Multi-View 3D Reconstruction from Uncalibrated Radially-Symmetric Cameras

Jae-Hak Kim [1*]   Yuchao Dai [1*]   Hongdong Li [1,2]   Xin Du [3]   Jonghyuk Kim [1]

[1] Australian National University and [2] NICTA, Australia

[3] Zhejiang University, China

Abstract

We present a new multi-view 3D Euclidean reconstruction method for arbitrary uncalibrated radially-symmetric cameras, which needs no calibration or any camera model parameters other than radial symmetry. It is built on the radial 1D camera model [25], a unified mathematical abstraction to different types of radially-symmetric cameras. We formulate the problem of multi-view reconstruction for radial 1D cameras as a matrix rank minimization problem. Efficient implementation based on alternating direction continuation is proposed to handle scalability issue for real-world applications. Our method applies to a wide range of omnidirectional cameras including both dioptric and catadioptric (central and non-central) cameras. Additionally, our method deals with complete and incomplete measurements under a unified framework elegantly. Experiments on both synthetic and real images from various types of cameras validate the superior performance of our new method, in terms of numerical accuracy and robustness.

1. Introduction

Having a wide field of view, omnidirectional cameras can be used to reconstruct broad scenes from few views, thus have been widely deployed to applications such as surveillance, robot navigation and 3D modeling of street scene. A large body of research has been devoted to the 3D reconstruction problem. However, existing methods are still not fully satisfactory and not flexible enough as most of the existing reconstruction methods rely on specific types of cameras. It is highly desired to have a unified and efficient reconstruction method for omnidirectional cameras.

This paper proposes a new multi-view 3D Euclidean reconstruction method for *generic types* of uncalibrated *radially-symmetric cameras*. It is built on the radial 1D camera model originally developed by Thirthala and Pollefeys [25]. Radial 1D camera model is shown to be a powerful mathematical abstraction, which makes our method generic enough to be applied to a variety of radially-symmetric cameras, be it a central or non-central, dioptric
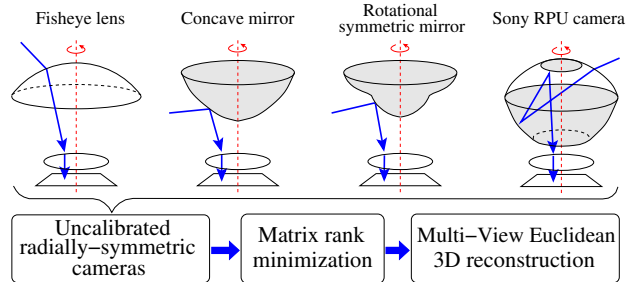

Figure 1. *Our method recovers 3D structure via matrix rank minimization from general types of uncalibrated radially-symmetric cameras – e.g. fisheye lens cameras, concave shape mirror based catadioptric cameras, noncentral cameras including spherical mirror or any radially-symmetric mirror shape based cameras, and multiple relfection surfaces based Sony RPU camera.*

or catadioptric, fisheye, projective or affine. We extend their framework to multi-view case, as opposed to the limit of three or four views at most. The advantage is then, our method is much more numerically stable and efficient, and much less sensitive to noise and perturbations. This is in sharp contrast to what was admitted in [25] that "However, it must be noted that currently the quadrifocal and mixed trifocal tensors are useful only from a theoretical stand-point" and "... hard to develop a robust automatic approach for real images". These drawbacks are tackled by our method.

Our 3D Euclidean reconstruction method adopts a stratification scheme. First, we formulate the problem of multi-view reconstruction for radial 1D cameras as a matrix rank minimization problem, and solve it through convex optimization (and semi-definite programming in particular followed by an efficient alternating direction continuation method). Second, the multi-view reconstruction is upgraded from projective to Euclidean by exploiting the internal constraints. Our method can handle both complete and incomplete measurements cases in a unified way elegantly.

1.1. Modeling radially-symmetric cameras

Due to the various types of omnidirectional camera design and construction, e.g. dioptric (lens-based) or catadioptric (mirror-lens system), central or non-central, most 3D reconstruction methods are specially designed for one

or a few particular types, and thus are not universally applicable to others. Meanwhile, novel types of omnidirectional optical devices are emerging too (*e.g.* Sony's panoramic lens module, which consists of complex refraction and reflection in the light ray path, as illustrated in Fig. 1), which also calls for a unified 3D reconstruction procedure.

There indeed exists a unified mathematical model to represent various types of omnidirectional cameras, so-called generalized camera model (GCM) [19, 22, 11], which models cameras as unconstrained sets of projection rays. However, the GCM does not suggest a unified way to handle 3D reconstruction from uncalibrated cameras. Although it represents incoming rays in a very generic way, the model is not very stable [13]. The radial 1D camera model studied in this paper is applicable to both central and non-central cases, and it is in fact a special case (symmetry version) of the GCM (*i.e.* "Axial camera" as defined in [21, 13, 11]).

In practice, very often omnidirectional cameras will manifest a certain type of symmetry, where radial symmetry being the dominant form. This is reasonable, because it is convenient to design, to manufacture and to use, an omnidirectional camera with a radially-symmetric field of view. There are two major classes of omnidirectional cameras: dioptric and catadioptric. The former one includes a wide-angle lens (fish-eye lens), and the latter one often consists of a perspective camera plus a curved mirror. In either of these two classes, it is often desirable and convenient to have a radially-symmetric field of view. Most of the commonly used omnidirectional cameras belong to this class, and this is the main focus of this paper.

To express radially-symmetric cameras in a unified way, Thirthala and Pollefeys [25] proposed the novel concept of "radial 1D camera" that maps a 3D point to a radial line. Tardif *et al*. [23] developed the varying focal length model, where each distortion circle and the associated viewing cone are considered as an individual perspective camera to provide great benefits to calibrate the camera with a collinearity condition and planar pattern, thus fitting the radial calibration purpose well. Ramalingam *et al*. [21] proposed the "axial camera", where all the rays intersect at a common line. The axial camera is an abstraction of stereo systems, non-central catadioptric cameras and pushbroom cameras.

