In this case, there is approximately 100 lattice points that are illuminated by the same projector pixel. There is a high probability that two neighboring points are lit by the same projector pixel. Thus, a 4-connected neighborhood is used in the smoothing term of Eq. 5. Note that, when radial distortion is present, the transitions do not form straight lines. The solution of Eq. 5 could be obtained using Graph Cut [7]. However, finding a label to every point is computationally intensive. It is more efficient to find a location for each label discontinuity using the Border-Cut algorithm[3]. In the next section, we show how to modify the Border-Cut framework for the energy function of Eq. 5. The likelihood term of Eq. 5 is

$$e_{v,b}(\mathbf{p}, d) = \sum_{c \in \mathcal{C}(\mathbf{p})} \frac{e^{o(d)(G^c_{v,b}(M^{wc}(\mathbf{p})) - R^c_{v,b}(M^{wc}(\mathbf{p})))}}{\#\mathcal{C}(\mathbf{p})} \quad (6)$$

where $\mathcal{C}(\mathbf{p})$ is the set of index of the cameras viewing point $\mathbf{p}$, $\#$ is the cardinality of a set and $o(d)$ is 1 when d is 0 and -1 otherwise. M^{wc} is the function that applies the homography mapping a point $\mathbf{p}$ from the calibration plane onto a pixel in camera c (obviously M^{wc} is not a matrix since $\mathbf{p}$ is in Euclidean coordinate). The homography for each camera is computed automatically using the ARTag fiducial system [5]. The intensity values of the image are assumed to range from zero to one. *Mutatis mutandis*, the energy function is defined similarly for horizontal patterns.

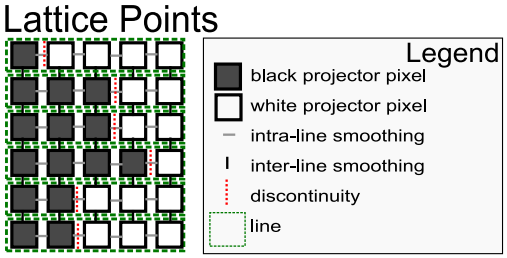

Figure 5. Line partitioning of the lattice.

4.1 Minimization using Border-Cut

The Border-Cut framework will be briefly presented for the energy function of Eq. 5 with vertical stripes. The modification for horizontal stripes is straightfoward, and more details about Border-Cut can be found in [3]. The points on the lattice $\mathcal{P}$ are partitioned into the set $\mathcal{L}$ of its horizontal lines. Each line α is a vector of points $(p^\alpha_1, \ldots, p^\alpha_N)$, N being the number of points on each line. Two lines are neighbors if two of their points are. The neighborhood of line α is denoted $\mathcal{N}_\alpha$ and is 2-connected since $\mathcal{N}_\mathbf{p}$ is 4-connected. This allows the use of efficient dynamic programming algorithms.

There are only two labels in our formulation: if we assume that there is a single label discontinuity in a line, then there is only one *discontinuity* point p^α_d in the line α preceding a point with a different label (see Fig. 5). Otherwise, it is possible to work on a subset of the lattice for which such a partition exists. A $\mathcal{D}$-configuration $d : \mathcal{L} \mapsto \mathcal{D}$ associates a *discontinuity* point to every line. The set $\mathcal{D}$ is simply $\{1,2,\ldots,N\}$ and represents the index of the *discontinuity* point. For simplicity, we assume that all the points on the left of the discontinuity have the same label and all those on the right have the other label. In our experiments, most of the time this is the case, since most of the error comes from discontinuity localization. The general case is presented in [3]. Allowing only one discontinuity per line with the pixel at the left of the discontinuity having label g, the energy to minimize is $E^g_{v,b}(b) =$

$$\underbrace{\sum_{\alpha \in \mathcal{L}} e^g_{v,b}(\alpha, b(\alpha))}_{\substack{\text{discontinuity} \\ \text{likelihood}}} + \underbrace{\sum_{\alpha \in \mathcal{L}} \sum_{\alpha' \in \mathcal{N}_\alpha} s_d(\alpha, \alpha', b(\alpha), b(\alpha'))}_{\text{inter-line smoothing}} \quad (7)$$

with respect to b, where the discontinuity likelihood e_g contains the likelihood of points on the same line and intra-line smoothing, which is the smoothing of neighboring point belonging to the same line (see Fig. 5). More details can be found in [3]. The inter-line smoothing (see Fig. 5) is the pixel smoothing of neighboring points belonging to different lines (see Fig. 5).

