Calibration of Axial Fisheye Cameras Through Generic Virtual Central Models

Pierre-André Brousseau and Sébastien Roy
Département d'informatique et de recherche opérationnelle
Université de Montréal
pierre-andre.brousseau@umontreal.ca, roys@iro.umontreal.ca

Abstract

Fisheye cameras are notoriously hard to calibrate using traditional plane-based methods. This paper proposes a new calibration method for large field of view cameras. Similarly to planar calibration, it relies on multiple images of a planar calibration grid with dense correspondences, typically obtained using structured light. By relying on the grids themselves instead of the distorted image plane, we can build a rectilinear Generic Virtual Central (GVC) camera. Instead of relying on a single GVC camera, our method proposes a selection of multiple GVC cameras which can cover any field of view and be trivially aligned to provide a very accurate generic central model. We demonstrate that this approach can directly model axial cameras, assuming the distortion center is located on the camera axis. Experimental validation is provided on both synthetic and real fisheye cameras featuring up to a 280° *field of view. To our knowledge, this is one of the only practical methods to calibrate axial cameras.*

1. Introduction

Fisheye lens calibration is a type of calibration which is becoming very common since the emergence of low cost high quality fisheye lenses for applications such as immersive imaging as well as industrial and automotive applications. These lenses can feature very large fields of view (FoV). Modern fisheyes can see as much as 280° (see Fig. 7), with large amount of radial distortion and significant axial displacement of the optical center, essentially making them non-single viewpoint, which do not comply with standard lens models [5, 16]. Simple OpenCv tools fail completely at calibrating these kinds of cameras and current methods are either planar and limited to 180° or non-planar which makes them impractical. This paper proposes a calibration method adapted to these kinds of lenses, which are described as generic axial cameras. We tackle a slightly more constrained version of the axial model, which we identify as *Generic Quasi-Central* camera.

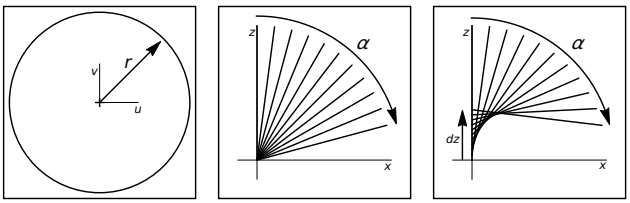


Figure 1: Central and Axial models. In the fisheye image (left), points at increased radius r correspond to increased ray angles α in a central camera (middle) as well as increased center displacement dz in a quasi-central camera (right).

Similarly to traditional planar calibration, the new method makes use of multiple dense correspondences of planar grids. It is known that a grid can be used as a *virtual image* to build a *Generic Virtual Central* cameras, which is free of distortion [2]. We propose to build multiple such virtual cameras, then calibrate and align them to provide an accurate solution to the full central camera model.

In order to recover the axial displacement of the optical center, an approach is presented for the generalization of the generic central model into the quasi-central model. This approach also relies on multiple GVC models, confirming their high accuracy. The paper is organized as follows. In Section 2, previous work and generic camera models are presented. The Generic Quasi-Central camera model is presented in Section 3. The calibration method for multiple Generic Virtual Central cameras is presented in Section 4, followed by experimental validation and results in Section 5.

2. Previous Work

There has been a lot of work on the calibration of fisheye lenses [7, 18]. Most of this work relies on global single viewpoint models. In this paper, the focus is on

generic and non-single viewpoint models.

Camera models can be divided in various ways. One of such way is to classify models according to how broad the impact of a set of parameters is in the image. The broader the impact the more global a model is. In this way, models can be classified as global, local and discrete classes. Discrete models are models where each individual pixel of the image requires its individual parameters. One such model is the *ray-based* model where it is required to explicitly represent each ray of a camera [18]. In this way, the generic imaging model as defined in [9, 17] is a discrete ray-based model where the camera captures images defined as a set of individual pixels each associated to a single ray of light in the 3D world. This unconstrained set of projection rays, each with its own direction and position in 3D, and associated image pixels, constitute the camera model and the expected result for the calibration algorithm [9, 15]. One possible calibration method is to solve 3D points along projection rays using 3D calibration objects. In the case of unknown pose of the calibration object, it is possible with 3 views of the object to solve a system of linear equations for the unknown poses and the projection rays in the case of a generic non-central camera. This system relies only on the collinearity constraints of points between the three views [9, 17]. The term non-central refers to the fact that the projection rays are unconstrained. Depending on the spatial distribution of the projection rays, a hierarchy of camera models emerges. Three of them are: central cameras where all projection rays intersect a single point (the optical center), axial cameras where all projection rays intersect a single line (the camera axis), and non-central cameras where the projection are entirely unconstrained [9, 15].

