

A Quasiconvex Formulation for Radial Cameras

Carl Olsson^{1,2}, Viktor Larsson³, Fredrik Kahl¹

¹Chalmers University of Technology, Sweden

²Lund University, Sweden

³Department of Computer Science, ETH Zürich, Switzerland

Abstract

In this paper we study structure from motion problems for 1D radial cameras. Under this model the projection of a 3D point is a line in the image plane going through the principal point, which makes the model invariant to radial distortion and changes in focal length. It can therefore effectively be applied to uncalibrated image collections without the need for explicit estimation of camera intrinsics.

*We show that the reprojection errors of 1D radial cameras are examples of quasiconvex functions. This opens up the possibility to solve a general class of relevant reconstruction problems globally optimally using tools from convex optimization. In fact, our resulting algorithm is based on solving a series of LP problems. We perform an extensive experimental evaluation, on both synthetic and real data, showing that a whole class of multiview geometry problems across a range of different cameras models with varying and unknown intrinsic calibration can be reliably and accurately solved within the same framework.*¹

1. Introduction and Related Work

Quasiconvexity in 3D reconstruction problems have been studied since the work of [18]. The paper made the crucial observation that with the pinhole projection, each residual has convex sublevel sets. Since this property is preserved by the max operation, it was proposed to solve the triangulation problem by minimizing the maximal residual (the L_∞ -norm) rather than the least squares error. The above mentioned paper works with regular pinhole cameras $\mathbb{P}^3 \rightarrow \mathbb{P}^2$. However, cameras with radial distortion introduce a mapping that is not projective, see Figure 1, and hence they have to be pre-calibrated and their corresponding

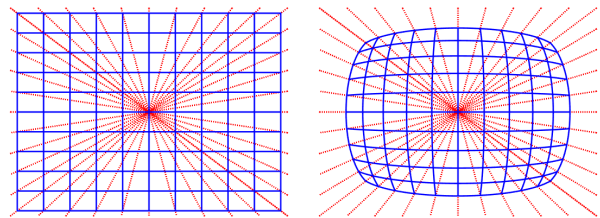


Figure 1. Radial distortion of an image. Image points move along lines going through the principal point.

images undistorted before we can apply the above framework. In this paper we instead study the 1D radial camera model that is invariant to changes in focal length and radial distortion [42]. The model encompasses any radially symmetric camera, which includes most camera types (e.g., pinhole, fisheye and catadioptric cameras). We show that this model also belongs to the quasiconvex framework, which opens up the possibility to solve a class of 3D reconstruction problems with unknown and varying focal length and radial distortion parameters, while still being able to guarantee global optimality of the obtained solution.

In [20, 22, 24], the work of [18] for triangulation was extended to a more general class of optimization problems in multiple view geometry. Some of the problems addressed within the same framework were triangulation, homography estimation, camera resectioning. In addition to studying the standard pinhole residual, the above papers showed how to handle angular errors as well as covariance weighted distance measures. In followup works the theory has been generalized to 1D cameras [12] and triangulation of planar structures [35]. The perhaps most interesting problem of this class is the so called known-rotation problem [40, 22, 46] where the orientation part of the camera is assumed to be known and optimization is performed over the positions of both the cameras and the 3D points. The ability to reliably solve this large scale problem lead to a new class of non-sequential methods [31, 11, 35, 32, 5]. These methods typically divide the reconstruction problem into

¹This work has been funded by the Swedish Research Council (grants no. 2016-04445, 2018-05375), the Swedish Foundation for Strategic Research (Semantic Mapping and Visual Navigation), the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation and an ETH Zurich Postdoctoral Fellowship.

two parts, where camera orientations are first determined from relative rotation estimates using rotation averaging [14, 13, 19, 31, 3, 43, 7, 16, 15] after which camera and 3D point positions are estimated. For traditional sequential reconstruction algorithms there is a trade-off between using views that are close enough to obtain sufficiently many matches and distant enough to obtain a stable enough geometry [33]. In addition they suffer from drift [8] that needs to be countered by running bundle adjustment multiple times during reconstruction. The non-sequential approaches are able to circumvent these issues since they consider all the data simultaneously [11].

While the least squares formulation of the above problems can in general have several local minima, see [22] for an example, one can in many cases use the L_∞ formulation to bound the hessian of the least squares formulation to verify optimality of a candidate solution [17].

The downside of minimizing the maximal error is the sensitivity to outliers stemming from mismatches. These typically have large reprojection errors and therefore the quality of the solution to the L_∞ -problem can be heavily affected by them. On the other hand, this makes the formulation efficient for outlier rejection. It can be shown [41] that the set of residuals that take the maximal value contains at least one outlier. In [34] this result was refined to show that there is a subset of $d + 1$ residuals, where d is the number of unknowns of the problem, that take the maximal value and contains at least one outlier. By searching over subsets of size $d + 1$ an outlier removal algorithm that works in polynomial time can be achieved [29]. The algorithm is however only practical for problems where d is small. For large scale problems robust relaxations have instead been proposed [36, 39, 45, 47, 9].

The first work to use the radial constraint was [44]. Later multiple view geometry with these cameras was studied in [42]. In [25] a factorization approach for simultaneously solving for cameras and points was proposed. The approach is based on a convex relaxation using the nuclear norm and as such it is relatively sensitive to noise. Minimal solvers for the absolute pose problem were presented in [26, 27] and a non-parametric radial distortion model was presented in [6]. In [30] the 1D radial camera was used to bootstrap the optimization in multi-camera calibration system. In a very recent work [28] Larsson et al. present a sequential reconstruction pipeline using radial cameras. They show that accurate metric solutions can be constructed from unordered image collections with varying and unknown distortion parameters and focal lengths. To our knowledge there is however no prior work that has observed or utilized the quasiconvexity of these problems.