When all the points on the left of the discontinuity have the same label and all those on the right have the other label, the 2-connected inter-line smoothing function between two neighbor lines is simply the difference of the discontinuity location multiplied by γ. Thus, the fast message passing presented in [4] can be used. The complexity of the entire algorithm of Border-Cut when n discontinuities of the Gray code are present is $\Theta(n.N.M)$ where N is the length of a line and M the number of lines. For our problem, even if the set $\mathcal{P}$ is huge the interval in which the discontinuities move is limited and is N is small.

Note that we do not need to hold the entire lattice in memory, and we only require to compute the cost function for a relatively small number of lattice points. The Border-Cut framework requires an initialization which is provided by the solution obtained without smoothing by decoding the Gray code independently in each image. The discontinuities are found in the image and then projected onto the lattice. Also, note that rather than using the discrete Border-Cut framework [3], the level set framework could have been used [14]. However, a discretization already occurs in the

image formation process and the sampling rate of the lattice is computed taking into account this discretization. We believe that our discrete setting is adequate and our experimental results yield a significant improvement in accuracy. When no radial distortion is expected, level set would allow us to include the constraint that straight lines should remain straight lines under projective transformed.

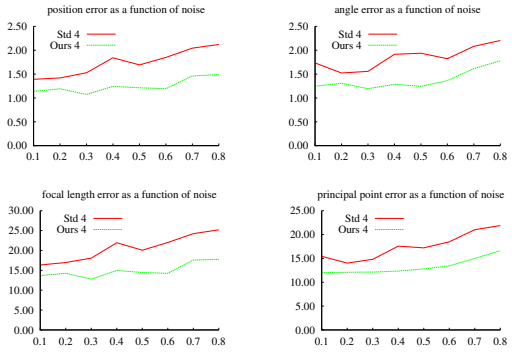

Figure 6. Results from simulation of the calibration of internal and external parameters. Top left) localization error (in mm) of the center of projection of the projector. Top right) orientation error (in degrees) of the optical axis of the projector Bottom left) focal length error (in pixels) Bottom right) localization error (in pixels) of the principal point.

5 Experimental results

We compare our approach for establishing the correspondence between projector and calibration target with the standard method of finding the fiducial markers of the calibration target in the camera's image, and decoding the Gray code at those points to obtain the projector-target correspondence [22, 11]. We choose the accuracy of projector calibration as a performance metric for our comparison. In the next sub-section we present the projector model that was used.

5.1 Projector calibration model

The projector model used in our experiments is similar to the one presented in [23] and [15]. A point in the world $P = (X^w, Y^w, Z^w)^T$ is first transformed into the coordinate system of the projector using $(X^p, Y^p, Z^p)^T = RP+T$ where R is a rotation matrix and T is a translation vector. The point $(X^p, Y^p, Z^p)^T$ is then projected onto the normalized image coordinate with $(x, y)^T = (X^p/Z^p, Y^p/Z^p)^T$. The radial distortion is then taken into account and the distortion point $(x', y')^T$ is computed using

$$x' = x\left[1 + k_1 r(x,y)^2 + k_2 r(x,y)^4\right] \quad (8)$$

$$y' = y\left[1 + k_1 r(x,y)^2 + k_2 r(x,y)^4\right] \quad (9)$$

where k_1 and k_2 are radial distortion coefficients and

$$r(x,y) = \sqrt{(r_x - x)^2 + (r_y - y)^2} \quad (10)$$

with radial center located at $(r_x, r_y)^T$. Finally, the pixel coordinates are $(f_x x' + s y' + c_x, f_y y' + c_y)^T$ where f_x and f_y are focal length, s is the skew and $(c_x, c_y)^T$ is the principal point. Note that, since we calibrate projectors using DLP chips which are manufactured with high accuracy and with square pixels, we set $f_x = f_y$ and $s = 0$. We use the planar calibration algorithm presented in [23]. Note that many calibration approaches for projectors do not model radial distortion. We provide results with and without radial distortion modeling.