1.2. Related works

To recover structure and motion for cameras with a wide circular field of view, Mičušík and Pajdla [17] estimated epipolar geometry of radially-symmetric cameras by solving a polynomial eigenvalue problem. Automatic estimation of two-view geometry, and 3D metric reconstruction from point correspondences are achieved in further. Lhuillier [12] presented fully automatic methods for estimating scene structure and camera motion from an image sequence acquired by a catadioptric system, where bundle adjustment

is applied to both central and non-central models.

When a specific camera model is available, 3D reconstruction can be achieved in a tailor-made style. For catadioptric cameras, with the information of the mirror (model and parameters), calibration and reconstruction can be done through computing the forward and backward projection. Geyer and Daniilidis [8] introduced the circle space representation for an image of points and lines in central catadioptric cameras, from which the epipolar constraint and catadioptric fundamental matrix are derived. However their method only works for two views or three views. Mičušík and Pajdla [16] developed accurate non-central and suitable approximate central models for specific mirrors, thus allowing to build a 3D metric reconstruction from two uncalibrated non-central catadioptric images. Agrawal *et al*. [1] computed the optical path from a given 3D point to the given viewpoint with an analytical solution. The analytical forward projection leads to 3D reconstruction via bundle adjustment. They further extended the method to general non-central off-axis camera placement [2].

On calibrating general radially-symmetric cameras, Tardif *et al*. [23] used the varying focal length model, presenting a plumbline type and plane-based calibration method. The plane based calibration for radially-symmetric cameras have also been studied in [20] and [24]. Generally, most of these methods investigate the interaction of the assumed model with the multi-view constraints. Hartley and Kang [9] proposed a parameter-free method to simultaneously calibrate the radial distortion function of a camera and other internal calibration parameters by using a planar calibration grid. However, their model is restricted to central cameras and assumes a known calibration grid. When the distortion model is available, radial distortion calibration and multi-view geometry can be solved with algebraic minimization methods such as [7].

2. Radial 1D camera model

The radial 1D camera model [25] (Fig. 2) is a much more general mathematical abstraction, which encompasses most of the fisheye cameras, central and non-central catadioptric cameras, perspective and affine cameras.

Definition: The radial 1D camera expresses the mapping of a 3D point in $\mathcal{P}^3$ onto a radial line in the image plane. The $\mathcal{P}^3 \rightarrow \mathcal{P}^1$ projection can be represented by a 2×4 matrix and has 7 degrees of freedom [25].

Under the radial 1D camera model, a 3D point $\mathbf{X}_j = [x_j, y_j, z_j, 1]^T$ is mapped to a distorted image measurement $\mathbf{x}_{ij}^d = [u_{ij}^d, v_{ij}^d]^T$ by a radial camera $\mathtt{P}_i \in \mathcal{R}^{2 \times 4}$:

$$\mathtt{P}_i \mathbf{X}_j = \phi_{ij} \mathbf{x}_{ij}^d , \qquad (1)$$

where ϕ_{ij} is a scale factor. We assume the centre of distortion is known and has been mapped to the origin. The projection lies on the line with direction vector $(v_{ij}^d, -u_{ij}^d)^T$.

The backward projection of the line is the plane containing the 3D point X_j and the ray passing through center of distortion $\mathbf{c}_i$ and $\mathbf{x}_{ij}^d$. Obviously, there is a scale ambiguity
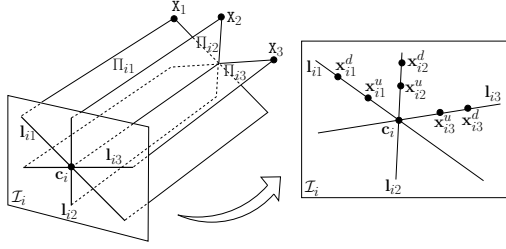

Figure 2. *Radial 1D camera model.*

for ϕ_{ij}. For the projection matrix P and 3D scene structure X, similar scale ambiguity exists, *i.e.*, a scaling projection matrix for each radial camera and scaling scene point individually will not change the 2D image measurements. Nevertheless, we can achieve 3D Euclidean reconstruction without ambiguity as shown in the following sections.

Discussion. The radial 1D camera can be thought of as projecting a bundle of planes containing the optical axis onto a bundle of radial lines passing through the radial centre in the image plane (Fig. 2). The radial 1D camera model encompasses most of the central and non-central omnidirectional cameras. This is because the only essential requirement in this model is that all points lie in one plane, of the bundle around the optical axis, project onto the same radial line (passing through the radial centre).

3. Multi-view reconstruction upto projectivity

In this paper, we target at multi-view Euclidean reconstruction from arbitrary radially-symmetric cameras. To achieve this, we adopt a stratification scheme. First, we achieve multi-view reconstruction upto projectivity through factorization, which offers great simplicity and elegancy. Second, by exploiting intrinsic constraints, the projective reconstruction is upgraded to Euclidean reconstruction. In this way, we do not need any specific camera and distortion model other than the radially-symmetric condition.

By collecting all the image measurements across different frames, we obtain a measurement matrix $M = [\mathbf{x}_{ij}^d]$ of size $2m \times n$ for m frames and n 3D points. Now the relationship Eq.-(1) can be compactly expressed in a matrix form as:

$$PX = [\phi_{ij} \otimes 1_{2\times 1}] \odot M, \qquad (2)$$

where $\otimes$ denotes the Kronecker product, $\odot$ denotes the Hadamard product, $P = [P_1^T \ P_2^T \cdots P_m^T]^T \in \mathcal{R}^{2m \times 4}$ and $X = [X_1 \ X_2 \cdots X_n] \in \mathcal{R}^{4 \times n}$.