2. Framework

2.1. 1D Radial Cameras and Objective Functions

Radial distortion is common in cameras with a wide field of view. Under this distortion model image points are displaced by rescaling their coordinates based on the distance from the principal point, as illustrated in Figure 1. Consequently, image lines going through the principal point are invariant to radial distortion. Consider a regular pinhole camera matrix of size 3×4 , $P = K [R \ t]$ with pose parameters (R, t) and intrinsic calibration

$$K = \begin{pmatrix} f & fs & x_0 \\ 0 & f\gamma & y_0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (1)$$

where f is the focal length, s skew and γ aspect ratio. The principal point (x_0, y_0) , if it is known, then the camera can be normalized by multiplying the image points and the camera matrix from the left with

$$N = \begin{pmatrix} 1 & 0 & -x_0 \\ 0 & 1 & -y_0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (2)$$

We can therefore assume that $(x_0, y_0) = (0, 0)$. Now consider a 3D point with homogeneous coordinates X which is projected with $P = K [R \ t]$, and let the 3-vector $[R \ t] X$, which are the local camera coordinates, be denoted by $(x_{cam}, y_{cam}, z_{cam})$

The resulting projection (in regular coordinates) is

$$\frac{f}{z_{cam}} (x_{cam} + sy_{cam}, \gamma y_{cam}). \quad (3)$$

It is clear that under radial distortion this point is on the line through the origin with the directional vector $(x_{cam} + sy_{cam}, \gamma y_{cam})$, which is independent of the third coordinate z_{cam} (the depth) and the focal length f . The 1D radial camera that projects 3D points to lines can therefore be modeled as $P_{2 \times 4} = K_{2 \times 2} [R_{2 \times 3} \ t_{2 \times 1}]$, where $R_{2 \times 3}$ and $t_{2 \times 1}$ are the two first rows of R and t respectively and $K_{2 \times 2}$ is top left 2×2 block of K . Comparing the expressions $K_{2 \times 2} [R_{2 \times 3} \ t_{2 \times 1}]$ and $K [R \ t]$ we see that the difference between these models is that the third row of R and the third entry of t is lacking in the radial camera. Since R is a rotation its third row can easily be computed from the first two. However, the third entry of t cannot be determined. This corresponds to a movement along the camera's viewing direction. Hence the radial camera model only allows us to recover the location of the camera center up to line in 3D.

2.2. Reprojection Errors

In this section we present the error function that we use for estimation. The same residual errors were used in [28]

but these were combined using the two-norm rather than the max-norm that we intend to use.

Suppose that we have observed a point in the image, represented by a 2-vector x . Under the radial camera model this point is assumed to be a sample from a line l with directional vector v given by the projection $v = P_{2 \times 4} X$. Under the Gaussian noise assumption the point on the line l that maximizes the likelihood of the measurement x is the orthogonal projection of x onto l , which is given by $\frac{vv^T}{\|v\|^2}x$. We therefore let the reprojection error be the distance between x and its projection given by

$$r(v) = \left\| \left(I - \frac{vv^T}{\|v\|^2} \right) x \right\|. \quad (4)$$

Note that since this objective function is scale invariant, it is not dependent on what particular directional vector v is used to represent l . Also, note that the objective function has a natural unit, namely pixels, so we do measure a distance in the image plane.

We will further restrict the vector v to the set $x^T v > 0$. Geometrically this means that the angle between x and v is less than 90° . For noise free data x and v have either the same ($x^T v > 0$) or opposite ($x^T v < 0$) directions. In case of opposite directions we see from (3) that $z_{cam} < 0$ which means that the 3D point X is behind the camera and cannot be visible. Figure 2 shows the geometric interpretation of our reprojection error.

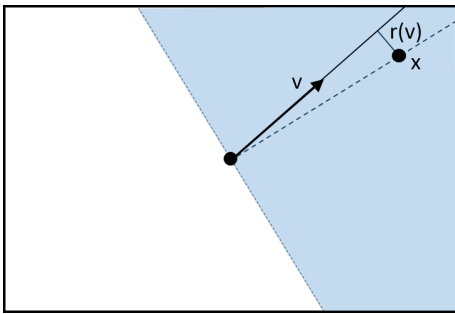


Figure 2. Geometric interpretation of the reprojection error $r(v)$. The constraint $v^T x > 0$ is shown as light blue. Note that rescaling v does not change $r(v)$.

In what follows we will show that the residual error function is indeed an example of a quasiconvex function which makes it possible to globally solve a class of reconstruction problems with standard methods from convex optimization.

2.3. Quasiconvexity and Optimization

A function f on a convex set \mathcal{C} is called quasiconvex if its sub-level sets

$$S_\alpha(f) = \{v \in \mathcal{C}; f(v) \leq \alpha\} \quad (5)$$

are convex for all α . From the definition it is easy to see that such a function has non-increasing objective values on a line segment $(1 - \lambda)v + \lambda v^*$ from an arbitrary point v to a global minimizer v^* . Therefore, there are no strict local minimizers other than the global ones (although the function can be constant on parts of the domain).

Optimization problems with multiple quasiconvex residuals can be reliably optimized under the max-norm. The reason is that quasiconvexity is preserved by the max operation, that is,

$$f(v) = \max_i f_i(v) \quad (6)$$

is quasiconvex on $\mathcal{C} = \cap_i \mathcal{C}_i$ if f_i is quasiconvex on \mathcal{C}_i . In Section 3 we describe how to optimize quasiconvex functions using the so called bisection algorithm.

We now show that residual functions of the type (4) are quasiconvex in the parameters v on the set $v^T x > 0$. First note that $P_v := \left(I - \frac{vv^T}{\|v\|^2} \right)$ is a symmetric matrix and a projection. Therefore

$$r(v)^2 = x^T P_v^T P_v x = x^T P_v x = \|x\|^2 - \frac{(v^T x)^2}{\|v\|^2}. \quad (7)$$

Rearranging terms gives

$$r(v)^2 \leq \epsilon^2 \Leftrightarrow (\|x\|^2 - \epsilon^2)\|v\|^2 \leq (v^T x)^2. \quad (8)$$

Note that if $\epsilon > \|x\|$ then the inequality is trivially fulfilled for all v . Under the assumption that $v^T x > 0$ we can equivalently write

$$r(v) \leq \epsilon \Leftrightarrow \left\| \sqrt{\|x\|^2 - \epsilon^2} v \right\| \leq v^T x. \quad (9)$$

For a fixed x the above constraint is a known as a second order cone and is well known to be convex [4].

