Closed-Form Optimal Two-View Triangulation Based on Angular Errors

Seong Hun Lee   Javier Civera

I3A, University of Zaragoza, Spain

{seonghunlee, jcivera}@unizar.es

Abstract

In this paper, we study closed-form optimal solutions to two-view triangulation with known internal calibration and pose. By formulating the triangulation problem as L_1 and L_∞ minimization of angular reprojection errors, we derive the exact closed-form solutions that guarantee global optimality under respective cost functions. To the best of our knowledge, we are the first to present such solutions. Since the angular error is rotationally invariant, our solutions can be applied for any type of central cameras, be it perspective, fisheye or omnidirectional. Our methods also require significantly less computation than the existing optimal methods. Experimental results on synthetic and real datasets validate our theoretical derivations.

1. Introduction

Recovering the position of a 3D point given its projections in two or more cameras is called triangulation. It constitutes a fundamental building block in stereo vision [33], simultaneous localization and mapping (SLAM) [22] and structure-from-motion (SfM) pipelines [29]. For large problems, reconstructing thousands (or millions) of points is not uncommon, so achieving fast and accurate triangulation is important for the performance of such systems.

If one assumes the exact knowledge of camera matrices and noiseless feature measurements, triangulation amounts to intersecting two backprojected rays that correspond to the same 3D point. In practice, however, this assumption is unrealistic, and the rays do not necessarily intersect. Therefore, a nontrivial method is required even for just two views.

The standard approach is to find the 3D point that minimizes a chosen cost function given the feature measurements. The most common are the L_1 norm (sum of magnitude), L_2 norm (sum of squares) and L_∞ norm (maximum) of image reprojection errors. While reasonable for perspective cameras, the image reprojection error does not generalize well to different camera types (e.g., omnidirectional or fully spherical panoramic cameras). This motivates the use of *angular* reprojection error, a rotationally invariant al-

ternative to the *image* reprojection error that is generic and independent of the projection geometry [21, 24, 27].

In this work, we derive, for the first time to our knowledge, the exact closed-form solutions to the L_1 and L_∞ optimal triangulation from two views based on the angular reprojection error. Unlike iterative methods (e.g., [13, 17]), the proposed methods guarantee global optimality without any iterations, and unlike polynomial methods (e.g., [10, 23, 32]), they do not involve finding the roots of a higher-degree polynomial. Hence, our methods simultaneously provide the global optimality, speed and simplicity. We also present our own derivation of the L_2 optimal solution that is much more compact and geometrically intuitive than the existing one [24]. Since all three methods are based on the angular error, they are not limited to standard perspective cameras and can also be used for fisheye, omnidirectional and fully spherical panoramic cameras.

The paper is organized as follows. In the next three sections, we discuss the related work and preliminaries. Section 5, 6 and 7 respectively present the closed-form solutions to the L_1, L_2 and L_∞ optimal triangulation. To make our paper compact and easily accessible, we separated the proofs from our main findings and put them in the appendix. Section 8 addresses the cheirality constraint. Finally, experimental results are provided in Section 9, followed by the conclusions in Section 10.

2. Related Work

The most widespread approach to triangulation is to find the 3D point that minimizes the L_2 norm of image reprojection errors [8]. Assuming that image points are perturbed by Gaussian noise, the L_2 optimal solution gives the maximum likelihood estimate (MLE). This can be obtained in closed form by solving a polynomial of degree 6 for two views [10] and degree 47 for three views [32]. Such polynomial methods are, however, computationally expensive and susceptible to ill-conditioning [17]. Besides, an iterative search for the roots may converge to a local minimum [10].

Another two-view method by Kanatani et al. [13] iteratively corrects the 2D projections of the points. Although this method was shown to be faster than the one by Hartley

and Sturm [10], it does not satisfy the epipolar constraint [19] in each iteration. Lindstrom [17] solved this problem with an improved iterative algorithm that is even more stable and faster. However, neither his method nor Kanatani's guarantees global optimality. Oliensis [24] showed that by formulating the problem as L_2 minimization of the sine of angular reprojection errors, an exact closed-form solution can be derived for two-view triangulation.

Instead of minimizing the L_2 norm, one may choose to minimize the L_1 norm of reprojection errors. The advantage of L_1 norm is that it is more robust to outliers as it places less emphasis on large errors [10, 11]. For two views, Hartley and Sturm [10] showed that the L_1 optimal solution can be obtained in closed form by solving a polynomial of degree 8. They also found that the L_1 optimization gives slightly more accurate 3D results than the L_2 optimization.

In geometric problems, another popular norm is the L_∞ norm. The L_∞ optimal solution corresponds to the MLE under the assumption of uniform noise in the image points [6]. The advantage of the L_∞ cost function over the L_2 cost is that it is relatively simpler and has a single minimum [9]. For the case of two views, Níster [23] showed that the optimal solution can be obtained in closed form by keeping the reprojection errors equal in the two views and solving the resulting quartic equation. A main drawback of the L_∞ cost is that it is relatively more sensitive to outliers [9]. This being said, such sensitivity was shown to be useful for outlier removal [30, 26, 16].

While most of the aforementioned works formulate their optimization problem in terms of the image reprojection error, the angular reprojection error is another popular choice. It embodies a better noise model for fisheye or omnidirectional cameras [24, 20]. Even for perspective cameras, the assumption of Gaussian noise is not justified [6], and the angular reprojection error is just as valid as the image reprojection error, if not more so. In the literature, it has been proposed to minimize the sine of angular reprojection errors in L_2 norm [24], the tangent in L_2 or L_∞ norm [7, 9, 12], and the cosine in negative L_1 norm [28, 3]. In contrast to these methods, our L_1 and L_∞ optimization do not involve trigonometric functions.

3. Preliminaries on 3D Geometry

Throughout the paper, we adopt the following notation: We use bold letters for vectors and matrices, and light letters for scalars. The Euclidean norm of a vector $\mathbf{v}$ is denoted by $\|\mathbf{v}\|$, and the unit vector by $\widehat{\mathbf{v}} = \mathbf{v}/\|\mathbf{v}\|$. The angle between two lines $\mathbf{L}_0$ and $\mathbf{L}_1$ is denoted by $\angle\,(\mathbf{L}_0,\ \mathbf{L}_1) \in [0, \pi/2]$.

The following vector identities will come in handy later:

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) \qquad (1)$$

$$\|\widehat{\mathbf{a}} \times \widehat{\mathbf{b}}\|^2 = 1 - (\widehat{\mathbf{a}} \cdot \widehat{\mathbf{b}})^2 \qquad (2)$$

We also make frequent use of the following formulas:

1. The distance between a point $\mathbf{p}$ and a plane $\Pi_0(\mathbf{x}) = \mathbf{n}_0 \cdot (\mathbf{x} - \mathbf{c}_0) = 0$ is given by $\|\mathbf{p} - \mathbf{r}_0\|$ where $\mathbf{r}_0$ is the projection of $\mathbf{p}$ onto Π_0. This is computed as follows:

$$\|\mathbf{p} - \mathbf{r}_0\| = |\widehat{\mathbf{n}}_0 \cdot (\mathbf{p} - \mathbf{c}_0)|. \qquad (3)$$

2. The distance between two skew lines $\mathbf{L}_0(s_0) = \mathbf{c}_0 + s_0\mathbf{m}_0$ and $\mathbf{L}_1(s_1) = \mathbf{c}_1 + s_1\mathbf{m}_1$ is given by $\|\mathbf{r}_0 - \mathbf{r}_1\|$ where $\mathbf{r}_0$ and $\mathbf{r}_1$ are the points on each line that form the closest pair. Letting $\mathbf{t} = \mathbf{c}_0 - \mathbf{c}_1$ and $\mathbf{q} = \mathbf{m}_0 \times \mathbf{m}_1$, this is computed as follows:

$$\|\mathbf{r}_0 - \mathbf{r}_1\| = |\mathbf{t} \cdot \widehat{\mathbf{q}}|. \qquad (4)$$

The two points can also be obtained individually [14]:

$$\mathbf{r}_0 = \mathbf{c}_0 + \frac{\mathbf{q} \cdot (\mathbf{m}_1 \times \mathbf{t})}{\|\mathbf{q}\|^2}\mathbf{m}_0, \qquad (5)$$

$$\mathbf{r}_1 = \mathbf{c}_1 + \frac{\mathbf{q} \cdot (\mathbf{m}_0 \times \mathbf{t})}{\|\mathbf{q}\|^2}\mathbf{m}_1. \qquad (6)$$

Equation (4) can be interpreted as the minimum amount of translation required for the two lines to intersect. In this work, it will be also important to know the minimum amount of rotation (or pivot) required for the two lines to intersect. We answer this question in the following lemma:

Lemma 1 (Minimum Pivot Angle for Intersection)
Given two skew lines $\mathbf{L}_0(s_0) = \mathbf{c}_0 + s_0\mathbf{m}_0$ *and* $\mathbf{L}_1(s_1) = \mathbf{c}_1 + s_1\mathbf{m}_1$, *let* $\mathbf{L}_0'$ *be the line that forms the smallest angle* $\theta_0 \in [0,\ \pi/2]$ *to* $\mathbf{L}_0$ *among all possible lines that intersect both point* $\mathbf{c}_0$ *and line* $\mathbf{L}_1$. *Then,* $\mathbf{L}_0'$ *is the projection of* $\mathbf{L}_0$ *onto the plane that contains* $\mathbf{c}_0$ *and* $\mathbf{L}_1$. *Furthermore, letting* $\mathbf{t} = \mathbf{c}_0 - \mathbf{c}_1$ *and* $\mathbf{n}_1 = \mathbf{m}_1 \times \mathbf{t}$,

$$\sin(\theta_0) = |\widehat{\mathbf{n}}_1 \cdot \widehat{\mathbf{m}}_0|. \qquad (7)$$

We call θ_0 **the minimum pivot angle for intersection**, *as it represents the smallest angle required for pivoting line* $\mathbf{L}_0$ *at* $\mathbf{c}_0$ *to make it intersect* $\mathbf{L}_1$.