We define $W = PX$ as the **weighted measurement matrix**, and equivalently re-write Eq.-(2) as:

$$W = \Phi \odot M = PX = \begin{bmatrix} P_1 \\ \vdots \\ P_m \end{bmatrix} \begin{bmatrix} X_1 & \cdots & X_n \end{bmatrix}, \quad (3)$$

where $\Phi = [\phi_{ij} \otimes 1_{2\times 1}] \in \mathcal{R}^{2m \times n}$. It is easy to check that both P and X have a rank at most 4, therefore the weighted measurement matrix W must have a rank at most 4. Thus, we have reached a factorization formulation similar to the factorization model for perspective cameras. Actually, a perspective camera model with or without distortion falls exactly into the radial 1D camera model. Note that Eq.-(3) can handle non-central cameras as well since it is a projection model by radial 1D cameras in [25].

3.1. Hadamard factorization based solution

Recall that the multi-view factorization model for radially symmetric cameras is expressed as: $PX = \Phi \odot M$. To deal with the scale ambiguity associated with the projection matrix and 3D points, we enforce the column-wise and row-wise normalization on Φ as $\Phi^T 1_m = m1_n$ and $\Phi 1_n = n1_m$, where 1_m and 1_n are vectors with 1 element-wise of length m and n, respectively. For general wide view angle omnidirectional cameras, the coefficient ϕ_{ij} is positive. Taking all the constraints into consideration, mathematically multi-view factorization for radially symmetric cameras is formulated as:

Problem 3.1. Find W and Φ subject to

$$W = \Phi \odot M,$$
$$\text{rank}(W) \leq 4,$$
$$\Phi^T 1_m = m1_n, \quad \Phi 1_n = n1_m,$$
$$\phi_{ij} > 0.$$

Once the scaling matrix Φ is recovered, the multi-view factorization problem can be solved via the singular value decomposition (SVD) as $W = PX$. Note that the solution is defined only up to a nonsingular 4×4 projective transformation H. Euclidean upgrading is explained in Section 4.

Minimization objective: Under the affine camera model, multi-view factorization achieves the maximum likelihood estimation (MLE) [10]. For the perspective camera model, multi-view factorization is actually minimizing an algebraic error, which can be viewed as an approximation to the geometric reprojection error. Under the radial 1D camera model, we cannot measure the geometric reprojection error but only an angular error. We define the angular error corresponding to a distorted image measurement $\mathbf{x}_{ij}^d$ as e_{ij}, which measures the angle between the measured ray and the reconstructed ray, *i.e.* $\angle(\mathbf{x}_{ij}^d, P_i X_j)$. Therefore, to evaluate the whole reconstruction, the performance measure is defined as the total angular error: $\sum_{i=1}^m \sum_{j=1}^n \angle(\mathbf{x}_{ij}^d, P_i X_j) = \sum_{i=1}^m \sum_{j=1}^n \arccos\left(\frac{<\mathbf{x}_{ij}^d, P_i X_j>}{\|\mathbf{x}_{ij}^d\|\|P_i X_j\|}\right)$. In Fig. 3, we illustrate the measure of the angular error. Similar measures have also been used in [12]. Note that the angular error is invariant to the scaling ambiguity associated with P and X.
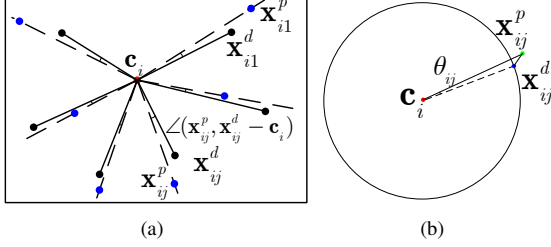
Figure 3. ***Angular error and its approximation with algebraic error.*** *(a) Angular error. (b) Approximating angular error with algebraic error.*

Normalization: Obviously, there is a scaling ambiguity in Φ. If one family of ϕ_{ij} yields a weighted measurement matrix of rank 4, another family ϕ'_{ij} given by $\phi'_{ij} = \delta_i \eta_j \phi_{ij}$ also yields a rank-4 weighted measurement matrix. Therefore normalization is needed to deal with these scale ambiguities. Column-wise and row-wise normalization of ϕ_{ij} deal with scale ambiguity with P and X respectively, which has been used in projective factorization problems *e.g.* [5].

In solving the multi-view factorization problem, we are actually relaxing the objective function to an algebraic error to approximate the geometric angular error. To achieve a good approximation, we propose to normalize all the image measurements $\mathbf{x}^d_{ij}$ to unit norm $\mathbf{x}^d_{ij} \leftarrow \mathbf{x}^d_{ij}/\|\mathbf{x}^d_{ij}\|$ as opposed to the normalization described in [10], Then, the factorization formulation expresses as: $P_i X_j = \phi_{ij} \mathbf{x}^d_{ij}$, where $\mathbf{x}^d_{ij}$ is a unit norm vector (direction). The underlying benefit is that normalization constraints $\sum_{i=1}^{m} \phi_{ij} = m$ and $\sum_{j=1}^{n} \phi_{ij} = n$ make all the scaling parameter ϕ_{ij} comparable, which gives a good approximation to the angular error as shown in Fig. 3(b), where $\mathbf{x}^p_{ij} = P_i X_j$.

4. Multi-view Euclidean Reconstruction

In this section, we upgrade the multi-view reconstruction from radially-symmetric camera to Euclidean reconstruction by exploiting constraints on the intrinsic camera matrix. Our basic assumptions are: 1) Radially-symmetric camera; 2) Identical aspect ratio, *i.e.* square pixel; 3) Center of distortion is known and aligned with the principal point; and 4) Zero skew.

Once we have recovered the scaling matrix Φ and the weighted measurement matrix W, the projection matrix P and scene structure X can be recovered through SVD as $W = PX$. However this decomposition is not unique as it is defined up to a nonsingular linear transformation $H \in \mathcal{R}^{4 \times 4}$, *i.e.* $P = \hat{P}H$ and $X = H^{-1}\hat{X}$ achieve identical image measurements W. If a reasonable upgrading matrix H is achieved, the Euclidean structure and motion can be recovered from structure matrix $\hat{X}$ and projection matrix $\hat{P}$.