Figure 3 shows the objective function $r(v)$ and its level contours. Note that without the constraint $v^T x > 0$ the sublevel sets are non-convex.

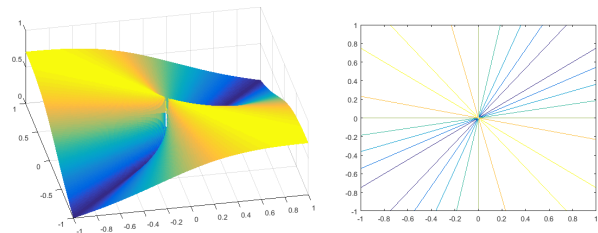


Figure 3. The objective function $r(v)$ (left) when $x = (1/2, 1/2)$ and its level contours (right).

It is easy to show that quasiconvexity is preserved under affine transformations. We therefore conclude this section by noting that any function of the form $\tilde{r}(u) = r(Au + B)$ is quasiconvex on $x^T (Au + B) > 0$. In Section 4 we show a number of multi-view geometry applications that fulfill this condition.

3. Optimization

In this section, a brief description of the so called bisection algorithm is given that we will use for optimization. We are interested in solving

$$\min_{v, \epsilon} \quad \epsilon \quad (10)$$

$$\text{s.t.} \quad f(v) \leq \epsilon, \quad (11)$$

where $f(v)$ is a quasiconvex function. If ϵ is fixed and we only minimize over v , then we obtain a convex feasibility problem since by definition the inequality in (11) is a convex constraint. Therefore testing whether there is a solution with objective value ϵ can be solved efficiently using standard methods. The bisection algorithm refines a lower bound ϵ_l and an upper bound ϵ_u on the optimal ϵ by solving a sequence of feasibility problems, see Algorithm 1. Recall that $S_\epsilon(f)$ denotes the set of solutions to the convex feasibility problem $f(v) \leq \epsilon$, cf. (5).

Data: ϵ_l, ϵ_u

Result: ϵ, v

while $\epsilon_u - \epsilon_l > \tau$ **do**

$\epsilon = \frac{\epsilon_u + \epsilon_l}{2};$
if $S_\epsilon(f) \neq \emptyset$ **then**
 $v \in S_\epsilon(f);$
 $\epsilon_u = \epsilon;$
else
 $\epsilon_l = \epsilon;$
end

end

Algorithm 1: The bisection algorithm for minimizing a quasiconvex function f .

The bisection algorithm which is known to have linear convergence is also the basis for other, computationally more efficient variants, see [2, 10, 46]. For instance, the Scaled Dinkelbach's Algorithm, evaluated in [2], has super-linear convergence.

3.1. Feasibility Tests Using Linear Programs

In case of 1D radial cameras we need to check feasibility of constraint sets of the type (9). We will show that these are equivalent to a linear constraint set and therefore the feasibility problem can be solved with linear programming.

The constraint (8) can equivalently be written

$$v^T \underbrace{((\|x\|^2 - \epsilon^2)I - xx^T)}_{:=M} v \leq 0. \quad (12)$$

If $x = (x_1, x_2)$ we let $x_\perp = (x_2, -x_1)$. The matrix xx^T has eigenvalues $\lambda_1 = \|x\|^2$ and $\lambda_2 = 0$ with corresponding eigenvectors $e = \frac{1}{\|x\|}x$ and $e_\perp = \frac{1}{\|x\|}x_\perp$. Therefore the

matrix M has eigenvalues $\gamma_1 = -\epsilon^2$ and $\|x\|^2 - \epsilon^2$ with eigenvectors e and e_\perp . Therefore $M = E\Gamma E^T$, where $E = [e \ e_\perp]$ and

$$\Gamma = \begin{pmatrix} -\epsilon^2 & 0 \\ 0 & \|x\|^2 - \epsilon^2 \end{pmatrix}. \quad (13)$$

If we let $\bar{v} = E^T v$ we see that (12) can be written

$$\bar{v}^T \Gamma \bar{v} \leq 0 \Leftrightarrow -\epsilon^2 \bar{v}_1^2 + (\|x\|^2 - \epsilon^2) \bar{v}_2^2 \leq 0. \quad (14)$$

Note that $\bar{v}_1 = v^T \frac{x}{\|x\|} > 0$. Solving for \bar{v}_2 gives

$$|\bar{v}_2| \leq \sqrt{\frac{\epsilon^2}{\|x\|^2 - \epsilon^2}} \bar{v}_1, \quad \bar{v}_1 > 0, \quad (15)$$

or equivalently

$$\pm v^T x_\perp \leq \sqrt{\frac{\epsilon^2}{\|x\|^2 - \epsilon^2}} v^T x, \quad v^T x > 0, \quad (16)$$

which corresponds to the three linear inequality constraints. In practice $v^T x > 0$ is replaced by $v^T x > \tau$ for some small positive number τ . For the different applications listed in Section 4 the v variables are linear expressions in the unknowns. It is therefore clear that the feasibility problem will be a linear program for these applications.

4. Applications

In this paper we have shown that the radial reprojection error (4) can be minimized in a quasiconvex optimization framework. This allows us to recover globally optimal solutions in a wide variety of geometric estimation problems where the radial reprojection model is valid. In the following sections we show some example applications where this framework can be used.

Experimentally we focus on evaluating the approach with respect to stability and invariance to radial distortion and changes in focal length for a wide range of different camera models and scene settings. In all experiments, we use MOSEK [1] to solve the LP problems and we set $\tau = 10^{-6}$ pixels in the bisection algorithm and initialize with $\epsilon_l = 0$ and $\epsilon_u = 200$.