Proof. Refer to Appendix A. ∎

4. Preliminaries on Two-View Triangulation

Consider two cameras C_0 and C_1 observing the same 3D world point $\mathbf{x}_w$. Let $\mathbf{c}_0$ and $\mathbf{c}_1$ be their positions in the world frame, and let $\mathbf{R}$ and $\mathbf{t}$ be the rotation matrix and translation vector that together transform a point from the camera frame C_0 to C_1, i.e., $\mathbf{x}_1 = \mathbf{R}\mathbf{x}_0 + \mathbf{t}$, where $\mathbf{x}_0 = [x_0, y_0, z_0]^\mathsf{T}$ and $\mathbf{x}_1 = [x_1, y_1, z_1]^\mathsf{T}$ correspond to $\mathbf{x}_w$ in camera frame C_0 and C_1, respectively. Since triangulation is impossible for zero translation, we set $\|\mathbf{t}\| = \|\mathbf{c}_0 - \mathbf{c}_1\| = 1$ without loss of generality. Let $\mathbf{u}_0 = (u_0, v_0, 1)^\mathsf{T}$ and $\mathbf{u}_1 = (u_1, v_1, 1)^\mathsf{T}$ be the homogeneous pixel coordinates of the estimated correspondence to $\mathbf{x}_w$ in each frame. Given the camera calibration matrix $\mathbf{K}$, the normalized image coordinates $\mathbf{f}_0 = [x_0/z_0, y_0/z_0, 1]^\mathsf{T}$ and $\mathbf{f}_1 = [x_1/z_1, y_1/z_1, 1]^\mathsf{T}$ are related to $\mathbf{u}_0$ and $\mathbf{u}_1$ by $\mathbf{u}_0 = \mathbf{K}\mathbf{f}_0$ and $\mathbf{u}_1 = \mathbf{K}\mathbf{f}_1$.
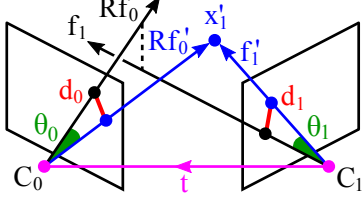
Figure 1. The difference between the observed features $(\mathbf{f}_0, \mathbf{f}_1)$ and the triangulation result $(\mathbf{f}'_0, \mathbf{f}'_1)$ can be quantified by either image reprojection errors (d_0, d_1) or angular reprojection errors (θ_0, θ_1).

The two backprojected rays in frame C_1, i.e., $\mathbf{r}_1(s_1) = s_1\mathbf{f}_1$ and $\mathbf{r}_0(s_0) = s_0\mathbf{R}\mathbf{f}_0 + \mathbf{t}$, do not necessarily intersect due to inaccuracies in the image measurements and camera matrices. For the rays to intersect, $\mathbf{f}_0$ and $\mathbf{f}_1$ must be corrected to $\mathbf{f}'_0$ and $\mathbf{f}'_1$ such that the epipolar constraint [19] is satisfied. It enforces the coplanarity of $\mathbf{f}'_1$, $\mathbf{R}\mathbf{f}'_0$ and $\mathbf{t}$, and is given by

$$\mathbf{f}'_1 \cdot (\mathbf{t} \times \mathbf{R}\mathbf{f}'_0) = 0. \tag{8}$$

The goal of the optimal triangulation is to *minimally* correct the feature rays so that they satisfy (8) and intersect at some point $\mathbf{x}'_1$ in frame C_1. What is meant by "minimal" depends on the chosen cost function and error criterion. Fig. 1 illustrates two most popular error criteria, namely the image reprojection error and the angular reprojection error. Formally, they are defined as follows:

$$d_i := \|\mathbf{u}_i - \mathbf{u}'_i\| = \|\mathbf{K}(\mathbf{f}_i - \mathbf{f}'_i)\|, \qquad \text{for } i = 0, 1 \tag{9}$$

$$\theta_i := \angle(\mathbf{f}_i, \mathbf{f}'_i) = \angle(\mathbf{K}^{-1}\mathbf{u}_i, \mathbf{K}^{-1}\mathbf{u}'_i) \quad \text{for } i = 0, 1 \tag{10}$$

In this work, we minimize the latter in L_1, L_2 and L_∞ norms. Once we have the optimal $\mathbf{f}'_0$ and $\mathbf{f}'_1$, the point of intersection $\mathbf{x}'_1$ can be obtained using either (5) or (6):

$$\mathbf{x}'_1 = \mathbf{t} + \underbrace{\frac{\mathbf{z} \cdot (\mathbf{t} \times \mathbf{f}'_1)}{\|\mathbf{z}\|^2}}_{\lambda_0} \mathbf{R}\mathbf{f}'_0 = \underbrace{\frac{\mathbf{z} \cdot (\mathbf{t} \times \mathbf{R}\mathbf{f}'_0)}{\|\mathbf{z}\|^2}}_{\lambda_1} \mathbf{f}'_1 \tag{11}$$

$$\text{with} \quad \mathbf{z} = \mathbf{f}'_1 \times \mathbf{R}\mathbf{f}'_0,$$

where λ_i equals the depth multiplied by $\|\mathbf{f}'_i\|$ for $i = 0, 1$.

Note that the epipolar constraint (8) is a necessary condition for intersecting the two rays, but not a sufficient one. Fig. 2 illustrates scenarios where the two rays are coplanar, but do not intersect. This happens when the intersection requires negative depth(s), violating the *cheirality* constraint [8]. In the following analysis (until Section 8), we will temporarily assume that satisfying the epipolar constraint (8) is sufficient for intersecting the rays.

5. Closed-Form L_1 Triangulation

The L_1 triangulation based on the angular reprojection error (10) finds the feature rays $\mathbf{f}'_0$ and $\mathbf{f}'_1$ that minimize $\theta_0 + \theta_1$ subject to the epipolar constraint (8). The following lemma reveals a surprising fact that $(\theta_0 + \theta_1)_{\min}$ is achieved by correcting either one of $\mathbf{f}_0$ or $\mathbf{f}_1$, but not both:
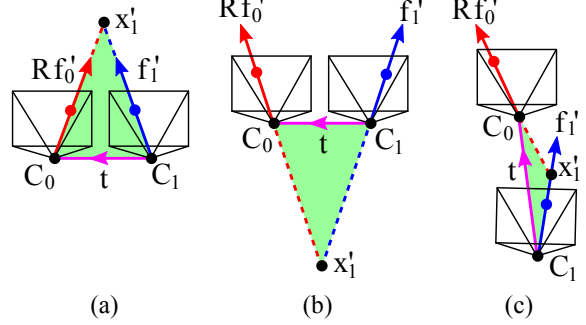

(a)   (b)   (c)

Figure 2. Example scenarios satisfying the epipolar constraint (8). The epipolar plane (shown in green) contains both the rays and the camera centers. Cheirality condition is violated in case (b) and (c).

Lemma 2 (L_1 Angle Minimization)
Given two skew lines $\mathbf{L}_0(s_0) = \mathbf{c}_0 + s_0\mathbf{m}_0$ *and* $\mathbf{L}_1(s_1) = \mathbf{c}_1 + s_1\mathbf{m}_1$, *consider any two intersecting lines that also pass* $\mathbf{c}_0$ *and* $\mathbf{c}_1$, *respectively, i.e.,* $\mathbf{L}'_0(s'_0) = \mathbf{c}_0 + s'_0\mathbf{m}'_0$ *and* $\mathbf{L}'_1(s'_1) = \mathbf{c}_1 + s'_1\mathbf{m}'_1$. *Let* $\mathbf{t} = \mathbf{c}_0 - \mathbf{c}_1$, $\mathbf{n}_0 = \mathbf{m}_0 \times \mathbf{t}$, $\mathbf{n}_1 = \mathbf{m}_1 \times \mathbf{t}$, $\theta_0 = \angle(\mathbf{L}_0, \mathbf{L}'_0)$ *and* $\theta_1 = \angle(\mathbf{L}_1, \mathbf{L}'_1)$. *Then,* $(\theta_0 + \theta_1)$ *is minimized for the following* $\mathbf{m}'_0$ *and* $\mathbf{m}'_1$:
If $\|\hat{\mathbf{m}}_0 \times \mathbf{t}\| \le \|\hat{\mathbf{m}}_1 \times \mathbf{t}\|$,

$$\mathbf{m}'_0 = \mathbf{m}_0 - (\mathbf{m}_0 \cdot \hat{\mathbf{n}}_1)\hat{\mathbf{n}}_1 \quad \text{and} \quad \mathbf{m}'_1 = \mathbf{m}_1. \tag{12}$$

Otherwise,

$$\mathbf{m}'_0 = \mathbf{m}_0 \quad \text{and} \quad \mathbf{m}'_1 = \mathbf{m}_1 - (\mathbf{m}_1 \cdot \hat{\mathbf{n}}_0)\hat{\mathbf{n}}_0 \tag{13}$$

Proof. Refer to Appendix B. ∎

By substituting $\mathbf{R}\mathbf{f}_0$ and $\mathbf{f}_1$ into $\mathbf{m}_0$ and $\mathbf{m}_1$ in the above lemma, the resulting $\mathbf{m}'_0$ and $\mathbf{m}'_1$ become the corrected rays $\mathbf{R}\mathbf{f}'_0$ and $\mathbf{f}'_1$ that satisfy the L_1 optimality, and $\mathbf{n}_0$ (or $\mathbf{n}_1$) becomes the normal of the corresponding epipolar plane.

