Tangent Sampson Error: Fast Approximate Two-view Reprojection Error for Central Camera Models

Mikhail Terekhov
ETH Zurich
mterekhov@ethz.ch

Viktor Larsson
Lund University
viktor.larsson@math.lth.se

Figure 1. The figure shows the level sets of the proposed Tangent Sampson error. In contrast to classical Sampson error, the proposed error function applies to arbitrary central camera models; in this case a 360 panoramic image and a fisheye camera. Given a relative pose and a pair of points (highlighted in red), we vary a point on one image while keeping the other fixed, and thus get the values on each image. A given color corresponds to a constant value of reprojection error, with purple indicating near-zero errors along the epipolar curve.

Abstract

In this paper we introduce the Tangent Sampson error, which is a generalization of the classical Sampson error in two-view geometry that allows for arbitrary central camera models. It only requires local gradients of the distortion map at the original correspondences (allowing for pre-computation) resulting in a negligible increase in computational cost when used in RANSAC or local refinement. The error effectively approximates the true-reprojection error for a large variety of cameras, including extremely wide field-of-view lenses that cannot be undistorted to a single pinhole image. We show experimentally that the new error outperforms competing approaches both when used for model scoring in RANSAC and for non-linear refinement of the relative camera pose.

1. Introduction

Estimating two-view geometry from a set of sparse 2D-2D correspondences is a common sub-problem in many vision applications. Each correspondence $(\boldsymbol{p}_1, \boldsymbol{p}_2) \in \mathbb{R}^2 \times \mathbb{R}^2$ constrains the geometry via the *epipolar constraint*,

$$(\boldsymbol{p}_2; 1)^\intercal \boldsymbol{E} (\boldsymbol{p}_1; 1) = 0, \tag{1}$$

where $\boldsymbol{E} = [\boldsymbol{t}]_\times \boldsymbol{R} \in \mathbb{R}^{3\times 3}$ is the essential matrix and the points $\boldsymbol{p}_1, \boldsymbol{p}_2$ are on the normalized image plane. This constraint ensures that the 2D-2D correspondence can be exactly triangulated (i.e. with zero reprojection error) using the camera pair extracted from $\boldsymbol{E}$. However when the measurement $(\boldsymbol{p}_1, \boldsymbol{p}_2)$ is inexact, it is natural to measure how close this correspondence is to being consistent with a given epipolar geometry. This is for example used for determining inlier matches during robust estimation, or as a loss to be minimized when refining an initial estimate of $\boldsymbol{E}$.

Under the assumption of i.i.d. Gaussian noise, the ML-estimate of the closest consistent correspondence, sometimes called the *gold-standard* estimate in the literature [12], is found by minimizing the squared reprojection error,

$$\mathcal{E}_{ML}^2 = \min_{\boldsymbol{X}\in\mathbb{R}^3} \|\boldsymbol{p}_1 - \pi(\boldsymbol{X})\|^2 + \|\boldsymbol{p}_2 - \pi(\boldsymbol{R}\boldsymbol{X} + \boldsymbol{t})\|^2, \tag{2}$$

where $\pi : \mathbb{R}^3 \to \mathbb{R}^2$ is the camera projection function. For identical pinhole cameras with unit aspect ratio and no

skew, this can be reformulated as finding the *corrected* normalized image points, i.e. the closest pair of 2D points that satisfy the epipolar constraint exactly,

$$\mathcal{E}_{ML}^2 = \min_{\hat{\boldsymbol{p}}_1, \hat{\boldsymbol{p}}_2 \in \mathbb{R}^2} \quad \|\boldsymbol{p}_1 - \hat{\boldsymbol{p}}_1\|^2 + \|\boldsymbol{p}_2 - \hat{\boldsymbol{p}}_2\|^2 \qquad (3)$$
$$\text{s.t.} \quad (\hat{\boldsymbol{p}}_2; 1)^\intercal \boldsymbol{E} (\hat{\boldsymbol{p}}_1; 1) = 0.$$

The closed-form solution to (3) can be found as a root to a degree-six polynomial (see [13]), or via iterative methods [20]. In robust estimators, e.g. RANSAC, it is often necessary to evaluate these residuals thousands (or potentially millions) of times for each image pair, which in practice means that cheaper approximate error metrics are favored.