4.1. Triangulation

To determine the position of a 3D point X from projections $v_i = P_i X$ for 2×4 radial cameras $P_i, i = 1, \dots, n$ we want to minimize

$$f(X) = \max_i r_i(P_i X) \quad (17)$$

where $r_i(\cdot)$ is the residual function (4) for point x_i and the optimization domain is over the set $x_i^T P_i X > 0, i =$

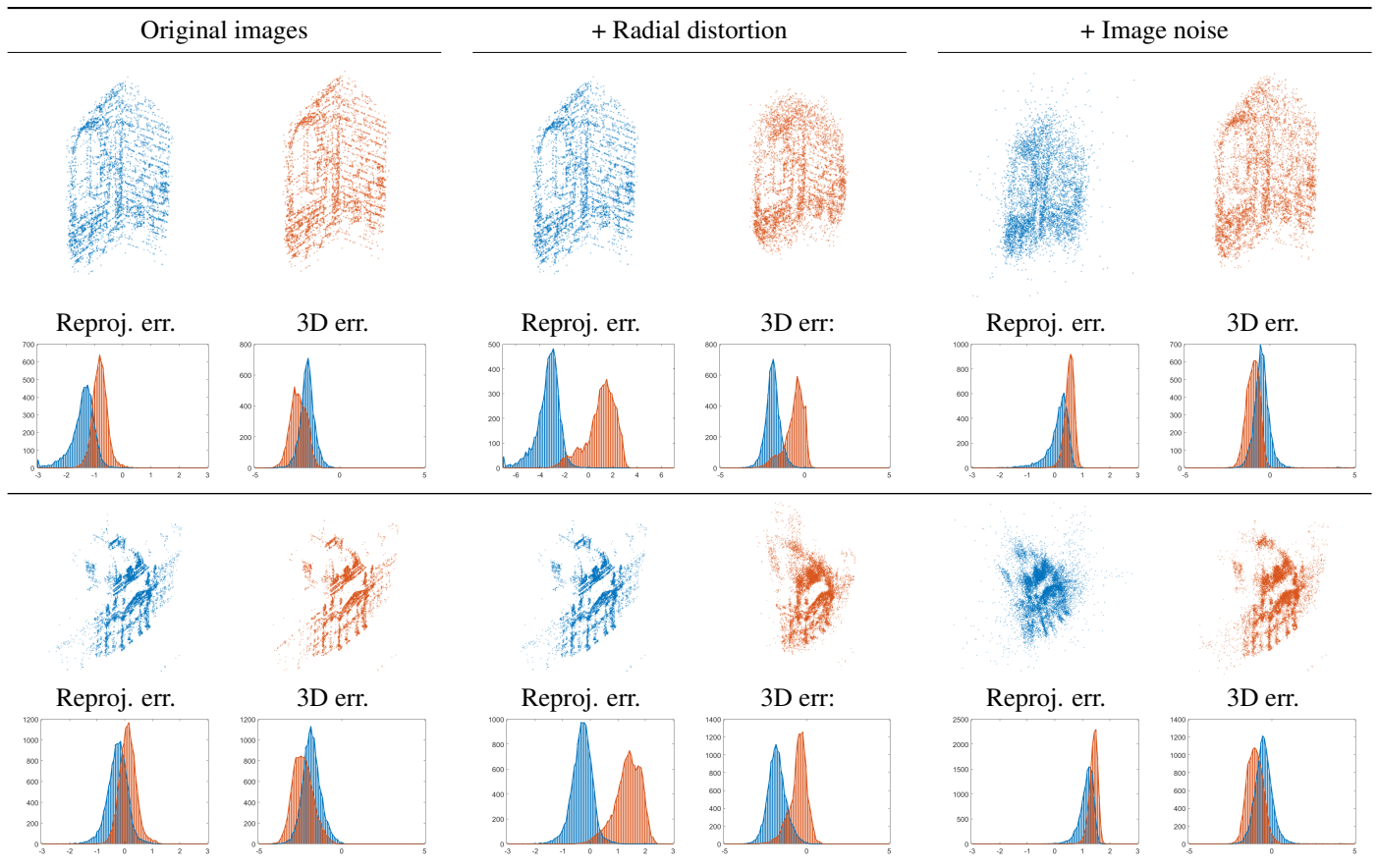


Figure 4. Triangulation results with the radial camera model (*blue*) vs. pinhole (*orange*). Histograms show log of the reprojection errors (*log pixels*) and log of the 3D errors (*log meters*). The radial camera model is invariant to the radial distortion (*middle column*) but is more sensitive to image noise (*right column*) due to the weaker geometric constraints. *Top row*: Corner (7618 pts, 9 cams). *Bottom row*: De Guerre (11178 pts, 35 cams).

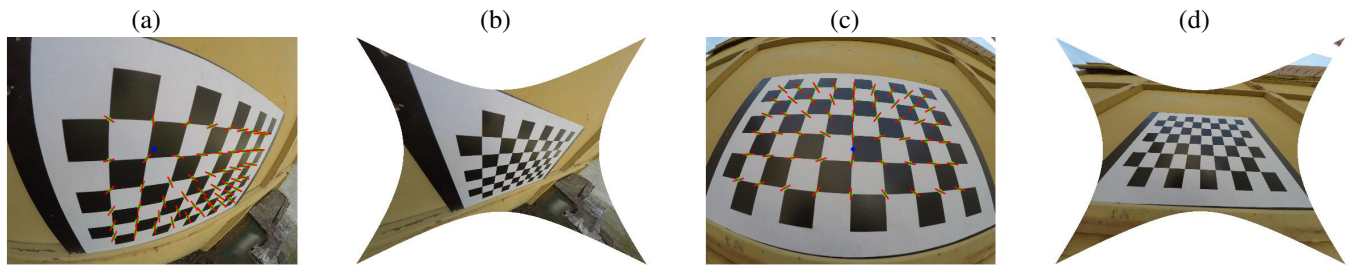


Figure 5. Estimated $\mathbb{P}^2 \rightarrow \mathbb{P}$ transform from a grid of planar points to a images containing a calibration pattern. Green points in (a) and (c) show matches to the grid. Red line segments show the lines obtained from the estimated transformation in the vicinity of a match. In (b) and (d), the undistorted images obtained from the estimated transformation are shown.

$1, \dots, n$. Note that the back projection of an image line is a plane in 3D. Since any three planes intersect in 3D the triangulation problem can only be uniquely solved when $n \geq 3$.