First let us decompose the upgrading matrix H into two parts as $H = [H_l | H_r]$, where H_l denotes the first three

columns of H while H_r for the fourth column. The absolute dual quadratic $Q = H_l H_l^\top$ projects to the dual image of the absolute conic $\omega_i^* = K_i K_i^\top$ as $\omega_i^* = P_i Q P_i^\top$ [10]. The objective here is to estimate Q directly from constraints on the intrinsic parameters. We use the intrinsic camera matrix as $K_i = \text{diag}(f_i, f_i)$ and $\omega_i^* = \text{diag}(f_i^2, f_i^2)$.

Based on the special structure of ω_i^*, we obtain the following two linear constraints on Q:
$$\begin{cases} \omega_i^*(1,2) = \omega_i^*(2,1) = 0, \\ \omega_i^*(1,1) - \omega_i^*(2,2) = 0. \end{cases}$$

Thus we can solve Q linearly. The Euclidean upgrading matrix H can be obtained through SVD of Q.

5. Efficient Implementations

In this section, we propose to solve multi-view projective reconstruction Problem 3.1 efficiently via semi-definite programming (SDP) and the augmented Lagrangian multiplier method (ALM).

5.1. Semi-definite Programming

Under noiseless and complete measurement case, our multi-view factorization Problem 3.1 with the rank constraint $\text{rank}(W) \leq 4$ can be equivalently written as a matrix rank minimization, $\min \text{rank}(W)$ (rank-4 is the upper bound for general configured radially-symmetric cameras).

Since the rank minimization problem is NP-hard in general, we propose to use the relaxed form as the objective function, *i.e.*, $\min \|W\|_*$. The nuclear norm $\|W\|_*$ of a matrix $W \in \mathcal{R}^{m \times n}$ is defined as the sum of singular values of W. Recently, nuclear norm minimization has been widely used in low-rank modeling such as projective factorization [5, 3] and robust principal component analysis [4]. The nuclear norm serves as a tight convex surrogate to the rank function, from which we develop efficient implementation.

We finally obtain a **trace norm minimization** formulation for Hadamard factorization :

$$
\begin{aligned}
&\min_{W,X,Y,\Phi} \frac{1}{2}(\text{tr}(X) + \text{tr}(Y)) \\
&\text{s.t.} \ \begin{pmatrix} X & W \\ W^T & Y \end{pmatrix} \succeq 0, \\
&W = \Phi \odot M, \\
&\Phi^T \mathbf{1}_m = m\mathbf{1}_n, \ \ \Phi \mathbf{1}_n = n\mathbf{1}_m, \\
&\phi_{ij} > 0.
\end{aligned}
\tag{4}
$$

where auxiliary variables X and Y are introduced to equivalently represent the nuclear norm as trace norm under the positive semi-definite constraint [6].

This is a standard SDP optimization problem, thus can be solved using any off-the-shelf SDP solvers such as SDPT3 [26]. However, due to excessive memory and computational requirement, these state-of-the-art SDP solvers still cannot handle large scale real-world problems.

5.2. Augmented Lagrangian Multiplier

To further speed up the implementation, we resort to the Augmented Lagrangian Multiplier (ALM) method which has been widely used in the low-rank representation [14, 15, 5]. The nuclear norm of $\mathtt{W}$ can be expressed in terms of its factors $\mathtt{U}$ and $\mathtt{V}$ as, $\|\mathtt{W}\|_* = \min_{\mathtt{U},\mathtt{V}:\mathtt{W}=\mathtt{UV}^T} \frac{1}{2}(\|\mathtt{U}\|_F^2 + \|\mathtt{V}\|_F^2)$. If $\mathrm{rank}(\mathtt{W}) = k \le \min(m, n)$, then the minimum above is attained at a factor decomposition $\mathtt{W} = \mathtt{U}_{m\times k}\mathtt{V}_{n\times k}^T$. Thus, we slightly modify our formulation Eq.-(4) following the low rank parametrization and reach

$$\min \frac{1}{2}\|\mathtt{UV}^T - \Phi \odot \mathtt{M}\|_F^2 + \frac{\mu}{2}(\|\mathtt{U}\|_F^2 + \|\mathtt{V}\|_F^2)$$
$$\text{s.t. } \Phi^T \mathbf{1}_m = m\mathbf{1}_n, \ \Phi \mathbf{1}_n = n\mathbf{1}_m, \ \phi_{ij} > 0, \quad (5)$$

where $\mathtt{U} \in \mathcal{R}^{2m\times 4}$ and $\mathtt{V} \in \mathcal{R}^{n\times 4}$ give the explicit low rank representation. In the formulation, we apply a continuation technique (*i.e.*, homotopy approach) to accelerate the convergence of the algorithm. The parameter η_μ determines the rate of reduction of consecutive μ_l, as $\mu_{l+1} = \max\{\mu_l\eta_\mu, \overline{\mu}\}, l = 1, \cdots, L-1, 0 < \eta_\mu < 1$.

To efficiently solve Eq.-(5), we adopt the classic alternating direction method by introducing Lagrangian multipliers Υ and Γ. The resulting partial augmented Lagrangian formulation states as:

$$\mathcal{L}(\Phi, \mathtt{U}, \mathtt{V}, \Gamma, \Upsilon) = \frac{1}{2}\|\mathtt{UV}^T - \Phi \odot \mathtt{M}\|_F^2 + \frac{\mu}{2}(\|\mathtt{U}\|_F^2 + \|\mathtt{V}\|_F^2)$$
$$+ \langle \Gamma, \Phi^T \mathbf{1}_m - m\mathbf{1}_n \rangle + \langle \Upsilon, \Phi \mathbf{1}_n - n\mathbf{1}_m \rangle$$
$$+ \frac{\beta}{2}(\|\Phi^T \mathbf{1}_m - m\mathbf{1}_n\|_F^2 + \|\Phi \mathbf{1}_n - n\mathbf{1}_m\|_F^2), \quad (6)$$