These approximate or surrogate error metrics can be divided into two categories: image-based (approximating the true reprojection error) or geometry-based (measuring some geometric quantity, e.g. angles, instead). In the first category we have the epipolar distance, i.e. distance from the point to the corresponding epipolar line. This can be interpreted as assuming zero noise on one of the keypoints. The error is often computed symmetrically and thus called the *symmetric epipolar distance*, having the form

$$\mathcal{E}_{SED}^2 = \frac{|\boldsymbol{d}_2^\intercal \boldsymbol{E} \boldsymbol{d}_1|^2}{\|\boldsymbol{E}_{12} \boldsymbol{d}_1\|^2} + \frac{|\boldsymbol{d}_2^\intercal \boldsymbol{E} \boldsymbol{d}_1|^2}{\|(\boldsymbol{E}^\intercal)_{12} \boldsymbol{d}_2\|^2}, \qquad (4)$$

where $\boldsymbol{E}_{12} \in \mathbb{R}^{2\times 3}$ refers to the first two rows of $\boldsymbol{E}$, and $\boldsymbol{d}_k = (\boldsymbol{p}_k; 1) \in \mathbb{R}^3$. Another image-based error is the *Sampson error*, which is a first-order approximation of $\mathcal{E}_{ML}$ (see Sec. 2.1 for details):

$$\mathcal{E}_S^2 = \frac{|\boldsymbol{d}_2^\intercal \boldsymbol{E} \boldsymbol{d}_1|^2}{\|\boldsymbol{E}_{12} \boldsymbol{d}_1\|^2 + \|(\boldsymbol{E}^\intercal)_{12} \boldsymbol{d}_2\|^2}. \qquad (5)$$

Both of these errors can after appropriate rescaling (using the focal lengths) be expressed in pixels and thus are natural to interpret in terms of the keypoint detection noise levels.

In the second category, an example is the *cosine error*,

$$\mathcal{E}_{CS}^2 = \frac{|\boldsymbol{d}_2^\intercal \boldsymbol{E} \boldsymbol{d}_1|^2}{\|\boldsymbol{d}_2\|^2 \, \|\boldsymbol{E} \boldsymbol{d}_1\|^2} + \frac{|\boldsymbol{d}_2^\intercal \boldsymbol{E} \boldsymbol{d}_1|^2}{\|\boldsymbol{E}^\intercal \boldsymbol{d}_2\|^2 \, \|\boldsymbol{d}_1\|^2}, \qquad (6)$$

which (symmetrically) measures the cosine between the normal of the epipolar plane $\boldsymbol{E}\boldsymbol{d}_1 \in \mathbb{R}^3$ and corresponding point $\boldsymbol{d}_2 \in \mathbb{R}^3$. Another geometric error is the algebraic residual,

$$\mathcal{E}_{ALG} = |\boldsymbol{d}_2^\intercal \boldsymbol{E} \boldsymbol{d}_1|, \qquad (7)$$

which was given a geometric interpretation in [19].

Most of the literature has focused on the pinhole camera case, and the image-based errors above ($\mathcal{E}_{SED}$ and $\mathcal{E}_S$) measure the error in the image plane assuming this model. For non-pinhole cameras, we either have to undistort the image such that $\mathcal{E}_{SED}$ and $\mathcal{E}_S$ are applicable, or use the geometry-based error metrics (e.g. $\mathcal{E}_{CS}$ or $\mathcal{E}_{ALG}$). However, undistorting to a pinhole image (which is not always possible for large field-of-view cameras) will introduce a radial-bias into the error (see Figure 2).