Figure 4 shows the result of solving the triangulation problem on two datasets. Here we have solved for all 3D-points that are co-visible in 4 or more views. In the left

column we show the result obtained with the original image observations. In this case the real images are largely free of radial distortion. The histograms show the logarithm (base 10) of the radial-reprojection error in pixels (both for the radial camera and the pinhole model) and the logarithm (base 10) of the 3D-errors (in meters). The 3D errors are computed with respect to a baseline reconstruction from [11].

In the middle column we have added synthetic radial distortion by applying the polynomial

$$f(r) = 1 + k_1 r^2 + k_2 r^4 + k_3 r^6, \quad (18)$$

with $(k_1, k_2, k_3) = (-1.5, 0.75, 0.5)$, to the normalized image coordinates. (Examples of the distorted images can be seen in Table 1.) In the case of the radial model the reconstruction quality is unaffected due to its invariance.

To the right we instead added synthetic noise with standard deviation 5 pixels to the real image coordinates. Because the radial model uses less data it is more sensitive to noise and in general gives larger errors. We also note that there is a small set of points (not shown in the 3D reconstruction) that are highly sensitive to noise giving very inaccurate 3D positions of as much as $10^4 - 10^5$ (meters).

4.2. Plane-Induced Radial Projections

If a planar object is projected into a scene under regular perspective projection there is a homography H between the points of the plane X_i and the corresponding points of the image x_i . To introduce invariance to radial distortion we instead apply a 2×3 matrix $H_{2 \times 3}$ to the homogeneous coordinates of the points in the plane and interpret the result as directional vectors of lines going through the principal point. Note that $H_{2 \times 3}$ can be viewed as a $\mathbb{P}^2 \rightarrow \mathbb{P}^1$ projection. The resulting objective function is therefore

$$f(H_{2 \times 3}) = \max_i r_i(H_{2 \times 3} X_i), \quad (19)$$

where $r_i(\cdot)$ is the residual error corresponding to the point x_i . In Figure 5 (a) and (c) we present the results of fitting a square grid of points to images of a calibration pattern captured by wide-field-of-view GoPro4 camera. Here the green points are measurements and the red line segments are the corresponding lines resulting from the projection. (For visualization purposes we only plot the part of the line which is close to its corresponding point.) The blue point is the principal point. Given the 2×3 homographies, the third row and radial distortion parameters can be estimated using the method proposed by Thirithala and Pollefeys [42]. Figure 5 (b) and (d) show the undistorted images with the estimated parameters from this method. The full data set contained 32 images and in Figure 6 we plot the radial reprojection errors (in pixels) over all images. We can see that the homography estimation accurately recovers the mapping from the checkerboard to the distorted image.

4.3. Resectioning

To determine an unknown radial camera matrix P from the projections $v_i = P X_i$ of 3D points $X_i, i = 1, \dots, n$ we minimize

$$f(P) = \max_i r_i(P X_i) \quad (20)$$

where again $r_i(\cdot)$ is the residual for point x_i and the optimization takes place over the set $x_i^T P X_i > 0, i = 1, \dots, n$.

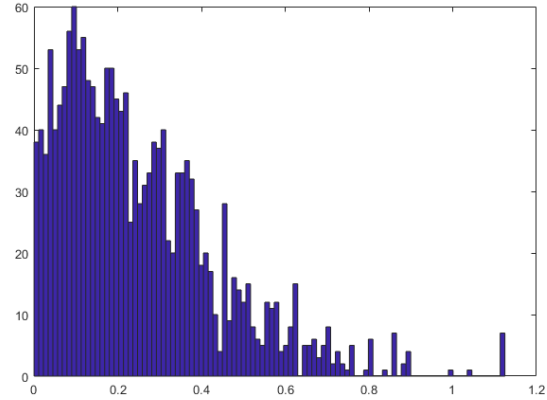


Figure 6. Histogram of residuals (in pixels) when fitting a grid of planar points to the GoPro images in Figure 5.

Note that here we are searching for a general 2×4 camera matrix which has 7 degrees of freedom and hence we need $n \geq 7$. Such a matrix can always be factorized into $K_{2 \times 2} [R_{2 \times 3} \ t_{2 \times 1}]$ where the matrices are as described in Section 2.1. Within our quasiconvex framework it is however not possible to constrain the entries of $K_{2 \times 2}$ since this results in additional non-convex constraints.

To evaluate the quasiconvex formulation for resectioning we consider three datasets (Munsterhof building, Grossmunster and Kirchengen) captured with a DSLR camera with a fisheye lens, which was also used in [28]. See Table 2 for some example images. The scenes are reconstructed with COLMAP [38] using ground truth intrinsic calibration (obtained from offline calibration). From the reconstruction we extract 2D-3D correspondences and for each image, we estimate a 1D radial camera. We then linearly estimate the third row of the projection matrix as well as the focal length (see [26, 27] for details), followed by non-linear refinement of the reprojection error. For the refinement we use the OpenCV Fisheye camera model (with distortion initialized to zero) since this camera model was used in the COLMAP reconstruction and fits the camera well. Figure 7 shows the estimated camera positions and the distribution of the reprojection errors. The estimated cameras align well with the ground truth poses and the reprojection error plots show that we can accurately self-calibrate the intrinsic parameters by bootstrapping the optimization from the radial estimate.

4.4. SfM with Known Rotations

The most interesting problem is perhaps to determine both camera positions t_i and 3D points X_j simultaneously. Here we minimize

$$f(t_1, \dots, t_m, X_1, \dots, X_n) = \max_{ij} r_{ij} \left(\begin{bmatrix} R_i & t_i \end{bmatrix} \begin{bmatrix} X_j \\ 1 \end{bmatrix} \right),$$

where $r_{ij}(\cdot)$ is the residual function (4) for image point x_{ij} and X_j are the regular Cartesian 3D coordinates of point j .