2.1. Generic Central Cameras

In the case of generic central cameras, calibration using planar calibration grids was demonstrated to work in practical cases [1, 8]. Using three calibration grids that are dense, the optical center as well as the pose of all three grids are solved. A 183° field of view fisheye camera was successfully calibrated using a total of 12 calibration grids to cover the whole field of view [8]. It is shown that a minimum of four points each matching three planes are required to solve plane poses. Furthermore, the linear equation system relies on collinearity of matching projection rays as well as cheirality (solved 3d points are located on the same side of the optical center) [8]. Our approach requires two planes per match, its linear system is more straightforward, and can generalize to axial models.

An example of image undistorsion is provided in [1] by intersecting a single plane with the previously computed projection rays. With a lens of only 60° FoV, the camera is probably very close to central, and its comparison to our axial and large FoV examples is not applicable.

The concept of using a grid as a virtual image for the purpose of a rectilinear planar calibration has been proposed in [2]. Once a grid is selected as the virtual image, if can be calibrated with all other grids, effectively providing the pose of all grids. However, using a single virtual image is limited to grids inside a small FoV. All grids not included must have their pose solved separately with an alternate algorithm which tends to be unstable and accumulate errors toward the edge of the FoV. This limitation is removed in our approach by introducing multiple generic virtual cameras and allows any field of view to be calibrated, up to a full sphere.

2.2. Generic Axial Cameras

In the context of generic axial cameras, calibration using planar calibration grids was demonstrated to work on simulated data [9, 11]. In this model, one has to solve for the camera axis and the poses of all grids. To establish the trifocal tensor to be solved, each image point must match three calibration grids and satisfy additional constraints: the camera axis must intersect all the projection rays, and the principal point, located at the distortion center, must be known to allow a prealignment of the planes. Subsequently, a bundle adjustment is used to minimize the distance between 3D points and their projection on their respective projection rays[11]. In the case of the spherical catadioptric camera, even though this camera seems to be axial, the solved rays are distributed in a very small region approximated as the optical center. The resolved poses seem to show some inconsistencies with the actual sequence of capture [10].

2.3. Dense planar correspondences

Calibration of generic cameras, since they are ray-based, require dense matches between the image and calibration grids. In practice, dense correspondences can be obtained by interpolating positions from closely spaced targets [8, 17], or structured light matching on a flat LCD monitor [1, 3, 12, 13]. The correspondences are provided in the form of a lookup table, where each pixel (u, v) of the fisheye image provides a match (x, y) on the planar grid. As illustrated in Fig. 2, (x, y) matches are represented as red and green, respectively. Blue is used as a mask.

3. Generic Quasi-Central Cameras

The goal of our method is to calibrate the generic axial model, as defined in [11], but this model is underconstrained. It allows a single 3d point to have two or more rays projecting to the image, making it impossible to inverse the image formation model. In [14], the term

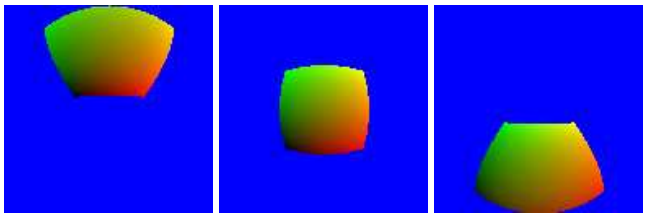


Figure 2: Examples of dense correspondence maps of planar grids obtained with structured light. In the 16-bit image, red and green encode x and y coordinates, respectively.

quasi-central is introduced for models where light rays intersect near a single optical center, as a way to approximate non-central cameras with central cameras.

In this paper, the term *quasi-central* defines a model where camera rays intersect on a common camera axis, and where the displacement along this axis is a monotonic function of the ray angle with the axis. This ensures the convergence of our projection model as well as the inversibility of the image formation model. The quasi-central model, depicted in Fig. 1(right), is thus somewhere between central and fully axial and applies to real fisheye cameras.

3.1. Image Formation Model

A generic axial camera model can be represented as a lookup table (LUT) assigning to each fisheye image position (u,v) a projection ray with parameters (α,β,dz). As illustrated in Fig. 3, the relationship between (u,v) and the ray (α,β,dz) is simple. A central camera features $dz=0$ while an axial or quasi-central camera provides an additional parameter dz, the vertical displacement of the optical center, which more accurately represents real fisheye lenses. A quasi-central camera requires a monotonic relation between dz and α, and full axial camera has no restriction on dz.

Typically, an equidistant fisheye has $r \propto \alpha$, and a rectilinear lens has $r \propto \tan\alpha$. Our model does not enforce a particular relationship between r and α or dz as it relies on a LUT to explicitly store individual ray information. Furthermore, it does not impose radial symmetry around the principal point.