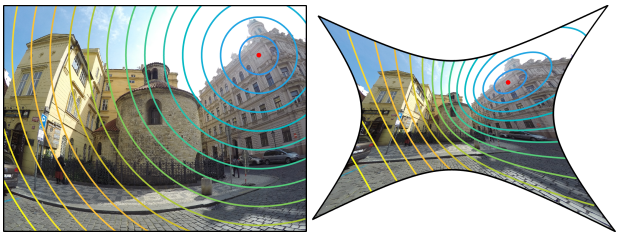

Figure 2. Undistorting the image deforms the level-sets of the reprojection error and introduces a radial bias in the optimization.

In this paper we instead propose the *Tangent Sampson* error $\mathcal{E}_{TS}$, a version of the classical Sampson error that has been generalized to handle arbitrary camera models. Compared to current alternatives it has the following benefits;

- It is **cheap to evaluate** (similar in cost to the original Sampson error) since it only requires computing the camera model forward-projection and Jacobian once (which can be pre-computed for each point).
- Since it **measures errors in the original image**, it is easier to set meaningful thresholds in pixels. This is in contrast to geometric error functions, where it is not always clear how to relate the thresholds (e.g. in angles) to the keypoint detection noise in pixels.
- It works for **arbitrary central camera models**, only requiring that we can compute a Jacobian of the camera projection function at each correspondence. This also allows the formulation to work for cameras modeled via look-up tables, where the Jacobian can be estimated via interpolation of finite differences.

In Section 2.1 we first recall the derivations for the original Sampson error. Next, in Section 3 we propose our generalization to arbitrary central camera models, and in Section 4 we experimentally show that the proposed error outperforms the competing methods on a variety of camera and lens configurations.

2. Related Work

The *Sampson error* was introduced in [23] to approximate the geometric distance from a point to a conic. The proposed approach was to linearize the equation of the conic around the given point, rendering a closed-form solution for the new distance, now from the point to a line. This idea proved useful for estimating distances to other implicit surfaces, notably those set by constraints of a fundamental matrix [22].

The fundamental matrix is only applicable to perspective cameras, and a lot of subsequent work in computer vision

focused on generalizing it to non-linear camera models, together with some notion of a reprojection error. A well-known example is the *division model* from Fitzgibbon [8]. It maps a point $\boldsymbol{p}$ on the normalized image plane to the corresponding bearing vector $\boldsymbol{d}$ by

$$\boldsymbol{d} = (\boldsymbol{p};\ \ 1 + \lambda \|\boldsymbol{p}\|^2). \tag{8}$$

The original paper [8] already discusses a way of approximating the gold-standard error for the new model: first, undistort the original points of the correspondence; next, apply the closed-form triangulation from [13], and distort the projections back onto the original image. The resulting points are then approximate minimizers of (2). A faster iterative scheme for the optimal error was later proposed in [16], which adapts the idea of [20], originally developed for pinhole cameras, to the division model.

Several camera models with closed-form expressions for epipolar curves were proposed. In [25], epipolar curves are represented by cubics, in the *rational* model [5] — by conics, and in [2] — by circles. With such models, we can use the Sampson approximation to estimate the distance from one point of the correspondence to the epipolar curve of the other, and vice versa. Note that this gives us an approximation of the symmetric epipolar distance, not the Sampson error, because for the latter we would need to optimize both point positions *simultaneously*, and this would move the epipolar curves as well.

All these models were proposed with the goal of searching for the intrinsic parameters together with the epipolar geometry. Indeed, [2] and [5] even provide their own versions of a fundamental matrix on a lifted image space. These matrices can be estimated directly from point correspondences, and then decomposed to provide the distortion parameters and relative pose up to an irreducible ambiguity. We, on the other hand, constrain ourselves to the calibrated case, and thus can afford much greater generality.