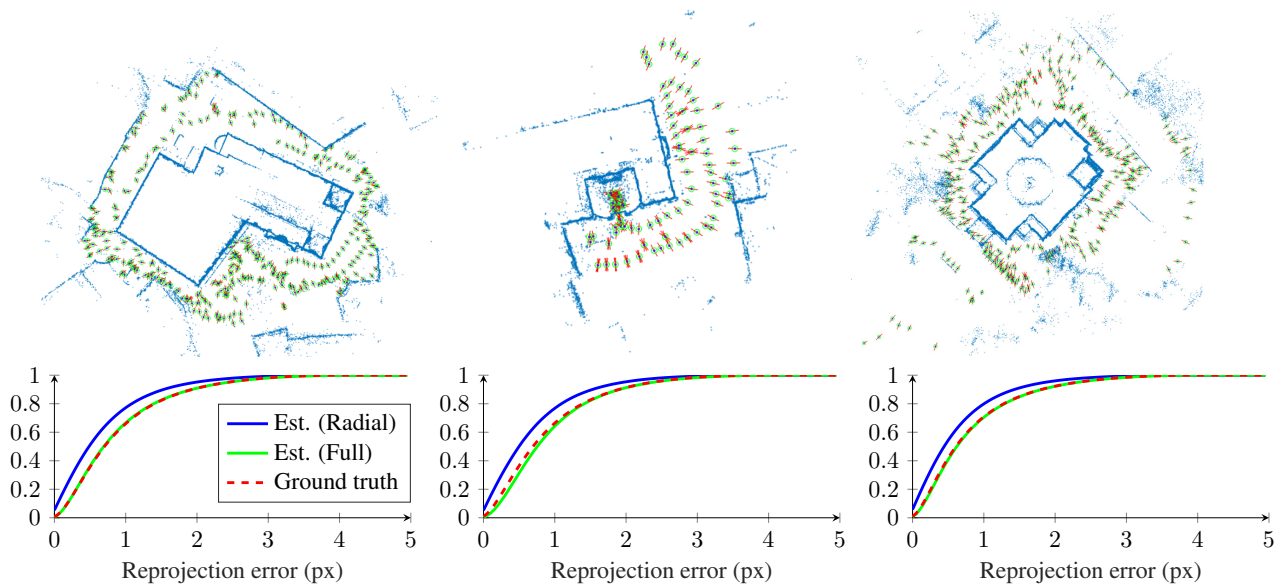


Figure 7. Camera resectioning with the radial camera model. The figure shows results for the three fisheye datasets (Munsterhof building, Grossmunster and Kircheng) considered in Section 4.3. *Top*: Estimated camera positions are shown in green and the ground truth in blue. The principal axes of the radial cameras are shown as red lines. *Bottom*: Cumulative distribution of the reprojection errors. Note that we accurately recover the intrinsic calibration by bootstrapping the optimization with the radial camera model estimate.

While it may seem artificial to assume that the rotational parts of the cameras are known this problem has been shown to be useful in the context of non-sequential reconstruction for regular pinhole models. In contrast to the previous examples this is a large scale application potentially involving millions of variables. Hence efficient methods with optimality guarantees are of interest here.

Table 1 shows the results of solving the known rotation problem. The datasets are taken from [11] and come with 3D reconstructions which we use as ground truth. We added synthetic radial distortion using (18) to these images and solved the problem using the radial camera formulation (with rotations extracted from the provided cameras). To simulate varying distortion we used $(k_1, k_2, k_3) = \sigma(-1.5, 0.75, 0.5)$ with a random $\sigma \in [0, 1]$. One example image with distortion for each data set is shown in Table 1. Column L_∞ of Table 1 shows the maximal reprojection errors (orthogonal point-to-line distance). For comparison we also report maximal reprojection errors obtained when using a regular pinhole camera on the same data.

Table 1 also shows the resulting reconstructions. Recall that the radial camera does not have any well defined camera center. Here the red lines show the viewing directions of the camera. To evaluate the quality of these reconstructions we registered the obtained 3D point cloud to the ground truth reconstruction using a rigid transformation. Note that due to the weaker camera formulation there will be points whose 3D positions are uncertain. In order not to make these affect the registration we use RANSAC with a rela-

tively generous threshold (0.1 times the standard deviation of the ground truth point cloud). We then computed the 3D-RMS errors (in meters), first with all the points, and then only with the inliers. In general the obtained reconstructions are close to the baseline, however, there are typically a number of points that do not make the inlier threshold. Note that the maximal reprojection error over all these data sets is less than 6 pixels (with image size 1296×1936).

For completeness we also test the known rotation problem on three datasets with real radial distortion taken from [28]. Resulting reconstructions are shown in Table 2. In this case the point clouds from [28] had extreme standard deviations and we therefore used a fixed threshold of 1 m for the RANSAC registration.

5. Conclusions

In this paper we have shown that the 1D radial reprojection error is a quasiconvex function. This extends previous work to any camera with radially symmetric distortion, enabling globally optimal minimization of reprojection errors for a more general class of cameras and reconstruction problems by simply solving a series of LP problems. We have presented an extensive empirical evaluation for several different camera models showing that the approach is reliable and provides accurate estimates of both scene geometry and camera positions despite varying and unknown intrinsic camera calibration. This makes it possible to construct structure from motion and SLAM systems based on the presented framework.


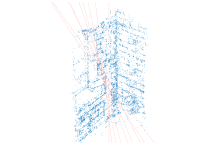



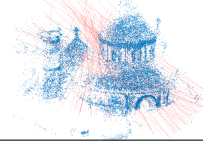

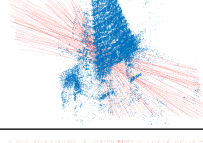
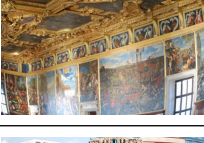
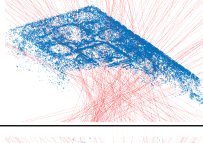