where $\beta > 0$ is a penalty parameter. Eq.-(6) is not jointly convex over $\mathtt{U}$, $\mathtt{V}$ and Φ, therefore we propose to minimize it with respect to $\mathtt{U}$, $\mathtt{V}$ and Φ one at a time while fixing the others. The iterations of ALM go as: $\mathtt{U}_+ \leftarrow \arg\min_{\mathtt{U}} \frac{1}{2}\|\mathtt{UV}^T - \Phi \odot \mathtt{M}\|_F^2 + \frac{\mu}{2}\|\mathtt{U}\|_F^2$, and $\mathtt{V}_+ \leftarrow \arg\min_{\mathtt{V}} \frac{1}{2}\|\mathtt{UV}^T - \Phi \odot \mathtt{M}\|_F^2 + \frac{\mu}{2}\|\mathtt{V}\|_F^2$. The update of $\mathtt{U}$ and $\mathtt{V}$ both have closed-form solutions as, $\mathtt{U}_+ = (\Phi \odot \mathtt{M})\mathtt{V}(\mathtt{V}^T\mathtt{V} + \mu\mathtt{I})^{-1}$ and $\mathtt{V}_+ = (\Phi \odot \mathtt{M})^T\mathtt{U}(\mathtt{U}^T\mathtt{U} + \mu\mathtt{I})^{-1}$.

To update ϕ_{ij}, we can also adapt a similar strategy and obtain a closed-form solution. However, the equation system is of size $2mn \times 2mn$. To further speed up the implementation, we propose to update ϕ_{ij} elementwise from the updated weighted measurement matrix $\mathtt{W}_+ = \mathtt{U}_+\mathtt{V}_+^T$ and normalized measurements $\mathtt{M}$ as $\phi_{ij}^{(k+1)} = (\mathbf{w}_{ij}^T\mathbf{m}_{ij})/(\mathbf{m}_{ij}^T\mathbf{m}_{ij})$. Then we normalize Φ to satisfy the column-sum and row-sum constraints. In this way, we even do not need to update the Lagrangian multipliers Γ, Υ and β. Therefore, we obtain an alternating direction continuation based algorithm for Hadamard factorization formulation for radially symmetric camera reconstruction, which alternates between updating $\mathtt{U}$, $\mathtt{V}$ and Φ. Refer to the supplementary material for implementation details.

5.3. Dealing with incomplete measurements

With a wide field of view camera, we can reconstruct 3D scene from a few images, which will create an incomplete measurement matrix in general and the SVD based method cannot be applied.

In this subsection, we extend our Hadamard factorization formulation from complete measurements case to incomplete measurements case. With missing data setting, we are given an incomplete measurement matrix $\mathtt{M} = [\mathbf{x}_{ij}^d]$, where the missing elements are completed with 0. To index the missing pattern and elements, we define a $0 - 1$ *mask matrix* Ω as

$$\Omega = [\omega_{ij}], \ \omega_{ij} = \begin{cases} 1 \in \mathcal{R}^2, & \text{if } \mathbf{x}_{ij}^d \text{ is available,} \\ 0 \in \mathcal{R}^2, & \text{if } \mathbf{x}_{ij}^d \text{ is missing.} \end{cases}$$

With these notations, the imaging process for radially-symmetric cameras with missing data can be compactly expressed as: $\Phi \odot \mathtt{M} = \Omega \odot \mathtt{W}$. Compared with the complete measurements case, we have to recover a completed low-rank matrix $\mathtt{W}$ and scaling matrix Φ such that $\Phi \odot \mathtt{M} = \Omega \odot \mathtt{W}$. For those missing positions, we do not need to estimate the corresponding depths, so we set $\phi_{ij} = 1$ whenever $\omega_{ij} = 0$. This setting guarantees that the column-sum and row-sum constraints on the available measurements and all the projective depths possess appropriate scales.

Semi-definite programming formulation can be derived by slightly modifying Eq.-(4). Once this SDP converges, the resultant $\mathtt{W}$ is a completed $2m \times n$ full matrix with no entries missing. Additionally, our alternating direction continuation based efficient implementation can also be extended to incomplete measurement case directly, achieving missing points handling ALM method (MALM). Implementation details are available in the supplementary material.

6. Experiments

6.1. Synthetic data

Catadioptric system. We performed experiments on both synthetic and real data. To generate the synthetic data we created 100 3D points placed randomly on 3 walls intersecting each other in 90 degrees, and a catadioptric system consisting of a perspective camera and a spherical mirror axially aligned is placed to capture the image of the 3D points on the walls. The catadioptric system is moved to capture 20 images at different views. Each image is obtained by the perspective projection of reflection on the spherical mirror. Note that this reflection by the spherical mirror creates axially aligned radial distortion in the image.

Methods SIESTA (the Simplest Iterative Extension of the Strum/Triggs algorithm [27, 18]), SDP (Semi-Definite Programming method), ALM (Augmented Lagrangian Multiplier method) and MALM (Missing points handling ALM method) are applied on the synthetic data and the
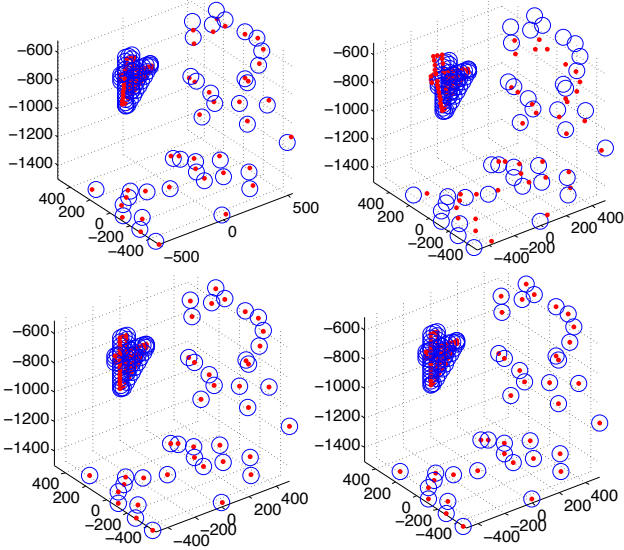
Figure 4. *Euclidean reconstruction results of the synthetic data.* *3D reconstruction (red dot) after registration with the ground truth (blue circle) by SIESTA (Top-left), SDP (Top-right), ALM (Bottom-left) and MALM (Bottom-right).*