There are also various heuristic generalizations of the symmetric epipolar distance to general non-linear camera models that have been used in practice. One approach is to project the bearing vector of one point of the correspondence onto the epipolar plane given by the other. The projected vector is then mapped back onto the image to provide an approximate corrected point, for which the error can be computed. Mathematically, this is expressed as

$$\boldsymbol{n}_1 = \boldsymbol{E}\boldsymbol{d}_1 / \|\boldsymbol{E}\boldsymbol{d}_1\|, \quad \boldsymbol{n}_2 = \boldsymbol{E}^\intercal \boldsymbol{d}_2 / \|\boldsymbol{E}^\intercal \boldsymbol{d}_2\|, \tag{9}$$

$$\begin{aligned} \mathcal{E}^2_{\pi SED} = &\left\|\boldsymbol{p}_1 - \pi\left(\boldsymbol{d}_1 - \boldsymbol{n}_2 \boldsymbol{n}_2^\intercal \boldsymbol{d}_1\right)\right\|^2 \\ &+ \left\|\boldsymbol{p}_2 - \pi\left(\boldsymbol{d}_2 - \boldsymbol{n}_1 \boldsymbol{n}_1^\intercal \boldsymbol{d}_2\right)\right\|^2 \end{aligned} \tag{10}$$

where $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ are the points on the images, $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ are the corresponding bearing vectors, $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ are the normals of the epipolar planes. We have encountered this formulation in private codebases, but we are not aware of its origin. We call it the *projective symmetric epipolar error*. While this does not give us exact distances to the epipolar curves even in the pinhole case, it often provides a reasonable approximation. The main problem with this approach is that it requires calculating $\pi(\boldsymbol{p})$ every time, making it rather computationally heavy for some camera models.

From the probabilistic perspective, the geometric error is only justified as long as the noise in the correspondences stays standard Gaussian and i.i.d. A natural generalization is to allow for arbitrary covariances in point positions. For example, when the keypoints are searched on a multi-resolution image pyramid, the standard deviation of the position can be taken as the downscaling factor of the level on which the point was found. For conics and fundamental matrices, covariance-aware error formulations were derived in [3]. We will show how this approach is generalized to our proposed error formulation in Section 3.2.

In [11] the authors propose another first-order approximation of the reprojection error. The method builds on the optimal triangulation method from [13] and differentates the sixth-order polynomial that has the optimal geometric error as its first root. The first-order approximation is then given by one step of the Newton-Raphson iteration on this polynomial starting from zero. In the paper, this error is only derived for conics. While it would also work for fundamental matrices, this does not generalize to the nonlinear camera model setting we are interested in.

Finally, we would like to note that the Sampson error can also be derived for homographies [4]. Our method is applicable there as well, but with larger computational cost, which somewhat defeats its original purpose.

2.1. The Sampson Approximation

Traditional Sampson error [22, 23] approximates the geometric error, i.e. the distance to the closest pair of points which satisfy the epipolar constraint. Given a correspondence $\boldsymbol{z} = (\boldsymbol{p}_1, \boldsymbol{p}_2) \in \mathbb{R}^2 \times \mathbb{R}^2$ and essential matrix $\boldsymbol{E} \in \mathbb{R}^{3\times 3}$, this can be formalized as

$$\begin{aligned} \mathcal{E}^2_{ML}(\boldsymbol{z}, \boldsymbol{E}) = \min_{\hat{\boldsymbol{z}}} \quad & \|\hat{\boldsymbol{z}} - \boldsymbol{z}\|^2 \\ \text{s.t.} \quad & C(\hat{\boldsymbol{z}}) = 0, \end{aligned} \tag{11}$$

where $\hat{\boldsymbol{z}} = (\hat{\boldsymbol{p}}_1, \hat{\boldsymbol{p}}_2) \in \mathbb{R}^2 \times \mathbb{R}^2$ is the corrected correspondence and the epipolar constraint is

$$C(\hat{\boldsymbol{z}}) = (\hat{\boldsymbol{p}}_2; 1)^\intercal \boldsymbol{E} (\hat{\boldsymbol{p}}_1; 1) = 0. \tag{12}$$

Since (11) does not admit a simple closed form solution, the idea in [23] is to linearize the epipolar constraint, $C(\hat{\boldsymbol{z}}) = 0$, at the original correspondence $\boldsymbol{z}$, i.e.