6. Closed-Form L_2 Triangulation

Considering that the angular errors are small in practice, the "relaxed" L_2 triangulation finds the feature rays $\mathbf{f}'_0$ and $\mathbf{f}'_1$ that minimize $\sin^2(\theta_0) + \sin^2(\theta_1)$ (instead of $\theta_0^2 + \theta_1^2$) subject to the epipolar constraint (8). Note that the small-angle approximation by $\sin(\theta)$ is more accurate than by $\tan(\theta)$ or $1 - \cos(\theta)$ that have been used in literature [7, 9, 12, 28, 3]. This is easily seen by comparing their Maclaurin expansions. As will be shown in the following lemma (and previously in [24]), the relaxation with the sine function allows us to derive the L_2 optimal solution in closed form.

Lemma 3 (L_2 Angle Minimization)
Given two skew lines $\mathbf{L}_0(s_0) = \mathbf{c}_0 + s_0\mathbf{m}_0$ *and* $\mathbf{L}_1(s_1) = \mathbf{c}_1 + s_1\mathbf{m}_1$, *consider any two intersecting lines that also pass* $\mathbf{c}_0$ *and* $\mathbf{c}_1$, *respectively, i.e.,* $\mathbf{L}'_0(s'_0) = \mathbf{c}_0 + s'_0\mathbf{m}'_0$ *and* $\mathbf{L}'_1(s'_1) = \mathbf{c}_1 + s'_1\mathbf{m}'_1$. *Let* $\mathbf{t} = \mathbf{c}_0 - \mathbf{c}_1$, $\theta_0 = \angle(\mathbf{L}_0, \mathbf{L}'_0)$ *and* $\theta_1 = \angle(\mathbf{L}_1, \mathbf{L}'_1)$. *Then,* $(\sin^2\theta_0 + \sin^2\theta_1)$ *is minimized for*

$$\mathbf{m}'_i = \mathbf{m}_i - (\mathbf{m}_i \cdot \hat{\mathbf{n}}')\hat{\mathbf{n}}' \quad \text{for} \quad i = 0, 1, \tag{14}$$

where $\hat{\mathbf{n}}'$ *is the second column of the* 3×3 *matrix* $\mathbf{V}$ *from*

$$\mathbf{U}\mathbf{S}\mathbf{V}^{\mathsf{T}} = SVD\left(\begin{bmatrix}\widehat{\mathbf{m}}_0 & \widehat{\mathbf{m}}_1\end{bmatrix}^{\mathsf{T}}\left(\mathbf{I} - \widehat{\mathbf{t}}\,\widehat{\mathbf{t}}^{\mathsf{T}}\right)\right). \qquad (15)$$

Proof. Refer to Appendix C. ∎

Analogously to the L_1 method, substituting $\mathbf{R}\mathbf{f}_0$ and $\mathbf{f}_1$ into $\mathbf{m}_0$ and $\mathbf{m}_1$ in the above lemma gives $\mathbf{R}\mathbf{f}_0' = \mathbf{m}_0'$ and $\mathbf{f}_1' = \mathbf{m}_1'$ that satisfy the L_2 optimality.

7. Closed-Form L_∞ Triangulation

The L_∞ triangulation based on the angular reprojection error (10) finds the feature rays $\mathbf{f}_0'$ and $\mathbf{f}_1'$ that minimize $\max(\theta_0, \theta_1)$ subject to the epipolar constraint (8). The following lemma states that this is achieved when $\theta_0 = \theta_1$:

Lemma 4 (L_∞ Angle Minimization)
Given two skew lines $\mathbf{L}_0(s_0) = \mathbf{c}_0 + s_0\mathbf{m}_0$ and $\mathbf{L}_1(s_1) = \mathbf{c}_1 + s_1\mathbf{m}_1$, consider any two intersecting lines that also pass $\mathbf{c}_0$ and $\mathbf{c}_1$, respectively, i.e., $\mathbf{L}_0'(s_0') = \mathbf{c}_0 + s_0'\mathbf{m}_0'$ and $\mathbf{L}_1'(s_1') = \mathbf{c}_1 + s_1'\mathbf{m}_1'$. Let $\mathbf{t} = \mathbf{c}_0 - \mathbf{c}_1$, $\mathbf{n}_a = (\widehat{\mathbf{m}}_0 + \widehat{\mathbf{m}}_1) \times \mathbf{t}$, $\mathbf{n}_b = (\widehat{\mathbf{m}}_0 - \widehat{\mathbf{m}}_1) \times \mathbf{t}$, $\theta_0 = \angle(\mathbf{L}_0, \mathbf{L}_0')$ and $\theta_1 = \angle(\mathbf{L}_1, \mathbf{L}_1')$. Then, $\max(\theta_0, \theta_1)$ is minimized when $\theta_0 = \theta_1$. This is achieved for

$$\mathbf{m}_i' = \mathbf{m}_i - (\mathbf{m}_i \cdot \widehat{\mathbf{n}}')\widehat{\mathbf{n}}' \quad for \quad i = 0, 1, \qquad (16)$$

where
$$\mathbf{n}' = \begin{cases} \mathbf{n}_a & if \ \|\mathbf{n}_a\| \geq \|\mathbf{n}_b\| \\ \mathbf{n}_b & otherwise \end{cases} \qquad (17)$$

Proof. Refer to Appendix D. ∎

Analogously to the previous two methods, substituting $\mathbf{R}\mathbf{f}_0$ and $\mathbf{f}_1$ into $\mathbf{m}_0$ and $\mathbf{m}_1$ gives $\mathbf{R}\mathbf{f}_0' = \mathbf{m}_0'$ and $\mathbf{f}_1' = \mathbf{m}_1'$ that satisfy the L_∞ optimality.

8. Cheirality, Parallax and Outliers

We have used the term *lines* instead of *rays* in all lemmas so far, ignoring the cheirality constraint [8]. We argue that if the optimal solution violates the cheirality constraint, the most reasonable choice is to simply discard the result. In the following, we provide the rationale for this choice.

Fig. 4 illustrates five scenarios where the optimal solution violates the cheirality constraint. In case (a), both rays have negative depths at the optimal intersection. Increasing the allowed angular reprojection error, the first intersection with positive depths occurs when the two corrected rays become parallel, resulting in a point at infinity. This point cannot be triangulated, so it should be discarded.

In the remaining cases, the optimal intersection involves only one of the rays having a negative depth. Following the same procedure, the first intersection with positive depths occurs either at infinity (case (b)), at one of the camera centers (case (c)), along the ray parallel to the translation (case (d)), or at a point somewhere else (case (e)).

In case (b), (c) and (d), the newly triangulated point has either infinite, zero or ambiguous depth, so it is reasonable to discard it. In case (e), we found that reattempting the triangulation with positive depths yields either a very large

---

Input: Calib. matrix ($\mathbf{K}$), relative pose ($\mathbf{R}, \mathbf{t}$), and a match ($\mathbf{u}_0, \mathbf{u}_1$) from two views ($C_0, C_1$).
Output: Triangulated 3D point ($\mathbf{x}_1'$) in ref. frame C_1.
1) $\mathbf{f}_0 \leftarrow \mathbf{K}^{-1}\mathbf{u}_0$, $\mathbf{f}_1 \leftarrow \mathbf{K}^{-1}\mathbf{u}_1$, $\mathbf{m}_0 \leftarrow \mathbf{R}\mathbf{f}_0$, $\mathbf{m}_1 \leftarrow \mathbf{f}_1$.
2) **For L_1 triangulation:**
　　If $\|\widehat{\mathbf{m}}_0 \times \mathbf{t}\| \leq \|\widehat{\mathbf{m}}_1 \times \mathbf{t}\|$, use (12) to obtain $\mathbf{m}_0'$ and $\mathbf{m}_1'$. Otherwise, use (13).
　For L_2 triangulation:
　　Compute $\mathbf{m}_0'$ and $\mathbf{m}_1'$ from (14) and (15).
　For L_∞ triangulation:
　　Compute $\mathbf{m}_0'$ and $\mathbf{m}_1'$ from (16) and (17).
3) $\mathbf{R}\mathbf{f}_0' \leftarrow \mathbf{m}_0'$ and $\mathbf{f}_1' \leftarrow \mathbf{m}_1'$.
4) Check cheirality:
　(i) Obtain λ_0 and λ_1 from (11).
　(ii) Discard the point and terminate if either $\lambda_0 \leq 0$ or $\lambda_1 \leq 0$.
5) Check angular reprojection errors:
　(i) $\theta_0 \leftarrow \angle(\mathbf{R}\mathbf{f}_0, \mathbf{R}\mathbf{f}_0')$ and $\theta_1 \leftarrow \angle(\mathbf{f}_1, \mathbf{f}_1')$.
　(ii) Discard the point and terminate if $\max(\theta_0, \theta_1) > \epsilon_1$ for some small ϵ_1.
6) Check parallax:
　(i) $\beta \leftarrow \angle(\mathbf{R}\mathbf{f}_0', \mathbf{f}_1')$
　(ii) Discard the point and terminate if $\beta < \epsilon_2$ for some small ϵ_2.
7) Compute and return $\mathbf{x}_1'$ from (11).