As illustrated in Fig. 3, the projection of a world point (x,y,z) to a camera ray (α,β,dz) is

$$
\begin{aligned}
(\alpha,\beta,dz) &= M(u,v)\\
(u,v,dz) &= M^{-1}(\alpha,\beta)\\
(\alpha,\beta) &= f(x,y,z)\\
(u,v,dz)^k &= \underset{k=0,1,\ldots}{\mathbf{FP}} \left| \begin{array}{ll} M^{-1}(f(x,y,z)) & k=0\\ M^{-1}(f(x,y,z-dz^{k-1})) & k>0 \end{array} \right.
\end{aligned}
\tag{1}
$$

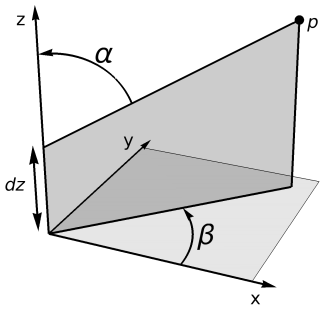


Figure 3: Generic quasi-central model. The projection ray of a 3D point $p=(x,y,z)$ is represented as (α,β,dz). We have $(\alpha,\beta)=f(x,y,z-dz)$.

where M is the image formation model, M^{-1} is its inverse, f is the projection model such that

$$
\begin{aligned}
(\alpha,\beta) = f(x,y,z) = (&\Theta(\langle x,y,z\rangle,\langle 0,0,1\rangle),\\
&\Theta(\langle x,y,0\rangle,\langle 1,0,0\rangle))
\end{aligned}
\tag{2}
$$

and **FP** is the fixed point function starting at $k=0$ which iterates until the result stops changing. We define $\Theta(\mathbf{a},\mathbf{b})$ as the angle between vectors $\mathbf{a}$ and $\mathbf{b}$. Notice that for a central camera, dz is always 0 so the iteration stops at $k=0$. For quasi-central, dz is unknown so it is initialized at 0. Its value will converge to the correct dz, and thus the correct (u,v) in the image, after a few iterations, assuming dz is a monotonic function of α. This assumption is realistic in practice for real lenses.

4. Calibrating multiple GVC cameras

Fig. 4 illustrates generic central and non-central camera models, where the image plane is non-planar to represent radial distortion. These models are incompatible with a linear planar calibration approach. As seen at the right of Fig. 4, [2] proposes to use one of the calibration grids, here depicted in red, as a *virtual image plane*, which we refer to as a Generic Virtual Central camera.

Instead of solving a single GVC for a full fisheye as in [2], we propose to use multiple GVC cameras, each using a minimal number of calibration grids, to represent the full generic central camera. This central camera calibration serves as the basis for the generalization to the generic quasi-central model.

4.1. Planar calibration for GVC model

As described in [2], *active grids* provide dense correspondence maps for the purpose of calibration. Given three grids, it is possible to find their respective poses using planar calibration [19]. Fig. 5 illustrates this for three

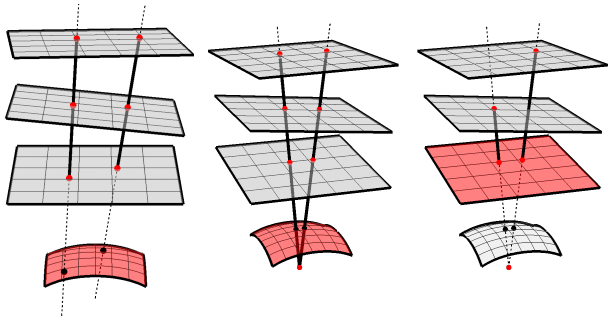


Figure 4: Left) Generic Non-Central camera. Middle) Generic Central camera (rays intersect at optical center). Right) Generic Virtual Central camera, where the plane in red takes the role of a virtual image plane.

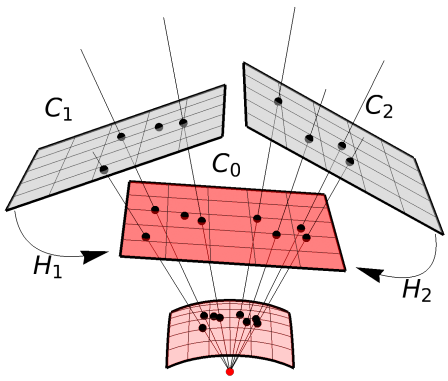


Figure 5: Calibration grid C_0 will act as the image plane in a planar calibration step with checkerboards C_1 and C_2, yielding the pose of C_1 and C_2 and the internal parameters ($\mathbf{K}$) for C_0.

grids C_0, C_1 and C_2, from which C_0 is designated *virtual image*. Notice that it is important to ensure that the relative pose of a grid with respect to the virtual image is never a pure translation, as this would yield a degenerate configuration and a failed calibration [6].