$$\begin{aligned} \mathcal{E}^2_{S}(\boldsymbol{z}, \boldsymbol{E}) = \min_{\hat{\boldsymbol{z}}} \quad & \|\hat{\boldsymbol{z}} - \boldsymbol{z}\|^2 \\ \text{s.t.} \quad & C(\boldsymbol{z}) + \boldsymbol{J}_C(\hat{\boldsymbol{z}} - \boldsymbol{z}) = 0, \end{aligned} \tag{13}$$

where $\boldsymbol{J}_C = \partial C(\boldsymbol{z})/\partial \boldsymbol{z}$ is the Jacobian of the constraint, evaluated at $\boldsymbol{z}$. Introducing a Lagrangian for (13),

$$\mathcal{L}(\hat{\boldsymbol{z}}, \lambda) = \|\hat{\boldsymbol{z}} - \boldsymbol{z}\|^2 + \lambda \left(C(\boldsymbol{z}) + \boldsymbol{J}_C(\hat{\boldsymbol{z}} - \boldsymbol{z})\right), \quad (14)$$

we get the first-order constraints as

$$\hat{\boldsymbol{z}} = \boldsymbol{z} - \frac{\lambda}{2}\boldsymbol{J}_C^\intercal, \quad C(\boldsymbol{z}) + \boldsymbol{J}_C(\hat{\boldsymbol{z}} - \boldsymbol{z}) = 0. \quad (15)$$

Inserting the first equation into the second yields

$$C(\boldsymbol{z}) - \frac{\lambda}{2}\|\boldsymbol{J}_C\|^2 = 0 \implies \hat{\boldsymbol{z}} = \boldsymbol{z} - \frac{C(\boldsymbol{z})}{\|\boldsymbol{J}_C\|^2}\boldsymbol{J}_C^\intercal. \quad (16)$$

Thus the minimum in (13) is given by

$$\mathcal{E}_S^2 = \left\| \frac{C(\boldsymbol{z})}{\|\boldsymbol{J}_C\|^2}\boldsymbol{J}_C^\intercal \right\|^2 = \frac{C(\boldsymbol{z})^2}{\|\boldsymbol{J}_C\|^2}. \quad (17)$$

Replacing $C(\boldsymbol{z})$ with the epipolar constraint, we arrive at the classical formula for the Sampson error:

$$\mathcal{E}_S^2 = \frac{((\boldsymbol{p}_2;1)^\intercal \boldsymbol{E}(\boldsymbol{p}_1;1))^2}{\|\boldsymbol{E}_{12}(\boldsymbol{p}_1;1)\|^2 + \|(\boldsymbol{E}^\intercal)_{12}(\boldsymbol{p}_2;1)\|^2}. \quad (18)$$

3. Generalizing the Sampson Error

The Sampson error derived in the previous section assumes a pinhole projection model. However, many cameras, in particular those with wide field-of-view, require a more complex model. One popular choice is the Kannala-Brandt [14] model in which a 3D point $\boldsymbol{d} = (X, Y, Z)^\intercal$ in the camera coordinate system is projected as

$$\pi(\boldsymbol{d}) = \frac{\theta_d}{\sqrt{X^2+Y^2}}\begin{pmatrix} f_x X \\ f_y Y \end{pmatrix} + \begin{pmatrix} c_x \\ c_y \end{pmatrix} \quad (19)$$

$$\theta_d = \theta(1 + k_1\theta^2 + k_2\theta^4 + k_3\theta^6 + k_4\theta^8) \quad (20)$$

$$\theta = \text{atan2}(\sqrt{X^2+Y^2}, Z), \quad (21)$$

where $k_1, \ldots, k_4, f_x, f_y, c_x, c_y$ are the intrinsic parameters.