Table 1. Summary of the proposed methods.

error, a point near the epipole or a low parallax angle. Typically, these are the indicators of low accuracy or an outlier [10, 8], so a reasonable choice is to discard the match. This procedure is outlined in Step 4–6 of Tab. 1.

9. Experimental Results

We evaluate the proposed methods in comparison to the midpoint method [2, 10], Hartley and Sturm's L_1 and L_2 method [10], Lindstrom's L_2 method with five iterations [17], and Níster's L_∞ method [23]. The evaluation was performed on both synthetic and real datasets. We generated the synthetic datasets as follows: A set of 8×4 point clouds of 2,500 points each are generated with a Gaussian radial distribution $\mathcal{N}(0, (d/4)^2)$ where d is the distance from the world origin. Each point cloud is centered at $[0, 0, d]^{\mathsf{T}}$ for $d = 2^n$ with $n = -1, 0, ..., +6$, and their image projections are perturbed by Gaussian noise $\mathcal{N}(0, \sigma^2)$ for $\sigma = 0.5, 1, 2, 4, 8$. The size and the focal length of the images are $1,024 \times 1,024$ pixels and 512 pixel, respectively. We have three configurations for the camera poses: (1) "orbital" - the cameras at $[\pm 0.5, 0, 0]^{\mathsf{T}}$ pointing at the point cloud center, (2) "lateral" - the cameras at $[\pm 0.5, 0, 0]^{\mathsf{T}}$ pointing at $[0, 0, \infty]^{\mathsf{T}}$, and (3) "forward" -

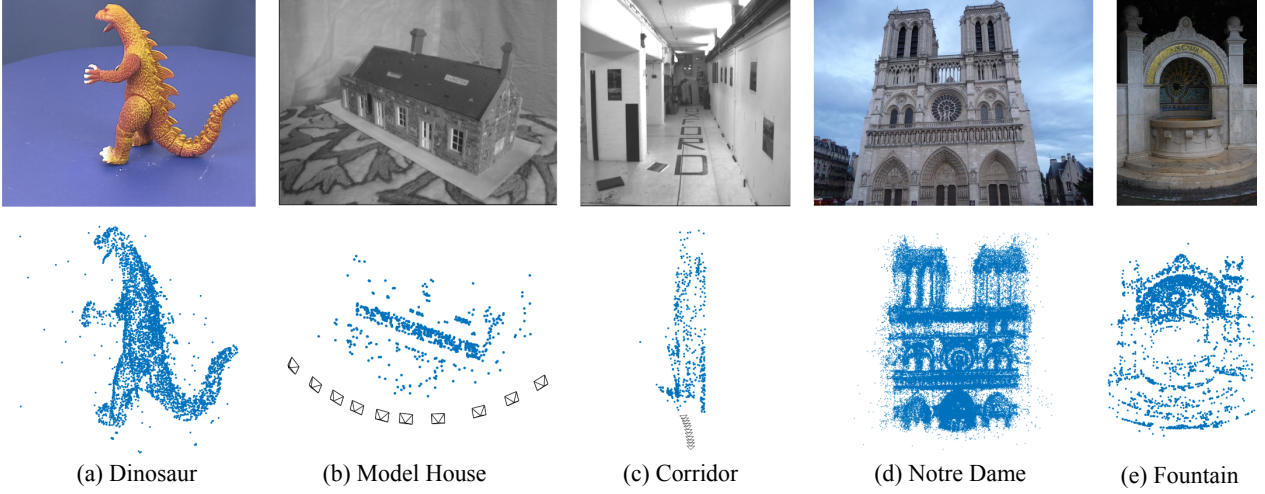(a) Dinosaur   (b) Model House   (c) Corridor   (d) Notre Dame   (e) Fountain

Figure 3. **Top row:** Real dataset images. **Bottom row:** Main segments of the median reconstruction results using the proposed L_1 method.
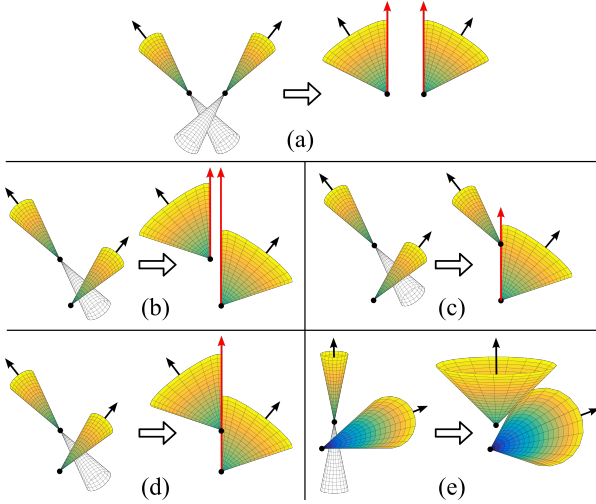

Figure 4. Five scenarios where the optimal solution violates the cheirality constraint, and the possible reattempts for triangulation.

(a) Orbital   (b) Lateral   (c) Forward

Figure 5. 3D triangulation errors before (top) and after (bottom) discarding the points with the lowest 5% parallax.

the cameras at $[0, 0, \pm 0.5]^\intercal$ pointing at the point cloud center. The poses are slightly perturbed with uniform noise $\mathcal{U}(0, 0.01)$. For real datasets, we used the Oxford Dinosaur, Model House and Corridor [1], Notre Dame [31] and Fountain [4, 25] dataset. In total, the synthetic and real datasets provide over 5.5 million unique triangulation problems in a wide variety of geometric configurations.

Tab. 2 provides the percentage of the total matches (from both synthetic and real datasets) for which each method yields the lowest error in given criterion. In 100 % of the total triangulation problems, all three of our methods yield the lowest errors in their corresponding optimal criterion. We also see that minimizing $\sin^2(\theta_0) + \sin^2(\theta_1)$ is very close to minimizing $\theta_0^2 + \theta_1^2$, as discussed in Section 6. Since our L_1 angular method is numerically stable, it sometimes finds better solutions than Hartley-Sturm's closed-form L_1

method [10] even in the L_1 image error criterion $(d_0 + d_1)$.

In Fig. 5, histograms are given for the 3D reconstruction errors on the synthetic datasets. It shows that 1) all methods exhibit similar 3D accuracy, and 2) discarding low-parallax points (Step 6 of Tab. 1) helps to remove large 3D errors. Qualitatively, we also found that the reconstructions of the real datasets look similar for all methods. Fig. 3 shows the reconstruction results using the proposed L_1 method.

We compare the speed of each algorithm in Tab. 3. The midpoint method is the fastest, as it directly computes the 3D point using (11) without correcting the feature rays or image points. Among the optimal methods, our L_1 and L_∞ methods are significantly faster than the rest, i.e., at least 1–2 orders of magnitude faster than the state-of-the-art [17].

5

	Midpoint [2, 10]	L_1 img [10]	L_2 img [10]	L_2 img 5 it. [17]	L_∞ img [23]	L_1 ang	L_2 ang	L_∞ ang
$\theta_0 + \theta_1$	-	-	-	-	-	**100 %**	-	-
$\theta_0^2 + \theta_1^2$	-	-	7e-5 %	5e-5 %	-	-	**99.9999 %**	-
$\sin^2(\theta_0) + \sin^2(\theta_1)$	-	-	-	-	-	-	**100 %**	-
$\max(\theta_0, \theta_1)$	-	-	-	-	-	-	-	**100 %**
$d_0 + d_1$	-	**70.84%**	0.002%	0.002%	-	29.16 %	-	-
$d_0^2 + d_1^2$	-	-	23.14 %	**76.86 %**	-	-	-	-
$\max(d_0, d_1)$	-	-	-	-	**100 %**	-	-	-

Table 2. Percentage of the total matches (from all synthetic and real datasets) for which each method yields the lowest error in given criterion. "img/ang": optimal in the image/angular errors. See the supplementary material for the results from individual datasets.

	Midpoint [2, 10]	L_1 img [10]	L_2 img [10]	L_∞ img [23]	L_2 img 2 it. [17]	L_2 img 5 it. [17]	L_1 ang	L_2 ang	L_∞ ang
Points/sec	42 M	65 K	92 K	270 K	1.4 M	520 K	29 M	670 K	14 M
Relative Speed	1.0	0.0016	0.0022	0.0064	0.033	0.013	0.71	0.016	0.33

Table 3. Speed of computing a 3D point. The relative speed is normalized by that of the midpoint method. Note that this does not take into account Step 4–6 of Tab. 1. All algorithms were implemented in C++ and run on a laptop CPU (Intel i7-4810MQ, 2.8 GHz).

10. Conclusions

In this work, we derived optimal closed-form solutions to the L_1, L_2 and L_∞ stereo triangulation based on the angular reprojection error. The proposed triangulation methods are extremely simple and fast, and they guarantee global optimality under respective cost functions. We believe that our findings will be particularly useful for large-scale SfM and real-time visual SLAM algorithms.