	Synthetic data (20 frames, 100 points)			
	SIESTA	SDP	ALM	MALM
2D angle error	0.0138°	0.0552°	0.016°	0.8106°
3D angle error	1.6118°	0.4967°	0.0991°	0.0571°
3D reg. error	1.60%	4.46%	0.13%	1.05e-10%
Aspect ratio err.	0.0599%	0.2827%	0.0073%	0.0571%
Skew error	1.21E-03	8.80E-04	4.60E-05	9.40E-04
Elapsed time	88 sec	776 sec	145 sec	39 sec

Table 1. *Analysis of the synthetic data experiments (without noise). The 2D angle error is the angle between the estimated radial line and the input radial line in the image. The 3D angle error is the angle difference between two estimated walls and the ground truth $90°$. The 3D registration error is a relative error defined as $||X - \bar{X}||/||\bar{X}||$, where X is a matrix of 3D points estimate and $\bar{X}$ is a matrix of the ground truth 3D points. The skew error is obtained from the estimated camera calibration matrix.*

results of Euclidean reconstruction are shown in Fig. 4. It shows that the ALM method outperforms against the SIESTA and SDP method in all error metrics. For each method, the angle difference between reconstructed 3D points on the walls and the ground truth, the 2D angle error from estimated radial lines, the 3D registration error, calibration error (aspect ratio and skew parameters) and the computation time are reported (refer to the caption of Table 1 for error metrics). The experiments were performed on a single thread of Intel Core 2.66GHz CPU machine. The experimental results are summarized in Table 1.

For statistical tests, experiments on the data by varying the level of Gaussian noise σ on image point coordinates are performed for all methods. Since the image point coordinates on the dataset are distributed between -1.3 and 1.9, the standard deviation of image noise level $\sigma \in [0, 0.001]$
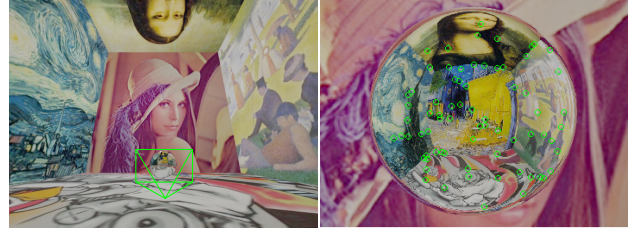

Figure 6. *Catadioptric simulation and rendered images. (Left) A catadioptric system consists of a spherical mirror and a camera (green frustrum) in 6 textured walls. (Right) A sample of the rendered image sequence and feature matches (green circles).*

	MirrorBall				RPU lens		
	f=41, p=72			p=107	f=17, p=51		
	SIESTA	SDP	ALM	MALM	SIESTA	SDP	ALM
2D angle error	0.771°	0.073°	0.001°	0.052°	0.191°	0.026°	0.026°
Aspect ratio err.	11.32%	0.14%	0.07%	0.06%	0.369%	0.002%	0.003%
Skew error	0.42	3.10E-05	3.40E-05	9.30E-05	0.0053	0.0019	0.0019
Elapsed time	380 sec	1,004 sec	185 sec	339 sec	89 sec	49 sec	8 sec

Table 2. *Results of the synthetic 'MirrorBall' and real 'rooftop' data from Sony RPU omnidirectional camera module. The numbers in headings indicate the number of frames (f) and points (p).*
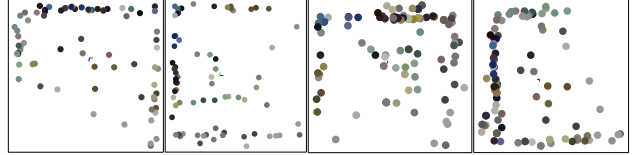

Figure 7. *3D reconstruction of 'MirrorBall' by ALM (41 frames and 72 points) and MALM (41 frames and 107 points with 14% missing data). (First and second) Top and side view of the Euclidean reconstruction by the ALM method from the image sequence (Fig. 6-right). (Third and fourth) The same views of 3D reconstruction by MALM and missing points are recovered.*

is chosen (Note that, for an image of size 1280×1280, it corresponds to the maximum noise 1.2 pixels for 99.7 percent of the sample). The result is compared with the tensor based method [25], and as shown in Fig. 5, our proposed methods, in particular, the ALM method significantly outperforms other methods including the tensor based one.