From the homographies relating C_0 to C_1 and C_2, planar calibration will provide the pose of C_1 and C_2 as well as the internal parameters $\mathbf{K}_0$. Assuming the camera is central, grid C_0 shares its optical center with the real camera, but has its own pose in the world. Because it features a known pixel ratio and size, we can derive its pose from the internal parameters of $\mathbf{K}_0$, which is of the form

$$\mathbf{K}_0 = \begin{pmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{pmatrix} \tag{3}$$

Because the projection model assumes that the optical center is at the origin of the world $(0,0,0)^\top$, we can infer its origin to be at $(-c_x,-c_y,f)^\top$ and its orientation aligned with the world.

4.2. Grouping grids into GVC

In this planar calibration scheme, three planes are the minimum required as one of them is the virtual image plane and the two others define the homographies. More planes could be used, as it is the case in [2], but one must realize that rectilinear cameras are not well defined at large field of views (and certainly cannot reach 180° FoV), even without radial distortion. Moreover, it is most probable that grids oriented closer to the far edges of the fisheye FoV will have few correspondences with the virtual image plane, making the homography unstable. This is why we propose to always use three planes for a GVC model, and then solve multiple GVC models which are subsequently aligned together. This ensures optimal planar calibration with reasonable FoV for each GVC camera, while allowing very large FoV once grouped together.

The correspondence maps need to be separated into triplets and some correspondence maps must be part of more than one triplet. In order to do that, we construct a fully connected graph between every correspondence map and every other correspondence map were the edge weights are the negative number of matching correspondences between both maps. Subsequently, this graph can be reduced by taking its minimum spanning tree. The plane with the most edges (or number of matches in the case of a tie) is determined as the global reference plane and is marked as a visited node. From the global reference plane, two neighboring nodes are selected and this triplet of calibration planes is assembled. The selected nodes are then marked as visited. In further iterations, each new triplet needs to count one visited and two unvisited neighbors. Each new visited node is marked as such. Special cases can happen, most notably when a visited node has only a single unvisited neighboring node. In such a case, the triplet will be composed of the visited node, the unvisited neighboring node and a previously visited neighboring node. In this way, all triplets of planes are assembled into a collection of GVC cameras which are then calibrated independently. For each triplet, any grid can be used as the virtual image plane with the one exception being the first triplet where the global reference plane needs to be the virtual image.

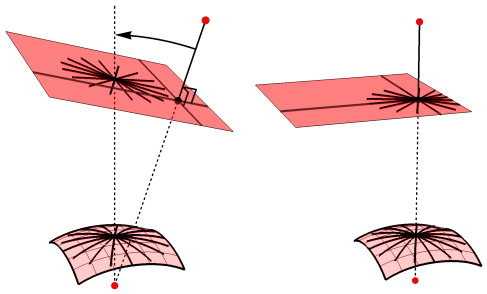


Figure 6: Principal point alignment. Left) The optical axis of the GVC is not aligned with the fisheye axis. Right) The principal point allows to align the GVC axis.

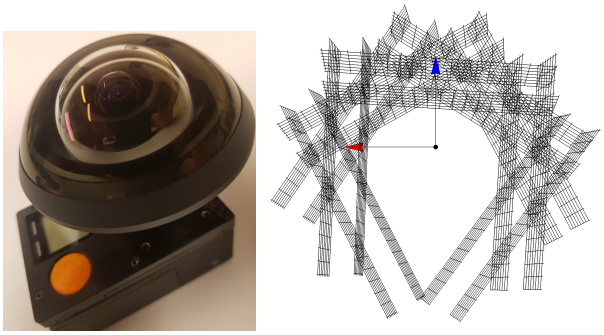


Figure 7: Left) Entaniya 280° Fisheye Lens mounted on a GoPro HERO 4. Right) Recovered poses of the calibration grids around the fisheye. Notice the near and far sets of planes.

4.3. Merging multiple GVC cameras

By calibrating all GVC cameras independently, we obtain the positions and orientations of all the grids but in their respective GVC triplet reference. Some planes belong to more than one triplet and were calibrated multiple times thus have multiple poses. Since each plane has only a single pose in the world, they can be rigidly moved into alignment. The reference systems are merged using only these planes as landmarks, not their centers of projection. This provides a way to align all triplets in the world. For a central camera, the multiple centers of projections will be a single point. For a quasi-central camera, each center of projection will be axially displaced along the camera axis, as each relates to a local region of the FoV. After the merge, a single grid is kept as the reference plane. The optical axis of the merged GVC camera is not yet defined in that reference system. Although, the displacement of the centers of projection describes the camera axis, in practice these displacements are small, even nonexistent for a central camera, they are not reliable to estimate the camera axis. Finding this axis is addressed in the following section.