Acknowledgement

Appendix

A. Proof of Lemma 1

Consider a right circular cone with apex $\mathbf{c}_0$ and axis $\mathbf{L}_0$, lying sideways on a plane Π that contains $\mathbf{c}_0$ and line $\mathbf{L}_1$ (see Fig. 6). The equation of the plane is given by

$$\Pi(\mathbf{x}) = \mathbf{n}_1 \cdot (\mathbf{x} - \mathbf{c}_0) = 0 \quad \text{with} \quad \mathbf{n}_1 = \mathbf{m}_1 \times \mathbf{t}. \quad (18)$$

The line of intersection between the plane and the cone forms the smallest angle to $\mathbf{L}_0$ among all possible lines on the plane that pass $\mathbf{c}_0$. That is, it forms the smallest angle to $\mathbf{L}_0$ among all possible lines that pass both $\mathbf{c}_0$ and $\mathbf{L}_1$. Hence, this line of intersection must be $\mathbf{L}_0'$. Now, consider a point $\mathbf{a}_0 = \mathbf{c}_0 + \widehat{\mathbf{m}}_0$ located one unit away from $\mathbf{c}_0$ along $\mathbf{L}_0$. Let $\mathbf{r}$ be the projection of $\mathbf{a}_0$ onto plane Π. According to lemma 5 in Appendix E, the point $\mathbf{r}$ must be located
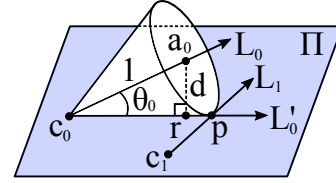
Figure 6. The angle θ_0 is the smallest angle required for pivoting line $\mathbf{L}_0$ at point $\mathbf{c}_0$ to make it intersect $\mathbf{L}_1$.

along $\mathbf{L}_0'$. Let $d = \|\mathbf{a}_0 - \mathbf{r}\|$, i.e., the distance between $\mathbf{a}_0$ and plane Π. Then, we obtain $\sin(\theta_0)$ as follows:

$$\sin(\theta_0) = d \overset{(3)}{=} |\widehat{\mathbf{n}}_1 \cdot (\mathbf{a}_0 - \mathbf{c}_0)| = |\widehat{\mathbf{n}}_1 \cdot \widehat{\mathbf{m}}_0|. \quad \blacksquare$$

B. Proof of Lemma 2

One of the following is true when $(\theta_0 + \theta_1)$ is minimized:

1. $\mathbf{L}_0' \neq \mathbf{L}_0$ and $\mathbf{L}_1' = \mathbf{L}_1 \longleftrightarrow \theta_0 > 0$ and $\theta_1 = 0$.
2. $\mathbf{L}_0' = \mathbf{L}_0$ and $\mathbf{L}_1' \neq \mathbf{L}_1 \longleftrightarrow \theta_0 = 0$ and $\theta_1 > 0$.
3. $\mathbf{L}_0' \neq \mathbf{L}_0$ and $\mathbf{L}_1' \neq \mathbf{L}_1 \longleftrightarrow \theta_0 > 0$ and $\theta_1 > 0$.

Suppose, for the sake of argument, that one of the first two statements is true. In the first case, lemma 1 states that $\mathbf{m}_0'$ is obtained by projecting $\mathbf{m}_0$ onto the plane with the normal $\mathbf{m}_1 \times \mathbf{t}$, which leads to (12) and

$$\sin(\theta_0) = \frac{|\widehat{\mathbf{m}}_0 \cdot (\mathbf{m}_1 \times \mathbf{t})|}{\|\mathbf{m}_1 \times \mathbf{t}\|} = \frac{|\widehat{\mathbf{m}}_0 \cdot (\widehat{\mathbf{m}}_1 \times \mathbf{t})|}{\|\widehat{\mathbf{m}}_1 \times \mathbf{t}\|}. \quad (19)$$

Likewise, in the second case, lemma 1 leads to (13), and

$$\sin(\theta_1) = \frac{|\widehat{\mathbf{m}}_1 \cdot (\mathbf{m}_0 \times \mathbf{t})|}{\|\mathbf{m}_0 \times \mathbf{t}\|} \overset{(1)}{=} \frac{|\widehat{\mathbf{m}}_0 \cdot (\widehat{\mathbf{m}}_1 \times \mathbf{t})|}{\|\widehat{\mathbf{m}}_0 \times \mathbf{t}\|}. \quad (20)$$

Now, the question is how to determine which of the two statements is true. Comparing the right-hand side of (19) and (20), we find that if $\|\widehat{\mathbf{m}}_0 \times \mathbf{t}\| \leq \|\widehat{\mathbf{m}}_1 \times \mathbf{t}\|$, then

$$\min_{\theta_0|\theta_1=0} \theta_0 \leq \min_{\theta_1|\theta_0=0} \theta_1, \tag{21}$$

and $\min(\theta_0 + \theta_1)$ is equal to the left-hand side of (21), indicating that the first statement is true. Naturally, the second statement is true otherwise. Note that there is an ambiguity if $\|\widehat{\mathbf{m}}_0 \times \mathbf{t}\| = \|\widehat{\mathbf{m}}_1 \times \mathbf{t}\|$, and the solution is optimal whichever case is considered. This concludes the proof of lemma 2 for the first two cases.

We will now prove that the third case never occurs. Given some angle θ_1, minimizing $(\theta_0 + \theta_1)$ is equivalent to minimizing θ_0. This is the identical situation as the first case if we replace $\mathbf{L}_1$ by $\mathbf{L}_1'$. We know from the proof of lemma 1 that pivoting a line to intersect another with minimum angle can be modeled by a cone lying sideways on a plane. The top of Fig. 7 illustrates this. Similarly, the bottom of Fig. 7 illustrates the minimization of θ_1 with respect to $\mathbf{L}_0'$ given θ_0. Now, since both planes touching each cone contain the same two intersecting lines $\mathbf{L}_0'$ and $\mathbf{L}_1'$, they must be the same plane. Let this plane be Π'. According to lemma 1, $\mathbf{L}_0'$ is the projection of $\mathbf{L}_0$ onto plane Π'. Therefore,

$$\widehat{\mathbf{m}}_0' = \frac{\widehat{\mathbf{m}}_0 - (\widehat{\mathbf{m}}_0 \cdot \widehat{\mathbf{n}}')\widehat{\mathbf{n}}'}{\|\widehat{\mathbf{m}}_0 - (\widehat{\mathbf{m}}_0 \cdot \widehat{\mathbf{n}}')\widehat{\mathbf{n}}'\|}, \tag{22}$$

where $\widehat{\mathbf{n}}'$ is the unit normal of plane Π'. Since Π' contains both $\mathbf{c}_0$ and $\mathbf{c}_1$, $\widehat{\mathbf{n}}'$ is perpendicular to $\mathbf{t} = \mathbf{c}_0 - \mathbf{c}_1$. Hence, computing the dot product with $\mathbf{t}$ on each side of (22) yields

$$\mathbf{t} \cdot \widehat{\mathbf{m}}_0' = \frac{\mathbf{t} \cdot \widehat{\mathbf{m}}_0}{\|\widehat{\mathbf{m}}_0 - (\widehat{\mathbf{m}}_0 \cdot \widehat{\mathbf{n}}')\widehat{\mathbf{n}}'\|}. \tag{23}$$

Note that $\|\widehat{\mathbf{m}}_0 - (\widehat{\mathbf{m}}_0 \cdot \widehat{\mathbf{n}}')\widehat{\mathbf{n}}'\|$ corresponds to the magnitude of the projection of $\widehat{\mathbf{m}}_0$ onto plane Π' for non-zero θ_0, so it must be smaller than $\|\widehat{\mathbf{m}}_0\| = 1$. Thus,

$$|\mathbf{t} \cdot \widehat{\mathbf{m}}_0'| = \frac{|\mathbf{t} \cdot \widehat{\mathbf{m}}_0|}{\|\widehat{\mathbf{m}}_0 - (\widehat{\mathbf{m}}_0 \cdot \widehat{\mathbf{n}}')\widehat{\mathbf{n}}'\|} > |\mathbf{t} \cdot \widehat{\mathbf{m}}_0|. \tag{24}$$

Using (2), this inequality can be written as

$$\|\mathbf{t} \times \widehat{\mathbf{m}}_0'\| < \|\mathbf{t} \times \widehat{\mathbf{m}}_0\|. \tag{25}$$

Analogously, we can also derive

$$\|\mathbf{t} \times \widehat{\mathbf{m}}_1'\| < \|\mathbf{t} \times \widehat{\mathbf{m}}_1\|. \tag{26}$$

Now, suppose that

$$\min_{\theta_0,\theta_1} (\theta_0 + \theta_1) = \theta_0^* + \theta_1^* \quad \text{with} \quad \theta_0^*, \theta_1^* > 0. \tag{27}$$