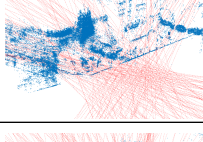

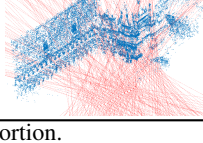

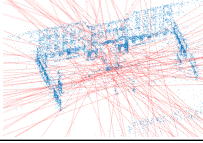

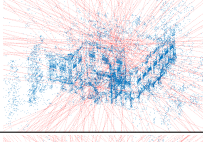

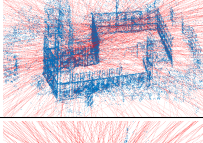

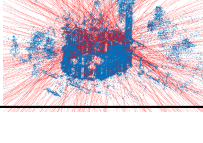
Corner (6558 pts, 9 cams.)						
	L_∞ :	Inlier nr:	3D-RMS (all):	3D-RMS (inliers):		
Radial:	1.8	6545	0.049	0.028		
Pinhole:	27.0	5631	0.137	0.065		
De Guerre (10298 pts, 35 cams.)						
	L_∞ :	Inlier nr:	3D-RMS (all):	3D-RMS (inliers):		
Radial:	3.7	9182	4.20	0.297		
Pinhole:	37.4	6501	3.17	0.508		
Round Church (36467 pts, 92 cams.)						
	L_∞ :	Inlier nr:	3D-RMS (all):	3D-RMS (inliers):		
Radial:	4.6	33124	2.00	0.270		
Pinhole:	83.1	21503	2.87	0.535		
Sri Thendayuthapani (53505 pts, 98 cams.)						
	L_∞ :	Inlier nr:	3D-RMS (all):	3D-RMS (inliers):		
Radial:	5.3	50389	3.75	0.348		
Pinhole:	73.7	35298	2.57	0.735		
Doge Palace Council Chamber (73110 pts, 176 cams.)						
	L_∞ :	Inlier nr:	3D-RMS (all):	3D-RMS (inliers):		
Radial:	5.1	71528	1.40	0.376		
Pinhole:	97.5	45828	4.02	1.087		
Sri Veeramakaliamman (53176 pts, 157 cams.)						
	L_∞ :	Inlier nr:	3D-RMS (all):	3D-RMS (inliers):		
Radial:	5.2	51265	8.06	2.794		
Pinhole:	134.5	29437	35.85	6.222		
Doge Palace Yard (39890 pts, 241 cams.)						
	L_∞ :	Inlier nr:	3D-RMS (all):	3D-RMS (inliers):		
Radial:	5.4	35334	1.583	0.678		
Pinhole:	91.5	28646	11.88	1.651		
Table 1. Results on the known rotation formulation with synthetic and varying radial distortion.						
Fisheye Facade (14411 pts, 148 cams.)						
	L_∞ :	Inlier nr:	3D-RMS (all):	3D-RMS (inliers):		
Radial:	3.5	13364	70100000	0.189		
Pinhole:	-	-	-	-		
Munsterhof Building (27020 pts, 126 cams.)						
	L_∞ :	Inlier nr:	3D-RMS (all):	3D-RMS (inliers):		
Radial:	3.16	25233	189000	0.208		
Pinhole:	-	-	-	-		
Grossmunster (87663 pts, 373 cams.)						
	L_∞ :	Inlier nr:	3D-RMS (all):	3D-RMS (inliers):		
Radial:	3.5	85967	91200	0.105		
Pinhole:	-	-	-	-		
Kirchenge (120993 pts, 369 cams.)						
	L_∞ :	Inlier nr:	3D-RMS (all):	3D-RMS (inliers):		
Radial:	3.2	115290	1.771	0.173		
Pinhole:	-	-	-	-		

Table 2. Results on the known rotation formulation with real radial distortion.