4.4. Aligning with the Principal point

By finding the center of distortion, it is possible to align the optical axis of the merged GVC camera to the optical axis of the fisheye camera. This point could be different than the principal point [4], but in practice we consider them to be the same [19].

The center of distortion is defined as the point in the fisheye camera with the least distortion, where straight lines in the world should remain straight in the camera image, as illustrated in Fig. 6. We find the center of distortion by locating straight lines in the fisheye image which correspond to collinear matches in their planar grid. The merged GVC camera is then rotated such that the virtual principal point, corresponding to the fisheye image principal point, lies on its Z-axis. In theory, this alignment is not absolutely necessary, except to simplify the representation of radial symmetry and axial displacement.

4.5. Generalizing to quasi-central camera model

For the axial fisheye cameras, the above method has a severe weakness because homographies assume that all rays intersect at a single point. In the case of axial cameras that is in fact false, since there will be multiple optical centers distributed along a common axis. In this situation, the homographies will in fact provide a single optical center representing the average of all these optical centers.

In order to generalize from central to quasi-central, an optimization procedure is devised. Let us consider the poses of the planar calibration grids $C_i : (\mathbf{R}_i, \mathbf{T}_i)$ and the calibration grid set defined as $\Pi = \{C_1, \ldots, C_n\}$. For given coordinates (u, v) in the fisheye camera image, the 3D point $\mathbf{p}^i_{u,v}$ located on plane i corresponding to the fisheye image point (u, v) is extracted directly from the lookup tables $LUT^i(u, v)$.

In this way the 3D point in the world centered around the optical center is defined as

$$\mathbf{q}^i_{u,v} = \mathbf{R}_i \mathbf{p}^i_{u,v} + \mathbf{T}_i \tag{4}$$

This leads on to defining a line between two points i and j such as

$$L^{i,j}_{u,v} = \left\langle \mathbf{q}^i_{u,v}, \mathbf{q}^j_{u,v} \right\rangle \tag{5}$$

One special line is the camera axis L_z defined as the Z-axis. Furthermore, we rely on two functions, $\Theta(L_a, L_b)$ (see Eq. 2) and $d(L_a, L_b)$, to provide the angle and distance, respectively, between two lines L_a and L_b. The

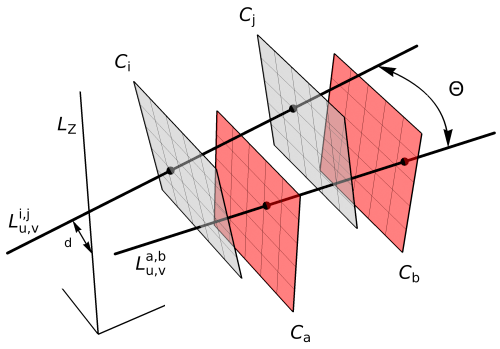


Figure 8: Minimization of planes i and j must reduce distance d with the Z-axis, as well as Θ for parallelism with the matching line from the same (u, v).

minimization problem is formulated as

$$\arg\min_{\Pi} \sum_{\{i,j\}}^{|\Omega|} \sum_{u,v} d(L^{i,j}_{u,v}, L_Z) + \gamma \sum_{\{i,j\}}^{|\Omega|} \sum_{\{a,b\}}^{|\Omega|} \sum_{u,v} \Theta(L^{i,j}_{u,v}, L^{a,b}_{u,v}) \tag{6}$$

where Ω is the set of pairs of planes such that planes i and j have correspondences and are separated by a sufficient distance and where γ is a weighting parameter. This distance is required to ensure stability because lines generated between overlapping planes are very unstable. This has an important consequence in practice as it requires two sets of calibration grids, one closer and one further away from the camera (see Fig. 7(right)). The cost function is illustrated in Fig. 8 for two pairs of planes. The planes i and j must be positioned so the distance d between line $L^{i,j}_{u,v}$ and the Z-axis is 0. Also, the angle Θ must be minimized to ensure parallelism with a line $L^{a,b}_{u,v}$ related to the same fisheye image point (u, v).

This above minimization problem is presented in two separate terms which minimize different objectives. The first aims to enforce that all lines intersect the camera axis. The second aims to enforce parallelism between lines corresponding to a common (u, v) point in the fisheye image, since this is the basic assumption of the generic camera model. This second term is required since the optimization does not solve all parameters globally, so multiple conflicting solutions for plane poses are possible.

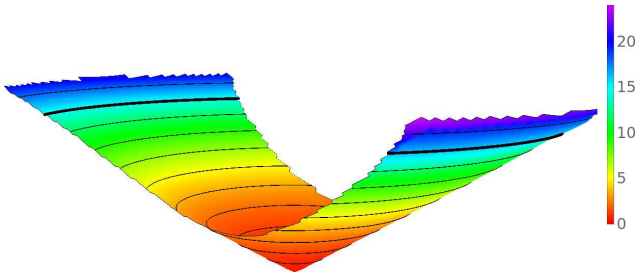


Figure 9: 3D view of the lookup table for the Synthetic Axial camera with parabolic displacement. Z-axis is α with isocontours at every 10° (90° is in bold). Axial displacement dz is color coded from 0 to 24 plane pixels (see legend).