5.2 Simulated data

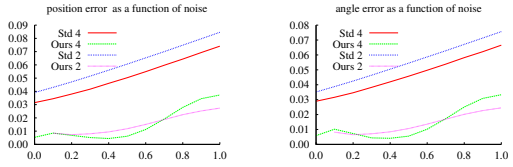

Figure 7. Results from simulation of the calibration of the external parameters when internal ones are known. Left) localization error (in mm) of the center of projection of the projector. Right) orientation error (in degrees) of the optical axis of the projector

We first provide results with simulated data where the images of the structured light patterns were corrupted. When decoding Gray code, the patterns with the narrowest stripes are the ones that are most error prone. In our simulation, each pixel of the image viewing those patterns have a probability of having it labeled from $\mathcal{L}$ (white stripe or black stripe) changed by one picked from a uniform distribution. Also, when an error occurs, the decoding is usually ambiguous and never just plain wrong. Our simulations reflect this. Figure 6 shows the error when the calibration uses 4 homographies for both our method and the standard one. The probability of a pixel to be corrupted ranges from 10% to 80%. Note that our method outperforms the standard approach [22, 11]. Figure 7 contains the results of simulations of the recovery of external parameters (pose) of a projector when internal parameters are known. We simulate the recovery for our algorithm and the standard one using 2 and 4 cameras. Also, since the pose sometime needs to be recovered in less controled environments, we allow corruption level to reach 100%, that is, all the low order bits of the Gray code have labels randomly assigned in each camera. Our method allows a significant improvement for all levels of corruption.

5.3 Real data

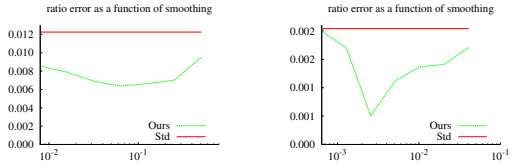

Figure 8. Variation of the error ratio as a function of smoothing for projector A (left) and C (right) with $\omega = 2$ (results are similar for $\omega = \sqrt{2}$).

A calibration plane was imaged by 4 cameras while Gray codes are projected on it. The resolution of the cameras is 1024×768. The projector was fixed on two translation tables allowing 2 DOF. The displacement was independently measured by a laser tracker accurate to 0.025 mm. Note that the laser tracker is only used to validate the results. We calibrated three projectors of different models, all based on DLP technology. The system setup is illustrated in Fig. 1. Note that in our experiments, the lattice is aligned with the checkerboard and the distance between adjacent points of the lattice is less than 0.2 mm. Moreover, 3 others were obtained by moving the projector in a triangular path in front of the fixed calibration plane. The displacement for the projectors A, B and C is respectively 339.94, 336.28 and 344.49 mm, at a projection distance of approximately 1.2 m. The displacement of the projector was measured by the laser tracker and compared to the one obtained from the homographies [23].

Table 1 shows the results for two values of ω, and the best values of γ for our algorithm and the standard algorithm. The results are provided for calibration with and without radial distortion modeling. Detailed results are presented in Table 2 for projector C. For two of the projectors, the accuracy is reduced when using the model with radial distortion. Projector A uses only one half of the lens, thus making the estimation of the distortion very difficult and unstable. For all projectors, our approach yields a much lower error rate. For projector C (using radial distortion modeling) we could achieve an error ratio of $\frac{8}{10000}$ which is more than 4 times less than the other method. For this projector, when using our method, the addition of radial distortion to the calibration model allows to improve the accuracy by a factor of 4. For the standard method, the accuracy is only improved by a factor of 2. Figure 1 shows some artifacts present in a decoded Gray code image using the standard approach. These artifacts impact the accuracy of the calibration as was shown in Table 1. Table 4 contains metrics measured with a laser tracker of the calibration target used for projector A,B and C. Table 3 contains the translation error obtained by comparing the values of the laser tracker and those obtained through pose estimation with a bundle adjustment. Our approach allows a decrease in the error by factors 2.8, 3.3 and 4.0 for projector A, B and C respectively. Figure 8 shows, for two projectors, a stability analysis for the smoothing parameter γ for our algorithm, giving the error in displacement ratio over a broad range of values. The error value for the standard approach is also shown. Our approach gives lower error over the complete range.