Catadioptric rendered images. We generated a dataset 'MirrorBall' which is an image sequence rendered by a computer graphics application to simulate a catadioptric system in the configuration shown in Fig. 6. A spherical mirror is placed in front of a perspective camera inside 6 textured walls, and the reflection on the spherical mirror is rendered. The rendered images are used in feature tracking and the matches are collected as input for our methods. As shown in Table 2 and Fig. 7, the ALM method is the best against the SIESTA and SDP method. An experiment using the MALM method is performed on the 'MirrorBall' dataset with 14% missing points. For the 41 frames of the dataset, about 1.5 times of number of points is increased and correctly estimated by the MALM method.
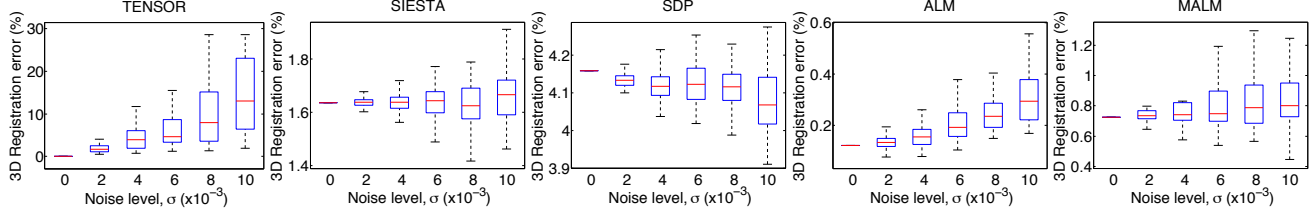
Figure 5. ***3D registration errors (compared with the ground truth) under varying noise level.*** *(From left to right) The tensor based method [25], SIESTA, SDP, ALM and MALM method. All experiments were performed 100 times per each different level of noise.*
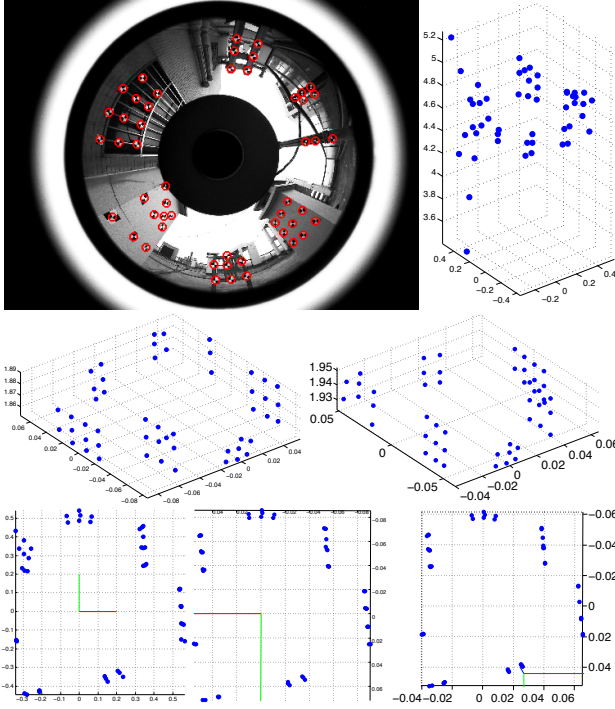

Figure 8. ***Euclidean reconstruction of the 'rooftop' sequence.*** *(Top-left) The 'rooftop' image with features from Sony RPU lens. (Top-right) Side view of the reconstruction by SIESTA. (Middle) Side view of the reconstruction by SDP and ALM method (from left to right). (Bottom) Top view of the reconstruction by SIESTA, SDP and ALM method (from left to right).*

6.2. Real data

Sony RPU camera module. Sony RPU camera module consisting of multiple reflection surfaces and a camera to create a panoramic image was used to capture a 'rooftop' sequence of 17 images. Features in the images were manually selected as shown in Fig. 8-(Top-left). Note that the 3D points are reconstructed correctly in parallel as shown in Fig. 8. The quantitative result is summarized in Table 2.

Omnidirectional camera (catadioptric). An omnidirectional image sequence captured by a catadioptric camera (Kumotek VS-C14U-80-ST) was used in our experiment. We captured images of a paper that was rolled to be a cylin-
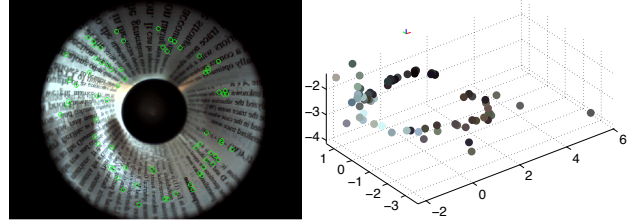

Figure 9. ***Omnidirectional image and 3D reconstruction.*** *(Left) An image of a cylinder shape of a paper captured by the omnidirectional camera (catadioptric). (Right) 3D reconstruction by the ALM method. It reports the 2D angle error* $0.038°$*, aspect ratio error 0.047% and skew error 1.7e-03 (in 280 sec).*

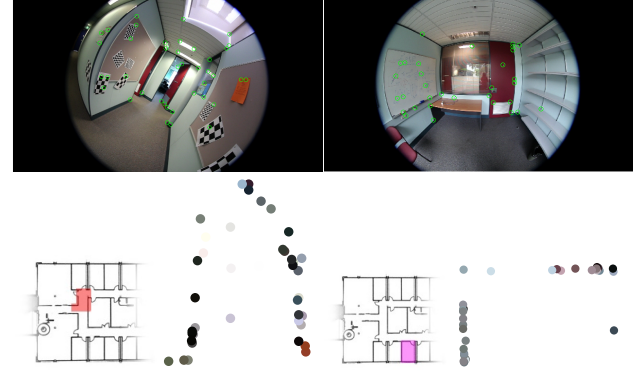Figure 10. ***Euclidean reconstruction results of fisheye lens.*** *(First column) The dataset 'CheckCorrNorth' by Canon fisheye 8mm lens. A top view of the reconstruction by SIESTA (171 frames, 40 points) and the area of the reconstruction in the building map. (Second column) The dataset 'OfficeScene' results by SIESTA. (492 frames, 32 points). Colour of dots indicates the pixel colour value of the feature from the input image sequence.*

der shape, then the catadioptric camera was inserted inside the rolled paper. The image and the 3D reconstruction are shown in Fig. 9. The shape of the 3D reconstruction by the ALM method shows a circular shape correctly.

Canon fisheye lens. Image sequences were captured by Canon fisheye lens at indoor scenes such as corridors and a room as shown in Fig. 10. The results are summarized in Table 3. For 3D reconstruction tests, we measured the distance of two ends from the scene (*e.g.* the height and the width of the corridor, the depth of the room, etc). Since we have a Euclidean reconstruction up to scale, the ratio of two dis-

	CheckCorrNorth (f=171, p=40)		OfficeScene (f=492, p=32)	
	SIESTA	ALM	SIESTA	ALM
2D angle error	0.121°	0.121°	0.062°	0.058°
3D ratio error	17.46%	4.55%	7.25%	3.32%
Aspect ratio err.	0.10%	0.06%	0.02%	0.09%
Skew error	0.021	0.005	6.10E-04	2.90E-04
Elapsed time	603 sec	154 sec	2,368 sec	433 sec