5. Experimental results

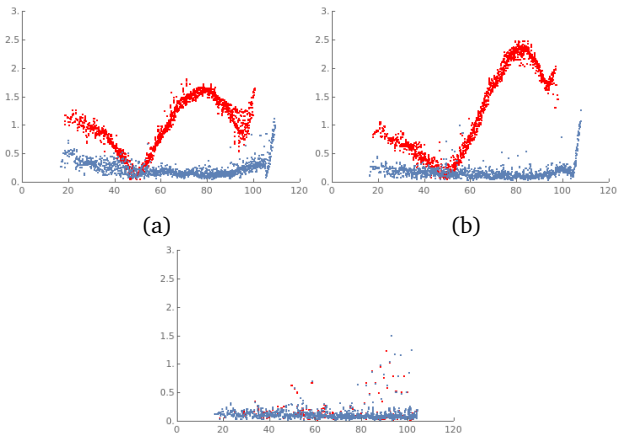
We provide two different types of experimental results. First, synthetically generated fisheye cameras, both central and quasi-central, are used to validate our calibration method. Second, calibration results of a real large FoV camera are provided demonstrating practical use of the quasi-central model and calibration.

Comparing experimental results with other methods proved to be challenging. In [10, 11], they calibrate an axial camera model where the camera is in fact two central pinhole cameras in a stereo configuration which is not quasi-central. Since our method requires quasi-central, no comparison can be performed. Their further work in [8] constrains their model to a central camera but its 183° FoV cannot be compared to ultra-wide. Finally, the method in [17] describes a framework for generic calibration and is theoretical in nature. Although they establish the groundwork for truly generic camera models, in the case of fisheye cameras, the results shown are preliminary as only a very limited FoV is calibrated.

5.1. Synthetic Calibration

The synthetic fisheyes have a 220° field of view with three different axial displacement modalities, respectively a linear axial displacement of $\frac{20}{r}$, a parabolic displacement of $\frac{20}{r^2}$ and no axial displacement (a central camera). For each fisheye cameras, the projection rays associated with each pixel were generated. Theses rays are then intersected with various planar grids at various poses to generate dense correspondence maps (see Fig. 2). The calibration is done using only these correspondence maps.

Image formation model. The result of Fig. 9 is the recovered image formation model $M(u, v)$ (see Eq. 1) for the synthetic camera with the parabolic axial displacement. We observe that the result closely matches the camera specification. The cone shape indicates an equidistant



(a) (b) (c)

Figure 10: Reprojection error in pixels, as a function of α, for calibrating a central model (in red) and a quasi-central model (in blue). Cameras are fisheye 220° with various axial displacements: A) linear, B) parabolic, C) none.

fisheye and the axial displacement is recovered accurately as parabolic.

Reprojection error. Fig. 10 presents the reprojection error for all three synthetic cameras, obtained by reprojecting all the points of a single calibration grid into the fisheye image. The selected grid is the same for all 3 cameras and was chosen because it fills a large portion of the FoV and includes areas where the radial distortion and axial displacement are highest.

When applying central model calibration on axial cameras (red curves in Fig. 10a and 10b) we observe high reprojection errors with two dipping points, indicative of the *average pose* recovered by this inadequate model. This is similar to the error observed when fitting a line to a portion of a second-degree curve. The fit will be wrong, except maybe at two points.

When applying quasi-central model on the same cameras, shown as blue curves in Fig. 10a and 10b, the curves are near constant at a low reprojection error (0.5 pixels) which also corresponds to the lowest point on the red curve. This demonstrates that our proposed quasi-central calibration method performs as intended.

For the central camera, depicted in Fig. 10c, both central and quasi-central calibration feature low reprojection errors, demonstrating that both methods perform as expected.

Table. 1 provides a global performance estimate at a glance. For synthetic as well as real cameras, calibrations with low reprojection errors are achieved. Moreover, using the axial model on axial cameras further reduce the error.

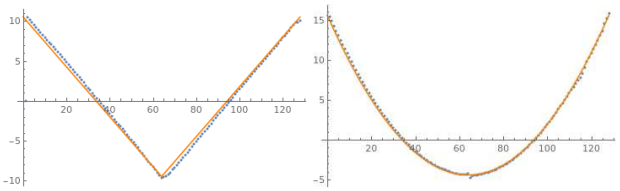


Figure 11: Axial displacement for the synthetic axial cameras, linear (left) and parabolic (right). In orange are the true curves and in blue are the axial displacements recovered across a vertical slice of the fisheye image.