	calibration error					
	with radial			no radial		
	mm	ratio	reproj.	mm	ratio	reproj.
Projector A						
$\omega = \sqrt{2}$	**0.914**	0.00807	0.44	**0.879**	0.00776	0.62
$\omega = 2$	0.956	0.00844	0.38	0.892	0.00787	0.50
Std	**1.388**	0.01225	0.43	**1.527**	0.01348	0.61
Projector B						
$\omega = \sqrt{2}$	**0.954**	0.00851	0.27	**0.764**	0.00681	0.35
$\omega = 2$	0.953	0.00850	0.28	0.800	0.00714	0.34
Std	**1.135**	0.01012	0.27	**1.054**	0.00940	0.34
Projector C						
$\omega = \sqrt{2}$	0.086	0.00075	0.37	0.262	0.00228	0.57
$\omega = 2$	**0.058**	0.00051	0.34	**0.229**	0.00200	0.52
Std	**0.234**	0.00204	0.32	**0.478**	0.00416	0.49

Table 1. Best results in reprojective error and displacement ratio using two different values of ω.

	detailed calibration results for Projector C	
parameter	Std	Ours
focal (pixel)	2033.433	2037.928
center (pixel)	(492.408, 929.146)	(492.810, 929.784)
radial center (pixel)	(0.003196, -0.115120)	(0.003024, -0.112467)
(k_1,k_2)	(-0.106738, 0.569304)	(-0.106022, 0.555041)
position world 1 (mm)	(708.98, 528.33, 996.87)	(709.69, 528.53, 998.98)
position world 2 (mm)	(691.56, 532.17, 897.94)	(692.15, 532.45, 899.74)
position world 3 (mm)	(592.69, 532.67, 913.31)	(593.34, 532.77, 915.20)

Table 2. Detailed results for the calibration of projector C using radial distortion model.

6 Conclusion

A new procedure for establishing the correspondence between a projector and a calibration target was proposed. It is based on a multi-camera energy-based approach where many low resolution cameras having overlapping views of the calibration target are used. Our formulation includes a data term that takes into account all camera images and a smoothing term that regulates noise. Moreover, the optimization is performed on the calibration target coordinate system. We demonstrated the validity of our approach on simulated data and on real imagery using off-the-shelf and low cost equipment; our results were validated using a laser tracker. Our results show that performing projector calibration using the correspondence obtained by our method

	Error in millimeter as measured by the laser tracker	
parameter	Std	Ours
	calibration result for Projector A (no radial)	
translation 1	3.903	0.326
translation 2	1.627	0.113
translation 3	0.102	1.553
total	**5.632**	**1.992**
	calibration result for Projector B (no radial)	
translation 1	0.690	1.487
translation 2	2.981	0.216
translation 3	2.385	0.110
total	**6.056**	**1.813**
	calibration result for Projector C (with radial)	
translation 1	0.383	0.049
translation 2	0.284	0.107
translation 3	0.036	0.019
total	**0.703**	**0.175**

Table 3. Error on the length of translation (in mm) as measured by calibration. Note that the computation is done through a bundle adjustment thereby explaining why some individual values may increase.

	calibration plane used for each projector	
	A	B & C
RMS	0.187 mm	0.244 mm
Form	0.550 mm	1.135 mm

Table 4. Standard metrics used to evaluate the fitness of the calibration plane computed using the laser tracker.

allows a significant improvement in accuracy. As for future work, we would like to investigate how our method would improve the accuracy of a 3D structured light scanner.

Acknowledgements

We would like to thank our colleagues Luc Cournoyer, Michel Picard and Daniel Gamache for their support.

References

[1] V. Arnold. *Mathematical Methods of Classical Mechanics*. Springer-Verlag, 1989.