We now consider arbitrary central camera models and derive a generalization of the Sampson error. We assume that we have some forward projection function $\pi : \mathbb{R}^3 \to \Omega$, where $\Omega \subset \mathbb{R}^2$ is the image plane. In the case of the Kannala-Brandt model this is given by (19), but our derivations work for any differentiable model with central projections, i.e. $\pi(\lambda \boldsymbol{d}) = \pi(\boldsymbol{d}) \quad \forall \lambda > 0$. Similarly we define the (normalized) unprojection, $\pi^{-1} : \Omega \to S^2$ as mapping image points onto the corresponding unit-vectors in 3D. The geometric error (11) then becomes $\mathcal{E}_{ML}(\boldsymbol{z}, \boldsymbol{E}) =$

$$\min_{\hat{\boldsymbol{z}}} \quad \|\hat{\boldsymbol{z}} - \boldsymbol{z}\|^2 \quad (22)$$

$$\text{s.t.} \quad C(\hat{\boldsymbol{z}}) = \pi^{-1}(\hat{\boldsymbol{p}}_2)^\intercal \boldsymbol{E} \pi^{-1}(\hat{\boldsymbol{p}}_1) = 0. \quad (23)$$

In practice, most camera models only define the forward projection π, and unprojection π^{-1} is solved for iteratively. For example, (19) does not have a closed-form inverse. This makes it non-trivial to directly compute its Jacobians required for the Sampson approximation. To avoid this computation we rely on the inverse function theorem. The variant that we will use can be found in the book [6]. It is tailored to functions that map higher-dimensional inputs to lower-dimensional outputs. We rephrase it here for reader's convenience, adapting the notation to our needs and losing some generality:

Theorem 1. *[6, Theorem 1F.6] Let $\pi : \mathbb{R}^3 \to \mathbb{R}^2$ be continuously differentiable in a neighborhood around $\boldsymbol{d} \in \mathbb{R}^3$. Further, suppose that its Jacobian $\boldsymbol{J}$ at $\boldsymbol{d}$ has full rank. Then, there exists a neighborhood V of $\bar{\boldsymbol{p}} = \pi(\boldsymbol{d})$ and a continuously differentiable function $\widetilde{\pi^{-1}} : V \to \mathbb{R}^3$ such that $\pi(\widetilde{\pi^{-1}}(\boldsymbol{p})) = \boldsymbol{p} \quad \forall \boldsymbol{p} \in V$ and whose Jacobian is given by $\boldsymbol{J}^\dagger$, the Moore-Penrose pseudoinverse of $\boldsymbol{J}$.*

The construction of $\widetilde{\pi^{-1}}$ in the proof from [6] also gives us that $\widetilde{\pi^{-1}}(\bar{\boldsymbol{p}}) = \boldsymbol{d}$. Note that π admits many right-inverses, but they only differ in lengths assigned to unprojected bearing vectors. Thus, $\widetilde{\pi^{-1}}$ that the theorem gives us need not coincide with π^{-1} that maps points to *unit* vectors. However, the following result shows that when we evaluate $\boldsymbol{J}$ at a unit vector $\boldsymbol{d} = \pi^{-1}(\bar{\boldsymbol{p}})$, $\boldsymbol{J}^\dagger$ actually gives us the Jacobian of π^{-1}:

Corollary 1. *Let $\boldsymbol{d} \in S^2$ and $\boldsymbol{J} = \frac{\partial \pi}{\partial \boldsymbol{d}} \in \mathbb{R}^{2\times 3}$. Then the Jacobian of π^{-1} at $\bar{\boldsymbol{p}} = \pi(\boldsymbol{d})$ is given by $\boldsymbol{J}^\dagger$.*

Proof. Since the only difference between $\widetilde{\pi^{-1}}$ from theorem 1 and π^{-1} is the norm of the output, we can write

$$\pi^{-1}(\boldsymbol{p}) = h(\widetilde{\pi^{-1}}(\boldsymbol{p})) \quad \forall \boldsymbol{p} \in V, \quad (24)$$

where $h : \boldsymbol{v} \mapsto \boldsymbol{v}/\|\boldsymbol{v}\|$ is the normalization function. We also have $\boldsymbol{d} = \widetilde{\pi^{-1}}(\bar{\boldsymbol{p}})$. Thus, the Jacobian of π^{-1} at $\bar{\boldsymbol{p}}$ is given by

$$\frac{\partial \pi^{-1}}{\partial \boldsymbol{p}} = \frac{\partial h}{\partial \boldsymbol{v}}\bigg|_{\boldsymbol{v}=\boldsymbol{d}} \frac{\partial \widetilde{\pi^{-1}}}{\partial \boldsymbol{p}} = (\boldsymbol{I} - \boldsymbol{d}\boldsymbol{d}^\intercal)\boldsymbol{J}^\dagger = \boldsymbol{J}^\dagger. \quad (25)$$