		Field of View Region			
Camera	**Calib**	**Center**	**Middle**	**Outer**	**Full**
Synthetic Central	Central	0.2	0.1	0.1	0.1
	Axial	0.2	0.1	0.1	0.1
Synthetic Axial Linear	Central	0.5	0.8	1.2	0.8
	Axial	0.4	0.2	0.2	0.3
Synthetic Axial Parabolic	Central	0.3	0.9	1.9	1.1
	Axial	0.2	0.4	0.2	0.4
Real Entaniya 280°	Central	4.2	4.9	6.4	5.0
	Axial	4.5	3.0	4.4	3.8

Table 1: Reprojection error in pixels, averaged for different regions of the field of view, for central and axial calibrations of the synthetic and real cameras. We observe that axial calibration yields lower error for axial cameras.

Axial displacement. Fig. 11 illustrates that the axial displacement is correctly recovered by the quasi-central model.

5.2. Real Camera Calibration

The experimental setup for the fisheye calibration uses a LG29UM57-P LCD monitor with a resolution of 2560x1080 pixels and a 0.2651mm dot pitch. The camera is a GoPro HERO 4, with an Entaniya 280° Fisheye lens attached, in 2.7K 4:3 Ultra Wide video resolution mode which results in a 2704 × 2028 full frame image (see Fig. 7(left)). The camera is mounted on a pan-tilt unit located at about 20cm from a monitor acting as an active calibration grid, and rotated to provide multiple calibration grids for different points of view. A second set of acquisitions is performed at 30cm from the monitor. Note that although the amount of rotation between planes and the approximate distances between the camera and the monitor are known, the calibration is done solely using the correspondence maps.

Image formation model. Fig. 12 is the recovered image formation model $M(u, v)$ (see Eq. 1) for the Entaniya 280°. We observe that the α angles are recovered well over the full field of view and indicate near equisolid geometry. The axial displacement is significant and also recovered well. It is compatible with the lens diameter of

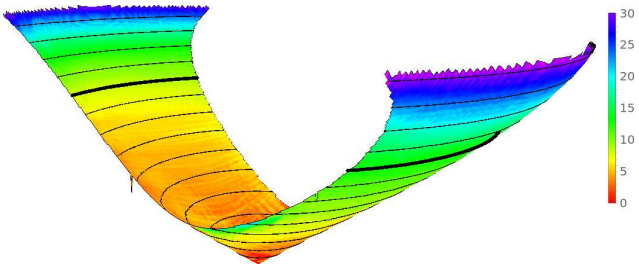


Figure 12: 3D view of the lookup table for the Entaniya 280°. Z-axis is α with isocontours at every 10° (90° is in bold). Axial displacement dz is color coded from 0 to 35mm (see legend).

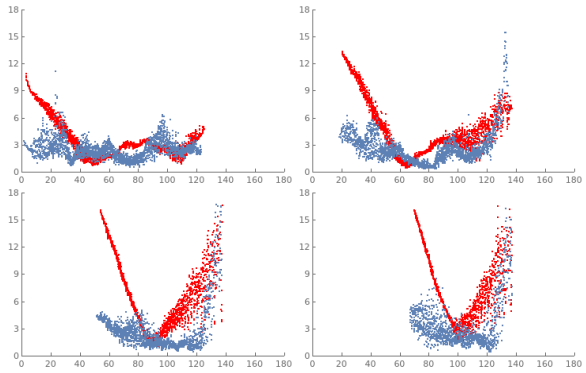


Figure 13: Reprojection error in pixels, as a function of α, of four calibration grids. Curves results from calibrating a central model (in red) and a quasi-central model (in blue) for a 280° fisheye.

64mm and the maximum incident angle of 50° above the horizon.

Reprojection error. Similarly to the synthetic camera cases, the reprojection error is computed on different calibration grids along the field of view. Four of them were chosen as representative of the expected results and together they cover the whole FoV.

Overall, we observe in Fig. 13 that for very wide FoV fisheyes, modeling the axial displacement greatly improves calibration. In all cases, the reprojection error for the quasi-central calibration can be averaged to 3 pixels or less, while the central calibration is up to 15 pixels. Notice the "V" shapes in the red curves which are steeper with increasing angle due to the averaging effect of the central calibration. This means that the farther away from the camera axis you are, the worst a central calibration performs, and the more essential quasi-central calibration becomes.

Axial displacement. To better assess the quality of the quasi-central calibration, Fig. 14 shows the recovered axial displacement in mm across the camera image. This illustrates that near the edges, the optical center has moved by up to 30mm from its initial position and confirms why a central approximation to such a camera is inaccurate. Notice that the curve is not symmetric around the principal point. We believe that this is a result of reaching an inaccurate local minimum during the optimization step. This was observed only for specific grids, suggesting some inaccuracy in the structured light matching.