Without loss of generality, let us assume that $\|\mathbf{t} \times \widehat{\mathbf{m}}_0\| \leq \|\mathbf{t} \times \widehat{\mathbf{m}}_1\|$. Then, (25) gives $\|\mathbf{t} \times \widehat{\mathbf{m}}_0'\| < \|\mathbf{t} \times \widehat{\mathbf{m}}_1\|$. As we discussed for the first two cases, this means that pivoting $\mathbf{L}_0'$ to intersect $\mathbf{L}_1$ takes smaller angle than pivoting $\mathbf{L}_1$ to intersect $\mathbf{L}_0'$, i.e., $\theta_0' < \theta_1^*$. Thus
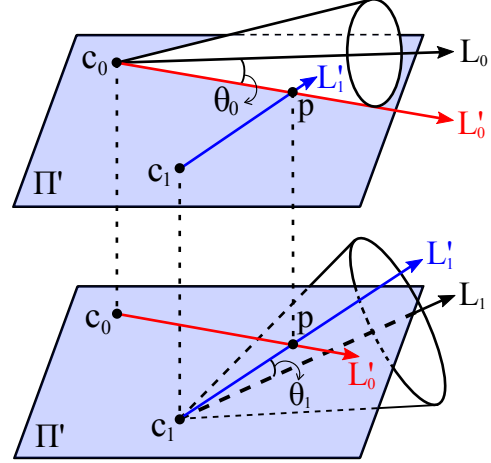
Figure 7. When two cones intersect at a single point $\mathbf{p}$ on their lateral surface, they are tangent to the same plane containing $\mathbf{p}$ and the apexes of each cone. For visualization purposes, we show the two cones lying on each side of the plane separately.

$$\theta_0^* + \theta_0' < \theta_0^* + \theta_1^*. \tag{28}$$

According to lemma 6 in Appendix E, pivoting a line twice for intersection takes equal or greater angle than the single minimum pivot angle. Therefore,

$$\min_{\theta_0|\theta_1=0} \theta_0 \leq \theta_0^* + \theta_0' < \theta_0^* + \theta_1^*, \tag{29}$$

which contradicts (27). Therefore, $(\theta_0 + \theta_1)$ is minimized when either θ_0 or θ_1 is zero. ∎

C. Proof of Lemma 3

Given some angle θ_1, $\left(\sin^2(\theta_0) + \sin^2(\theta_1)\right)_{\min}$ is achieved by minimizing θ_0 and vice versa. As discussed in the proof of lemma 2, this means that the underlying geometry at $\left(\sin^2(\theta_0) + \sin^2(\theta_1)\right)_{\min}$ can be represented by the two cones with apex $\mathbf{c}_0$, $\mathbf{c}_1$ and skew axes $\mathbf{L}_0$, $\mathbf{L}_1$, respectively, touching each side of the same plane on their lateral surface. This is visualized in Fig. 7. Let $\mathbf{n}'$ be the normal of plane Π'. From Lemma 1, we know that

$$\sin(\theta_0) = |\widehat{\mathbf{n}}' \cdot \widehat{\mathbf{m}}_0| \quad \text{and} \quad \sin(\theta_1) = |\widehat{\mathbf{n}}' \cdot \widehat{\mathbf{m}}_1|. \tag{30}$$

Combining these two equations, we get

$$\sin^2(\theta_0) + \sin^2(\theta_1) = \|\mathbf{M}^\mathsf{T}\widehat{\mathbf{n}}'\|^2$$
$$\text{with} \quad \mathbf{M} = \begin{bmatrix} \widehat{\mathbf{m}}_0 & \widehat{\mathbf{m}}_1 \end{bmatrix}. \tag{31}$$

Since plane Π' contains both $\mathbf{c}_0$ and $\mathbf{c}_1$, $\widehat{\mathbf{n}}'$ is perpendicular to $\mathbf{t} = \mathbf{c}_0 - \mathbf{c}_1$. Therefore, minimizing $(\sin^2(\theta_0) + \sin^2(\theta_1))$ is equivalent to solving the following equality-constrained quadratic programming problem:

$$\underset{\widehat{\mathbf{n}}'}{\arg\min} \|\mathbf{M}^\mathsf{T}\widehat{\mathbf{n}}'\|^2, \text{ s.t. } \|\widehat{\mathbf{n}}'\| = 1 \text{ and } \mathbf{t} \cdot \widehat{\mathbf{n}}' = 0. \tag{32}$$

In [5], it was shown that this problem can be solved using the method of Lagrange multipliers, and $\|\mathbf{M}^\mathsf{T}\widehat{\mathbf{n}}'\|^2$ is

minimized when $\widehat{\mathbf{n}}'$ is the eigenvector corresponding to the smallest nontrivial eigenvalue of $\mathbf{A} = (\mathbf{I} - \widehat{\mathbf{t}}\,\widehat{\mathbf{t}}^{\mathsf{T}})\mathbf{M}\mathbf{M}^{\mathsf{T}}$. Letting $\mathbf{P} = (\mathbf{I} - \widehat{\mathbf{t}}\,\widehat{\mathbf{t}}^{\mathsf{T}})$, it can be easily shown that $\mathbf{P} = \mathbf{P}^{\mathsf{T}} = \mathbf{P}^{\mathsf{T}}\mathbf{P} = \mathbf{P}\mathbf{P}^{\mathsf{T}}$. Hence, $\mathbf{A} = \mathbf{P}\mathbf{M}\mathbf{M}^{\mathsf{T}} = \mathbf{P}\mathbf{P}^{\mathsf{T}}\mathbf{M}\mathbf{M}^{\mathsf{T}}$. Note that for any square matrix $\mathbf{X}$ and $\mathbf{Y}$, the eigen-decomposition of $\mathbf{X}\mathbf{Y}$ is the same as that of $\mathbf{Y}\mathbf{X}$. This means that the eigenvectors of $\mathbf{A} = \mathbf{P}\left(\mathbf{P}^{\mathsf{T}}\mathbf{M}\mathbf{M}^{\mathsf{T}}\right)$ are the same as those of $\left(\mathbf{P}^{\mathsf{T}}\mathbf{M}\mathbf{M}^{\mathsf{T}}\right)\mathbf{P} = (\mathbf{M}^{\mathsf{T}}\mathbf{P})^{\mathsf{T}}(\mathbf{M}^{\mathsf{T}}\mathbf{P})$, i.e., the right-singular vectors of $\mathbf{M}^{\mathsf{T}}\mathbf{P}$. Therefore, letting $\mathbf{U}\mathbf{S}\mathbf{V}^{\mathsf{T}} = \mathrm{SVD}\left(\mathbf{M}^{\mathsf{T}}\mathbf{P}\right)$ with the diagonal entries of $\mathbf{S}$ in descending order, the optimal $\widehat{\mathbf{n}}'$ is given by the second column of $\mathbf{V}$. Finally, projecting $\mathbf{m}_0$ and $\mathbf{m}_1$ onto plane Π' leads to (14). ∎

D. Proof of Lemma 4

First, we show that $\theta_0 = \theta_1$ when $\max(\theta_0, \theta_1)$ is minimized: Consider two cones with apex $\mathbf{c}_0$, $\mathbf{c}_1$ and skew axes $\mathbf{L}_0$, $\mathbf{L}_1$. Constrain both their apertures to be 2θ. When $\theta = 0$, the they are simply two skew lines. As we gradually increase θ, they will grow at the same rate, and eventually, touch one another. Let $\theta = \theta'$ at this point. Now, suppose

$$\theta^* := \min_{\theta_0, \theta_1} \max(\theta_0, \theta_1) < \theta'. \tag{33}$$

The definition of θ^* implies that setting $\theta_0 = \theta_1 = \theta^*$ will make the two cones partially overlap in space (or at least meet at a point). However, the they do not meet when $\theta_0 = \theta_1 < \theta'$. This is a contradiction, so the inequality in (33) must be false, and θ^* must be equal to θ'. That is, $\theta_0 = \theta_1 = \theta'$ in order for $\max(\theta_0, \theta_1)$ to be minimized.

We can now represent the underlying geometry at $(\max(\theta_0, \theta_1))_{\min}$ as two congruent cones with skew axes, touching each side of the same plane Π' on their lateral surface. This is the situation shown in Fig. 7 for $\theta_0 = \theta_1$. Let $\mathbf{n}'$ be the normal of plane Π'. Then, from lemma 1, we get

$$\sin(\theta_0) = \sin(\theta_1) = |\widehat{\mathbf{n}}' \cdot \widehat{\mathbf{m}}_0| = |\widehat{\mathbf{n}}' \cdot \widehat{\mathbf{m}}_1|. \tag{34}$$

The last equality in (34) can be written as

$$(\widehat{\mathbf{m}}_0 + w\widehat{\mathbf{m}}_1) \cdot \widehat{\mathbf{n}}' = 0, \tag{35}$$

where w is -1 or 1, depending on the signs of $\widehat{\mathbf{n}}' \cdot \widehat{\mathbf{m}}_0$ and $\widehat{\mathbf{n}}' \cdot \widehat{\mathbf{m}}_1$. On the other hand, since plane Π' contains both $\mathbf{c}_0$ and $\mathbf{c}_1$, $\widehat{\mathbf{n}}'$ is perpendicular to $\mathbf{t} = \mathbf{c}_0 - \mathbf{c}_1$:

$$\mathbf{t} \cdot \widehat{\mathbf{n}}' = 0. \tag{36}$$

Combining (35) and (36), $\widehat{\mathbf{n}}'$ can be expressed as

$$\widehat{\mathbf{n}}' = \lambda(\widehat{\mathbf{m}}_0 + w\widehat{\mathbf{m}}_1) \times \mathbf{t} \tag{37}$$

where λ is the normalizing factor. Evaluating (37) at $w = 1$ and $w = -1$ gives two candidates for optimal $\widehat{\mathbf{n}}'$. The optimal solution is then determined by comparing the values of (34) with each candidate $\widehat{\mathbf{n}}'$, which amounts to choosing the solution with smaller λ. This procedure corresponds to (17). Finally, projecting $\mathbf{m}_0$ and $\mathbf{m}_1$ onto plane Π' with optimal $\widehat{\mathbf{n}}'$ leads to (16). ∎
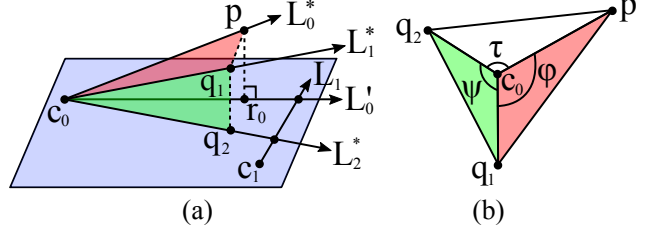
Figure 8. **(a)** Pivoting line $\mathbf{L}_0^*$ in two steps ($\mathbf{L}_0^* \to \mathbf{L}_1^* \to \mathbf{L}_2^*$) to make it intersect to another line $\mathbf{L}_1$. **(b)** A tetrahedron formed by point $\mathbf{c}_0$, p, $\mathbf{q}_1$ and $\mathbf{q}_2$.