References

- [1] *The MOSEK optimization toolbox for MATLAB manual*. 4
- [2] S. Agarwal, N. Snavely, and S.M. Seitz. Fast algorithms for L_∞ -problems in multiview geometry. In *Computer Vision and Pattern Recognition (CVPR)*, 2008. 4
- [3] F. Arrigoni, L. Magri, B. Rossi, P. Fragneto, and A. Fusiello. Robust absolute rotation estimation via low-rank and sparse matrix decomposition. In *International Conference on 3D Vision (3DV)*, 2014. 2
- [4] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. 3
- [5] Á.P. Bustos, T.-J. Chin, A. Eriksson, and I. Reid. Visual SLAM: Why bundle adjust? In *International Conference on Robotics and Automation (ICRA)*, 2019. 1
- [6] F. Camposco, T. Sattler, and M. Pollefeys. Non-parametric structure-based calibration of radially symmetric cameras. In *International Conference on Computer Vision (ICCV)*, 2015. 2
- [7] L. Carlone, R. Tron, K. Daniilidis, and F. Dellaert. Initialization techniques for 3D SLAM: A survey on rotation estimation and its use in pose graph optimization. In *International Conference on Robotics and Automation (ICRA)*, 2015. 2
- [8] K. Cornelis, F. Verbiest, and L. Van Gool. Drift detection and removal for sequential structure from motion algorithms. *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, 26(10):1249–1259, 2004. 2
- [9] Y. Dai, M. He, and H. Li. Two efficient algorithms for outlier removal in multi-view geometry using L_∞ norm. In *International Conference on Image and Graphics*, 2009. 2
- [10] Z. Dai, Y. Wu, F. Zhang, and H. Wang. A novel fast method for L_∞ problems in multiview geometry. In *European Conference on Computer Vision (ECCV)*, 2012. 4
- [11] O. Enqvist, F. Kahl, and C. Olsson. Non-sequential structure from motion. In *International Conference on Computer Vision Workshops (ICCVW)*, 2011. 1, 2, 5, 7
- [12] O. Enqvist, F. Kahl, C. Olsson, and K. Åström. Global optimization for one-dimensional structure and motion problems. *SIAM J. Imaging Sci.*, 3(4):1075–1095, 2010. 1
- [13] A. Eriksson, C. Olsson, F. Kahl, and T.-J. Chin. Rotation averaging and strong duality. In *Computer Vision and Pattern Recognition (CVPR)*, 2018. 2
- [14] A. Eriksson, C. Olsson, F. Kahl, and T.-J. Chin. Rotation averaging with the chordal distance: Global minimizers and strong duality. *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, 43(1):256–268, 2021. 2
- [15] J. Fredriksson and C. Olsson. Simultaneous multiple rotation averaging using Lagrangian duality. In *Asian Conference on Computer Vision (ACCV)*, 2012. 2
- [16] V. Govindu. Combining two-view constraints for motion estimation. In *Computer Vision and Pattern Recognition (CVPR)*, 2001. 2
- [17] R. Hartley, F. Kahl, C. Olsson, and Y. Seo. Verifying global minima for L_2 minimization problems in multiple view geometry. *International Journal of Computer Vision (IJCV)*, 101(2):288–304, 2012. 2
- [18] R. Hartley and F. Schaffalitzky. L_∞ minimization in geometric reconstruction problems. In *Computer Vision and Pattern Recognition (CVPR)*, 2004. 1
- [19] R. Hartley, J. Trumpf, Y. Dai, and H. Li. Rotation averaging. *International Journal of Computer Vision (IJCV)*, 103(3):267–305, 2013. 2
- [20] F. Kahl. Multiple view geometry and the L_∞ -norm. In *International Conference on Computer Vision (ICCV)*, pages 1002–1009, Beijing, China, 2005. 1
- [21] F. Kahl and J. August. Multiview reconstruction of space curves. In *International Conference on Computer Vision (ICCV)*, 2003.
- [22] F. Kahl and R. Hartley. Multiple-view geometry under the L_∞ -norm. *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, 30:1603–17, 09 2008. 1, 2
- [23] F. Kahl, A. Heyden, and L. Quan. Minimal projective reconstruction including missing data. *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, 23(4):418–424, 2001.
- [24] Q. Ke and T. Kanade. Quasiconvex optimization for robust geometric reconstruction. *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, 29(10):1834–1847, 2007. 1
- [25] J.-H. Kim, Y. Dai, H. Li, X. Du, and J. Kim. Multi-view 3D reconstruction from uncalibrated radially-symmetric cameras. In *International Conference on Computer Vision (ICCV)*, 2013. 2
- [26] Z. Kukeleva, M. Bujnak, and T. Pajdla. Real-time solution to the absolute pose problem with unknown radial distortion and focal length. In *International Conference on Computer Vision (ICCV)*, 2013. 2, 6
- [27] V. Larsson, T. Sattler, Z. Kukeleva, and M. Pollefeys. Revisiting radial distortion absolute pose. In *International Conference on Computer Vision (ICCV)*, 2019. 2, 6
- [28] V. Larsson, N. Zobernig, K. Taskin, and M. Pollefeys. Calibration-free structure-from-motion with calibrated radial trifocal tensors. In *European Conference on Computer Vision (ECCV)*, 2020. 2, 6, 7
- [29] H. Li. A practical algorithm for L_∞ triangulation with outliers. In *Computer Vision and Pattern Recognition (CVPR)*, 2007. 2
- [30] Y. Lin, V. Larsson, M. Geppert, Z. Kukeleva, M. Pollefeys, and T. Sattler. Infrastructure-based multi-camera calibration using radial projections. In *European Conference on Computer Vision (ECCV)*, 2020. 2
- [31] D. Martinec and T. Pajdla. Robust rotation and translation estimation in multiview reconstruction. In *Computer Vision and Pattern Recognition (CVPR)*, 2007. 1, 2
- [32] P. Moulon, P. Monasse, and R. Marlet. Global Fusion of Relative Motions for Robust, Accurate and Scalable Structure from Motion. In *International Conference on Computer Vision (ICCV)*, 2013. 1
- [33] D. Nistér. Reconstruction from uncalibrated sequences with a hierarchy of trifocal tensors. In *European Conference on Computer Vision (ECCV)*, 2000. 2
- [34] C. Olsson, O. Enqvist, and F. Kahl. A polynomial-time bound for matching and registration with outliers. In *Computer Vision and Pattern Recognition (CVPR)*, 2008. 2

- [35] C. Olsson and A. Eriksson. Triangulating a plane. In *Scandinavian Conference on Image Analysis (SCIA)*, 2011. [1](#)
- [36] C. Olsson, A. Eriksson, and R. Hartley. Outlier removal using duality. In *Computer Vision and Pattern Recognition (CVPR)*, 2010. [2](#)
- [37] C. Olsson, F. Kahl, and M. Oskarsson. The registration problem revisited: Optimal solutions from points, lines and planes. In *Computer Vision and Pattern Recognition (CVPR)*, 2006.
- [38] J.L. Schönberger and J.-M. Frahm. Structure-from-motion revisited. In *Computer Vision and Pattern Recognition (CVPR)*, 2016. [6](#)
- [39] Y. Seo, H. Lee, and S.W. Lee. Outlier removal by convex optimization for L_∞ -infinity approaches. In *Advances in Image and Video Technology*, 2009. [2](#)
- [40] K. Sim and R. Hartley. Recovering camera motion using l_∞ minimization. In *Computer Vision and Pattern Recognition (CVPR)*, 2006. [1](#)
- [41] K. Sim and R. Hartley. Removing outliers using the L_∞ norm. In *Computer Vision and Pattern Recognition (CVPR)*, 2006. [2](#)
- [42] S. Thirithala and M. Pollefeys. Radial multi-focal tensors. *International Journal of Computer Vision (IJCV)*, 96:195–211, 2012. [1](#), [2](#), [6](#)
- [43] R. Tron, B. Afsari, and R. Vidal. Intrinsic consensus on $SO(3)$ with almost-global convergence. In *IEEE Conference on Decision and Control (CDC)*, 2012. [2](#)
- [44] R. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Journal on Robotics and Automation*, 3(4):323–344, August 1987. [2](#)
- [45] J. Yu, A. Eriksson, T.-J. Chin, and D. Suter. An adversarial optimization approach to efficient outlier removal. In *International Conference on Computer Vision (ICCV)*, 2011. [2](#)
- [46] Q. Zhang, T.-J. Chin, and H.M. Le. A fast resection-intersection method for the known rotation problem. In *Computer Vision and Pattern Recognition (CVPR)*, 2018. [1](#), [4](#)
- [47] Q. Zhang, T.-J. Chin, and D. Suter. Quasiconvex plane sweep for triangulation with outliers. In *International Conference on Computer Vision (ICCV)*, 2017. [2](#)