[2] C. Chen, Y. Hung, C. Chiang, and J. Wu. Range acquisition using color structured lighting and stereo vision. *Image Vision Comput.*, 15:445–456, 1997.

[3] M.-A. Drouin, M. Trudeau, and S. Roy. Improving border localization of multi-baseline stereo using border-cut. In *IEEE Conf. on Comput. Vision and Pattern. Recogn.*, pages 511–518, 2006.

[4] P. F. Felzenszwalb and D. P. Huttenlocher. Efficient belief propagation for early vision. *Int. J. Comput. Vision*, 70(1):41–54, 2006.

[5] M. Fiala. ARTag revision 1. a fiducial marker system using digital techniques. Technical Report NRC/ERB-1117, National Research Council of Canada, 2004.

[6] J. Gluckman and S. K. Nayar. Rectifying transformations that minimize resampling effects. In *IEEE Conf. on Comput. Vision and Pattern. Recogn.*, 2001.

[7] D. M. Greig, B. T. Porteous, and A. H. Seheult. Exact maximum a posteriori estimation for binary images. *J. R. Statist. Soc.*, 51(2):271–279, 1989.

[8] S. Inokuchi, K. Sato, and F. Matsuda. Range imaging system for 3-D object recognition. In *Int. Conf. on Pattern Recogn.*, pages 806–808, 1984.

[9] T. P. Koninckx, I. Geys, T. Jaeggli, and L. Van Gool. A graph cut based adaptive structured light approach for real-time range acquisition. In *Int. Symp. on 3D Data Process., Vis. and Transmiss.*, 2004.

[10] T. P. Koninckx, P. Peers, P. Dutre, and L. Van Gool. Scene-adapted structured light. In *IEEE Conf. on Comput. Vision and Pattern. Recogn.*, pages 611–618, 2005.

[11] R. Legarda-Sáenz, T. Bothe, and W. P. Jüptner. Accurate procedure for the calibration of a structured light system. *Optical Engineering*, 43(2):464–471, 2004.

[12] Y. F. Li and S. Y. Chen. Automatic recalibration of an active structured light vision system. *IEEE J. Robotics and Automation*, 19(2), 2003.

[13] T. Okatani and K. Deguchi. Autocalibration of a projector-camera system. *IEEE Trans. Pattern Anal. and Mach. Intell.*, 27(12):1845–1855, 2005.

[14] S. J. Osher and R. P. Fedkiw. *Level Set Methods and Dynamic Implicit Surfaces*. Springer-Verlag, Berlin, 2002.

[15] M. Ribo, M.; Brandner. State of the art on vision-based structured light systems for 3d measurements. *Robotic Sensors: Robotic and Sensor Environments, 2005. International Workshop on*, pages 2–6, 30 Sept.-1 Oct. 2005.

[16] J. Salvi, J. Pages, and J. Batlle. Pattern codification strategies in structured light systems. *Pattern Recogn.*, 37(4):827–849, 2004.

[17] G. Sansoni, M. Carocci, and R. Rodella. Calibration and performance evaluation of a 3-D imaging sensor based on the projection of structured light. *IEEE Transactions on Instrumentation and Measurement*, 49(3):628–636, June 2000.

[18] P. F. Sturm and S. J. Maybank. On plane-based camera calibration: A general algorithm, singularities, applications. In *IEEE Conf. on Comput. Vision and Pattern. Recogn.*, pages 432–437, 1999.

[19] J. Tardif and S. Roy. A MRF formulation for coded structured light. In *3-D Digit. Imag. and Model.*, 2005.

[20] J.-P. Tardif, S. Roy, and M. Trudeau. Multi-projectors for arbitrary surfaces without explicit calibration nor reconstruction. In *3-D Digit. Imag. and Model.*, pages 217–224, 2003.

[21] J. Zhang, B. Curless, and S. Seitz. Rapid shape acquisition using color structured light and multi-pass dynamic programming. In *Int. Symp. on 3D Data Process., Vis. and Transmiss.*, pages 24–36, 2002.

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

[23] Z. Zhang. A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. and Mach. Intell.*, 22(11):1330–3334, 2000.