E. Other Geometric Lemmas

Lemma 5 (Cone-On-Plane Perpendicularity)
When a plane is tangent to a right circular cone, the line of intersection is the projection of cone's axis onto the plane.

Proof. Consider a cone with axis $\mathbf{L}_0$ and plane Π tangent to this cone. They are both symmetric with respect to the plane that contains $\mathbf{L}_0$ and the normal of Π. Let this plane of symmetry be Π_{sym}. For any circular cross-section of the cone, there is a single point touching Π. Therefore, this point must lie on Π_{sym}, and so must the line of intersection, $\mathbf{L}_0'$. It follows that Π_{sym} contains $\mathbf{L}_0$ and $\mathbf{L}_0'$. Since Π_{sym} is perpendicular to Π, $\mathbf{L}_0'$ is a projection of $\mathbf{L}_0$ onto Π. ∎

Lemma 6 (Single vs Multi-Pivot for Intersection)
Given two skew lines $\mathbf{L}_0(s_0) = \mathbf{c}_0 + s_0\mathbf{m}_0$ and $\mathbf{L}_1(s_1) = \mathbf{c}_1 + s_1\mathbf{m}_1$, let $\mathbf{L}_0'$ be the line that forms the smallest angle $\theta_0 \in [0, \pi/2]$ to $\mathbf{L}_0$ among all possible lines that intersect both point $\mathbf{c}_0$ and line $\mathbf{L}_1$. For any positive integer N, consider the following arbitrary lines passing $\mathbf{c}_0$ such that

$$\mathbf{L}_i^*(s_i^*) = \begin{cases} \mathbf{L}_0 & for \quad i = 0 \\ \mathbf{c}_0 + s_i^*\mathbf{m}_i^* & for \quad i = 1, 2, \cdots, N \end{cases}$$

where only $\mathbf{L}_N^*$ intersects $\mathbf{L}_1$. Then,

$$\theta_0 \le \sum_{i=1}^{N} \angle(\mathbf{L}_i^*, \mathbf{L}_{i-1}^*). \tag{38}$$

Proof. The right-hand side of (38) corresponds to the sum of N pivot angles that make line $\mathbf{L}_0$ to intersect $\mathbf{L}_1$. Fig. 8a depicts such operation for $N = 2$. Let $\phi = \angle(\mathbf{L}_0^*, \mathbf{L}_1^*)$, $\psi = \angle(\mathbf{L}_1^*, \mathbf{L}_2^*)$ and $\tau = \angle(\mathbf{L}_0^*, \mathbf{L}_2^*)$. Now, consider three arbitrary points p, $\mathbf{q}_1$ and $\mathbf{q}_2$ on $\mathbf{L}_0^*$, $\mathbf{L}_1^*$ and $\mathbf{L}_2^*$, respectively. A tetrahedron formed by these three points and $\mathbf{c}_0$ are shown in Fig. 8b. At a vertex of a tetrahedron, the three edges form three angles such that the sum of any two angles is greater than the third one [18, 15]. Thus, $\tau \le \phi + \psi$. Since θ_0 is the minimum pivot angle for intersection, we have $\theta_0 \le \tau \le \phi + \psi$, which proves (38) for $N = 2$. Now, for $N > 2$, we know that replacing the last two pivots by the corresponding single minimum pivot will produce $N-1$ pivots that take equal or smaller angle. Repeating this process until $N = 1$ proves (38) for any $N > 2$. ∎

References

[1] Oxford Multiview Datasets. http://www.robots.ox.ac.uk/~vgg/data/data-mview.html. 5

[2] P. A. Beardsley, A. Zisserman, and D. W. Murray. Navigation using affine structure from motion. In *Eur. Conf. on Computer Vision*, pages 85–96, 1994. 4, 6, 10, 11, 12

[3] G. Chesi. LMI-based estimation of scene points in vision systems with generalized cameras. *IEEE Trans. Automatic Control*, 59(11):2996–3001, 2014. 2, 3

[4] O. Enqvist, F. Kahl, and C. Olsson. Non-sequential structure from motion. In *IEEE Int. Conf. on Computer Vision*, pages 264–271, 2011. 5

[5] G. H. Golub. Some modified matrix eigenvalue problems. *SIAM Review*, 15(2):318–334, 1973. 7

[6] R. Hartley and F. Kahl. Optimal algorithms in multiview geometry. In Y. Yagi, S. B. Kang, I. S. Kweon, and H. Zha, editors, *Asian Conf. on Computer Vision*, pages 13–34, 2007. 2

[7] R. Hartley, F. Kahl, C. Olsson, and Y. Seo. Verifying global minima for L2 minimization problems in multiple view geometry. *International Journal of Computer Vision*, 101(2):288–304, 2013. 2, 3

[8] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, New York, NY, USA, 2 edition, 2003. 1, 3, 4

[9] R. I. Hartley and F. Schaffalitzky. L_∞ minimization in geometric reconstruction problems. In *IEEE Conf. on Computer Vision and Pattern Recognition*, pages 504–509, 2004. 2, 3

[10] R. I. Hartley and P. Sturm. Triangulation. *Comput. Vis. Image Underst.*, 68(2):146–157, 1997. 1, 2, 4, 5, 6, 10, 11, 12

[11] F. Kahl, S. Agarwal, M. K. Chandraker, D. Kriegman, and S. Belongie. Practical global optimization for multiview geometry. *International Journal of Computer Vision*, 79(3):271–284, 2008. 2

[12] F. Kahl and R. Hartley. Multiple-view geometry under the L_∞-norm. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(9):1603–1617, 2008. 2, 3

[13] K. Kanatani, Y. Sugaya, and H. Niitsuma. Triangulation from two views revisited: Hartley-Sturm vs. optimal correction. In *British Machine Vision Conference*, pages 173–182, 2008. 1

[14] Y. Kanazawa and K. Kanatani. Reliability of 3-D reconstruction by stereo vision. *IECE Trans. Inf. & Syst.*, E78-D(10):1301–1306, 1995. 2

[15] M. S. Klamkin. Vector proofs in solid geometry. *The American Mathematical Monthly*, 77(10):1051–1065, 1970. 8

[16] H. Li. A practical algorithm for L_∞ triangulation with outliers. *IEEE Conf. on Computer Vision and Pattern Recognition*, 2007. 2

[17] P. Lindstrom. Triangulation made easy. In *IEEE Conf. on Computer Vision and Pattern Recognition*, pages 1554–1561, 2010. 1, 2, 4, 5, 6, 10, 11, 12

[18] L. Lines. *Solid Geometry*. Macmillan and co., limited, 1935. 8

[19] H. C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293(5828):133–135, 1981. 2, 3

[20] B. Mičušík and T. Pajdla. Structure from motion with wide circular field of view cameras. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(7):1135–1149, 2006. 2

[21] E. Mouragnon, M. Lhuillier, M. Dhome, F. Dekeyser, and P. Sayd. Generic and real-time structure from motion using local bundle adjustment. *Image Vision Comput.*, 27(8):1178–1193, 2009. 1

[22] R. Mur-Artal and J. D. Tardós. ORB-SLAM2: an open-source SLAM system for monocular, stereo and RGB-D cameras. *IEEE Trans. Robotics*, 33(5):1255–1262, 2017. 1

[23] D. Níster. *Automatic Dense Reconstruction from Uncalibrated Video Sequences*. PhD thesis, KTH, 2001. 1, 2, 4, 6, 10, 11, 12

[24] J. Oliensis. Exact two-image structure from motion. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(12):1618–1633, 2002. 1, 2, 3

[25] C. Olsson, O. Enqvist, and F. Kahl. Stable structure from motion for unordered image collections. In *Image Analysis*, pages 524–535, 2011. 5

[26] C. Olsson, A. Eriksson, and R. Hartley. Outlier removal using duality. In *IEEE Conf. on Computer Vision and Pattern Recognition*, pages 1450–1457, 2010. 2

[27] A. Pagani and D. Stricker. Structure from motion using full spherical panoramic cameras. In *IEEE Int. Conf. on Computer Vision Workshops*, pages 375–382, 2011. 1