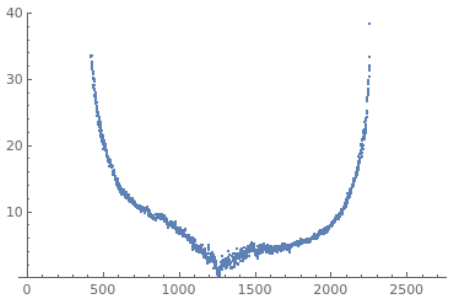


Figure 14: Axial displacement in mm for the 280° camera, recovered across a horizontal slice of the fisheye image.

6. Conclusion

This paper presented a new calibration method for fisheye cameras, with an underlying Generic Quasi-Central camera model. It can bypass any image distortion by using calibration planes as virtual images, the Generic Virtual Central cameras, which are perfectly rectilinear and can be solved with simple planar calibrations. The resulting calibration allows very large fields of view. This model is then generalized to recover axial optical center displacements. Although the optimization scheme which recovers grids poses is simple and could be further improved, results on a real lens demonstrate not only that axial calibration works, but that it significantly improves the reprojection error. In the future, we expect that the Quasi-Central model and multiple GVC cameras will find more uses such as generating perfect image rectifications, generalize to other models, improve fisheye image stitching for immersive imaging and enhance 3D reconstruction from the triangulation of fisheye images.

References

[1] Filippo Bergamasco, Luca Cosmo, Andrea Gasparetto, Andrea Albarelli, and Andrea Torsello. Parameter-free lens distortion calibration of central cameras. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3847–3855, 2017.

[2] Aubrey K Dunne, John Mallon, and Paul F Whelan. Efficient generic calibration method for general cameras with single centre of projection. *Computer Vision and Image Understanding*, 114(2):220–233, 2010.

[3] Chaima El Asmi and Sébastien Roy. Fast unsynchronized unstructured light. In *2018 15th Conference on Computer and Robot Vision (CRV)*, pages 277–284. IEEE, 2018.
[4] Richard Hartley and Sing Bing Kang. Parameter-free radial distortion correction with center of distortion estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(8):1309–1321, 2007.
[5] Katsushi Ikeuchi. *Computer vision: A reference guide*. Springer Publishing Company, Incorporated, 2014.
[6] H Malm and A Heyden. Plane-based calibration: The case of pure translation. In *Proc. Int. Workshop on Machine Vision Applications*, 2002.
[7] Luis Puig, Jesús Bermúdez, Peter Sturm, and José Jesús Guerrero. Calibration of omnidirectional cameras in practice: A comparison of methods. *Computer Vision and Image Understanding*, 116(1):120–137, 2012.
[8] Srikumar Ramalingam and Peter Sturm. Minimal solutions for generic imaging models. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. IEEE, 2008.
[9] Srikumar Ramalingam and Peter Sturm. A unifying model for camera calibration. *IEEE transactions on pattern analysis and machine intelligence*, 39(7):1309–1319, 2017.
[10] Srikumar Ramalingam, Peter Sturm, and Suresh K Lodha. *Generic calibration of axial cameras*. PhD thesis, INRIA, 2005.
[11] Srikumar Ramalingam, Peter Sturm, and Suresh K Lodha. Theory and calibration for axial cameras. In *Asian Conference on Computer Vision*, pages 704–713. Springer, 2006.
[12] Ryusuke Sagawa, Masaya Takatsuji, Tomio Echigo, and Yasushi Yagi. Calibration of lens distortion by structured-light scanning. In *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 832–837. IEEE, 2005.
[13] Joaquim Salvi, Jordi Pages, and Joan Batlle. Pattern codification strategies in structured light systems. *Pattern recognition*, 37(4):827–849, 2004.
[14] Miriam Schönbein, Tobias Strauß, and Andreas Geiger. Calibrating and centering quasi-central catadioptric cameras. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4443–4450. IEEE, 2014.
[15] Peter Sturm. Multi-view geometry for general camera models. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, pages 206–212. IEEE, 2005.
[16] Peter Sturm. Calibration of a non-single viewpoint system. *Computer Vision: A Reference Guide*, pages 66–69, 2014.
[17] Peter Sturm and Srikumar Ramalingam. A generic concept for camera calibration. In *European Conference on Computer Vision*, pages 1–13. Springer, 2004.
[18] Peter Sturm, Srikumar Ramalingam, Jean-Philippe Tardif, Simone Gasparini, Joao Barreto, et al. Camera models and fundamental concepts used in geometric computer vision. *Foundations and Trends in Computer Graphics and Vision*, 6(1–2):1–183, 2011.
[19] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence*, 22, 2000.