Table 3. ***Results of real sequences from fisheye lens.*** *Two real sequence results corresponding to Fig. 10. The numbers in headings indicate the number of frames (f) and the number of points (p) in each dataset.*
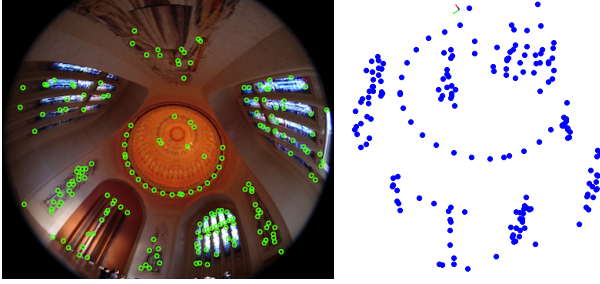

Figure 11. ***3D reconstruction of a dome structure (the Hall of Memory, Australian War Memorial).*** *(Left) One of 79 images and 189 features (green). (Right) 3D points recovered by ALM.*

tances is compared with the ground truth. The reports show that our proposed ALM achieves less than 5% error of 3D reconstruction compared with the ground truth. In Fig. 11, we show 3D reconstruction by our ALM method from 189 points and 79 images of the interior of a dome structure, the Hall of Memory at the Australian War Memorial.

7. Conclusion and discussion

In this paper, we propose a multi-view 3D Euclidean reconstruction method for radially-symmetric cameras. Since our method takes advantage of multiple views, significant improvement is achieved especially under noisy measurement case. With multi-view input, our method overcomes the previous theoretical boundary of 3D reconstruction from radial 1D cameras. Our method can handle different types of cameras (*e.g.* one view from a fish eye lens and another view from catadioptric camera), theoretically this is identical and solvable. A current limitation of our method is that the centre of distortion needs to be known or estimated from lens outer edges. As a future work, incorporating the estimation of the center of distortion in our framework may be studied to achieve more accurate estimation.

Acknowledgements

References

[1] A. Agrawal, Y. Taguchi, and S. Ramalingam. Analytical forward projection for axial non-central dioptric and catadioptric cameras. In *ECCV*, pages 129–143. 2010.

[2] A. Agrawal, Y. Taguchi, and S. Ramalingam. Beyond Alhazen's problem: Analytical projection model for non-central catadioptric cameras with quadric mirrors. In *CVPR*, pages 2993–3000, 2011.

[3] R. Angst, C. Zach, and M. Pollefeys. The generalized trace-norm and its application to structure-from-motion problems. In *ICCV*, pages 2502–2509, 2011.

[4] E. Candès, X. Li, Y. Ma, and J. Wright. Robust principal component analysis? *J. ACM*, 58(3):11:1–37, 2011.

[5] Y. Dai, H. Li, and M. He. Projective multiview structure and motion from element-wise factorization. *PAMI*, 35(9):2238–2251, 2013.

[6] M. Fazel, H. Hindi, and S. Boyd. A rank minimization heuristic with application to minimum order system approximation. In *Proceedings of the American Control Conference*, pages 4734–4739, 2001.

[7] A. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In *CVPR*, pages 125–132, 2001.

[8] C. Geyer and K. Daniilidis. Structure and motion from uncalibrated catadioptric views. In *CVPR*, pages 279–286, 2001.

[9] R. Hartley and S. B. Kang. Parameter-free radial distortion correction with center of distortion estimation. *PAMI*, 29(8):1309–1321, 2007.

[10] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, Second edition, 2004.

[11] J.-H. Kim, H. Li, and R. Hartley. Motion estimation for nonoverlapping multicamera rigs: Linear algebraic and L-infinity geometric solutions. *PAMI*, 32(6):1044–1059, 2010.

[12] M. Lhuillier. Automatic scene structure and camera motion using a catadioptric system. *CVIU*, 109(2):186 – 203, 2008.

[13] H. Li, R. Hartley, and J.-H. Kim. A linear approach to motion estimation using generalized camera models. In *CVPR*, pages 1–8, 2008.

[14] Z. Lin, M. Chen, and Y. Ma. The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-Rank Matrices. *ArXiv e-prints*, Sept. 2010.

[15] R. Liu, Z. Lin, F. De la Torre, and Z. Su. Fixed-rank representation for unsupervised visual learning. In *CVPR*, pages 598–605, 2012.

[16] B. Mičušík and T. Pajdla. Autocalibration 3d reconstruction with non-central catadioptric cameras. In *CVPR*, pages 58–65, 2004.

[17] B. Mičušík and T. Pajdla. Structure from motion with wide circular field of view cameras. *PAMI*, 28(7):1135–1149, 2006.

[18] J. Oliensis and R. Hartley. Iterative extensions of the Sturm/Triggs algorithm: Convergence and nonconvergence. *PAMI*, 29(12):2217–2233, 2007.

[19] R. Pless. Using many cameras as one. In *CVPR*, 2003.

[20] S. Ramalingam, P. Sturm, and E. Boyer. A factorization based self-calibration for radially symmetric cameras. In *3DPVT*, pages 480–487, 2006.

[21] S. Ramalingam, P. Sturm, and S. Lodha. Theory and calibration for axial cameras. In *ACCV*, pages 704–713. 2006.

[22] P. Sturm. Multi-view geometry for general camera models. In *CVPR*, pages 206–212, 2005.

[23] J.-P. Tardif, P. Sturm, and S. Roy. Self-calibration of a general radially symmetric distortion model. In *ECCV*, pages 186–199. 2006.

[24] J.-P. Tardif, P. Sturm, and S. Roy. Plane-based self-calibration of radial distortion. In *ICCV*, pages 1–8, 2007.

[25] S. Thirthala and M. Pollefeys. Radial multi-focal tensors. *Int J Comput Vis*, 96(2):195–211, 2012.

[26] K. Toh, M. Todd, and R. Tutuncu. SDPT3 — a Matlab software package for semidefinite programming. *Optimization Methods and Software*, 11:545–581, 1999.

[27] B. Triggs. Factorization methods for projective structure and motion. In *CVPR*, pages 845–851, 1996.