[28] S. Recker, M. Hess-Flores, and K. I. Joy. Statistical angular error-based triangulation for efficient and accurate multiview scene reconstruction. In *IEEE Workshop on Applications of Computer Vision*, pages 68–75, 2013. 2, 3

[29] J. L. Schönberger and J. Frahm. Structure-from-Motion revisited. In *IEEE Conf. on Computer Vision and Pattern Recognition*, pages 4104–4113, 2016. 1

[30] K. Sim and R. Hartley. Removing outliers using the L_∞ norm. In *IEEE Conf. on Computer Vision and Pattern Recognition*, volume 1, pages 485–494, 2006. 2

[31] N. Snavely, S. M. Seitz, and R. Szeliski. Photo tourism: exploring photo collections in 3D. *ACM Trans. Graph.*, 25(3):835–846, 2006. 5

[32] H. Stewénius, F. Schaffalitzky, and D. Nister. How hard is 3-view triangulation really? In *IEEE Int. Conf. on Computer Vision*, volume 1, pages 686–693, 2005. 1

[33] B. Tippetts, D. J. Lee, K. Lillywhite, and J. Archibald. Review of stereo vision algorithms and their suitability for resource-limited systems. *Journal of Real-Time Image Processing*, 11(1):5–25, 2016. 1

Supplementary Materials

Tab. 5–12 present the results of the error criteria comparison from each individual dataset considered in the main paper. See Tab. 4 for the statistics of the datasets. Note that combining the results in Tab. 5–12 leads to Tab. 2 of the main paper.

	Orbital	Lateral	Forward	Dinosaur	Model House	Corridor	Fountain	Notre Dame
Cameras	16	2	2	36	10	11	14	715
Points	10^5	10^5	10^5	4,983	672	737	5,302	127,431
Matches	10^5	10^5	10^5	27,080	5,550	12,139	58,815	5,120,077

Table 4. Datasets used. "Matches" denotes the number of pairwise instances where a point is visible in two camera views. Generally, the number of matches is larger than the number of points because the points may be visible in more than two views.

		Midpoint [2, 10]	L_1 img [10]	L_2 img [10]	L_2 img 5 it. [17]	L_∞ img [23]	L_1 ang	L_2 ang	L_∞ ang
Error Criterion	$\theta_0 + \theta_1$	-	-	-	-	-	100,000	-	-
	$\theta_0^2 + \theta_1^2$	-	-	-	-	-	-	100,000	-
	$\sin^2(\theta_0) + \sin^2(\theta_1)$	-	-	-	-	-	-	100,000	-
	$\max(\theta_0, \theta_1)$	-	-	-	-	-	-	-	100,000
	$d_0 + d_1$	-	81,654	-	-	-	18,346	-	-
	$d_0^2 + d_1^2$	-	-	23,435	76,565	-	-	-	-
	$\max(d_0, d_1)$	-	-	-	-	100,000	-	-	-

Table 5. **[Orbital]** The number of matches for which each method yields the lowest error in given criterion.

		Midpoint [2, 10]	L_1 img [10]	L_2 img [10]	L_2 img 5 it. [17]	L_∞ img [23]	L_1 ang	L_2 ang	L_∞ ang
Error Criterion	$\theta_0 + \theta_1$	-	-	-	-	-	100,000	-	-
	$\theta_0^2 + \theta_1^2$	-	-	-	-	-	-	100,000	-
	$\sin^2(\theta_0) + \sin^2(\theta_1)$	-	-	-	-	-	-	100,000	-
	$\max(\theta_0, \theta_1)$	-	-	-	-	-	-	-	100,000
	$d_0 + d_1$	-	99,814	1	2	-	183	-	-
	$d_0^2 + d_1^2$	-	-	11,485	88,515	-	-	-	-
	$\max(d_0, d_1)$	-	-	-	-	100,000	-	-	-

Table 6. **[Lateral]** The number of matches for which each method yields the lowest error in given criterion.

		Midpoint [2, 10]	L_1 img [10]	L_2 img [10]	L_2 img 5 it. [17]	L_∞ img [23]	L_1 ang	L_2 ang	L_∞ ang
Error Criterion	$\theta_0 + \theta_1$	-	-	-	-	-	100,000	-	-
	$\theta_0^2 + \theta_1^2$	-	-	4	3	-	-	99,993	-
	$\sin^2(\theta_0) + \sin^2(\theta_1)$	-	-	-	-	-	-	100,000	-
	$\max(\theta_0, \theta_1)$	-	-	-	-	-	-	-	100,000
	$d_0 + d_1$	-	7,625	-	-	-	92,375	-	-
	$d_0^2 + d_1^2$	-	-	34,785	65,215	-	-	-	-
	$\max(d_0, d_1)$	-	-	-	-	100,000	-	-	-

Table 7. **[Forward]** The number of matches for which each method yields the lowest error in given criterion.

10

Error Criterion	Midpoint [2, 10]	L_1 img [10]	L_2 img [10]	L_2 img 5 it. [17]	L_∞ img [23]	L_1 ang	L_2 ang	L_∞ ang
$\theta_0 + \theta_1$	-	-	-	-	-	27,080	-	-
$\theta_0^2 + \theta_1^2$	-	-	-	-	-	-	27,080	-
$\sin^2(\theta_0) + \sin^2(\theta_1)$	-	-	-	-	-	-	27,080	-
$\max(\theta_0, \theta_1)$	-	-	-	-	-	-	-	27,080
$d_0 + d_1$	-	27,080	-	-	-	-	-	-
$d_0^2 + d_1^2$	-	-	6,598	20,482	-	-	-	-
$\max(d_0, d_1)$	-	-	-	-	27,080	-	-	-

Table 8. **[Dinosaur]** The number of matches for which each method yields the lowest error in given criterion.

Error Criterion	Midpoint [2, 10]	L_1 img [10]	L_2 img [10]	L_2 img 5 it. [17]	L_∞ img [23]	L_1 ang	L_2 ang	L_∞ ang
$\theta_0 + \theta_1$	-	-	-	-	-	5,550	-	-
$\theta_0^2 + \theta_1^2$	-	-	-	-	-	-	5,550	-
$\sin^2(\theta_0) + \sin^2(\theta_1)$	-	-	-	-	-	-	5,550	-
$\max(\theta_0, \theta_1)$	-	-	-	-	-	-	-	5,550
$d_0 + d_1$	-	3,563	-	-	-	1,987	-	-
$d_0^2 + d_1^2$	-	-	2,766	2,784	-	-	-	-
$\max(d_0, d_1)$	-	-	-	-	5,550	-	-	-

Table 9. **[Model House]** The number of matches for which each method yields the lowest error in given criterion.

Error Criterion	Midpoint [2, 10]	L_1 img [10]	L_2 img [10]	L_2 img 5 it. [17]	L_∞ img [23]	L_1 ang	L_2 ang	L_∞ ang
$\theta_0 + \theta_1$	-	-	-	-	-	12,139	-	-
$\theta_0^2 + \theta_1^2$	-	-	-	-	-	-	12,139	-
$\sin^2(\theta_0) + \sin^2(\theta_1)$	-	-	-	-	-	-	12,139	-
$\max(\theta_0, \theta_1)$	-	-	-	-	-	-	-	12,139
$d_0 + d_1$	-	6,053	-	-	-	6,086	-	-
$d_0^2 + d_1^2$	-	-	5,999	6,140	-	-	-	-
$\max(d_0, d_1)$	-	-	-	-	12,139	-	-	-

Table 10. **[Corridor]** The number of matches for which each method yields the lowest error in given criterion.

Error Criterion	Midpoint [2, 10]	L_1 img [10]	L_2 img [10]	L_2 img 5 it. [17]	L_∞ img [23]	L_1 ang	L_2 ang	L_∞ ang
$\theta_0 + \theta_1$	-	-	-	-	-	58,815	-	-
$\theta_0^2 + \theta_1^2$	-	-	-	-	-	-	58,815	-
$\sin^2(\theta_0) + \sin^2(\theta_1)$	-	-	-	-	-	-	58,815	-
$\max(\theta_0, \theta_1)$	-	-	-	-	-	-	-	58,815
$d_0 + d_1$	-	32,557	-	-	-	26,258	-	-
$d_0^2 + d_1^2$	-	-	2,348	56,467	-	-	-	-
$\max(d_0, d_1)$	-	-	-	-	58,815	-	-	-

Table 11. **[Fountain]** The number of matches for which each method yields the lowest error in given criterion.

		Midpoint [2, 10]	L_1 img [10]	L_2 img [10]	L_2 img 5 it. [17]	L_∞ img [23]	L_1 ang	L_2 ang	L_∞ ang
Error Criterion	$\theta_0 + \theta_1$	-	-	-	-	-	5,120,077	-	-
	$\theta_0^2 + \theta_1^2$	-	-	-	-	-	-	5,120,077	-
	$\sin^2(\theta_0) + \sin^2(\theta_1)$	-	-	-	-	-	-	5,120,077	-
	$\max(\theta_0, \theta_1)$	-	-	-	-	-	-	-	5,120,077
	$d_0 + d_1$	-	3,654,620	121	123	-	1,465,213	-	-
	$d_0^2 + d_1^2$	-	-	1,190,827	3,929,250	-	-	-	-
	$\max(d_0, d_1)$	-	-	-	-	5,120,077	-	-	-

Table 12. **[Notre Dame]** The number of matches for which each method yields the lowest error in given criterion.