□

In the last equation we used $\boldsymbol{d}^\intercal \boldsymbol{J}^\dagger = \boldsymbol{0}$, which is true in general for any null-vector $\boldsymbol{d}$ of the matrix $\boldsymbol{J}$, but in our case can also be seen from the following lemma that gives a closed-form expression for $\boldsymbol{J}^\dagger$:

Lemma 1. *Let $\boldsymbol{d} \in \mathbb{R}^3$ and $\boldsymbol{J} = \frac{\partial \pi}{\partial \boldsymbol{d}} \in \mathbb{R}^{2\times 3}$. Then, $\boldsymbol{J}^\dagger$ is given by*

$$\boldsymbol{M} = \frac{1}{\boldsymbol{d} \cdot (\boldsymbol{g}_x \times \boldsymbol{g}_y)} \left[(\boldsymbol{g}_y \times \boldsymbol{d}), \quad (\boldsymbol{d} \times \boldsymbol{g}_x)\right], \quad (26)$$

where $\boldsymbol{g}_x, \boldsymbol{g}_y \in \mathbb{R}^3$ are the two rows of the Jacobian $\boldsymbol{J}$.

Proof. Simple calculations show that $\boldsymbol{J}\boldsymbol{M} = \boldsymbol{I}_{2\times 2}$. It remains to show that $\boldsymbol{M}\boldsymbol{J} \in \mathbb{R}^{3\times 3}$ is symmetric. For ease of notation introduce $\alpha = \boldsymbol{d} \cdot (\boldsymbol{g}_x \times \boldsymbol{g}_y)$. Then we have

$$\alpha \boldsymbol{M}\boldsymbol{J} = (\boldsymbol{g}_y \times \boldsymbol{d})\boldsymbol{g}_x^\intercal + (\boldsymbol{d} \times \boldsymbol{g}_x)\boldsymbol{g}_y^\intercal = \tag{27}$$

$$= [\boldsymbol{d}]_\times \left(-\boldsymbol{g}_y\boldsymbol{g}_x^\intercal + \boldsymbol{g}_x\boldsymbol{g}_y^\intercal\right) = [\boldsymbol{d}]_\times [\boldsymbol{g}_y \times \boldsymbol{g}_x]_\times. \tag{28}$$

Since the projection function π does not depend on the norm of $\boldsymbol{d}$, i.e. $\pi(\boldsymbol{d}) = \pi(\lambda \boldsymbol{d}) \;\; \forall \lambda > 0$, we have that $\boldsymbol{J}\boldsymbol{d} = \boldsymbol{0}$. Since $\boldsymbol{d}$ is then orthogonal to both $\boldsymbol{g}_x$ and $\boldsymbol{g}_y$ we have $\boldsymbol{g}_y \times \boldsymbol{g}_x = \gamma \boldsymbol{d}$ for some $\gamma \in \mathbb{R}$. Inserting into (28) we get

$$\boldsymbol{M}\boldsymbol{J} = \frac{\gamma}{\alpha}[\boldsymbol{d}]_\times^2, \tag{29}$$

which is a symmetric matrix. □

The expression (26) can be used to efficiently compute the Jacobian $\boldsymbol{J}_C$ of the constraint function C from (23), and finally to arrive at the Sampson approximation of the reprojection error in any differentiable camera model:

$$\mathcal{E}_{TS}^2 = \frac{C(\boldsymbol{z})^2}{\|\boldsymbol{J}_C\|^2} = \frac{(\boldsymbol{d}_2^\intercal \boldsymbol{E} \boldsymbol{d}_1)^2}{\|\boldsymbol{d}_2^\intercal \boldsymbol{E} \boldsymbol{J}_1^\dagger\|^2 + \|\boldsymbol{d}_1^\intercal \boldsymbol{E}^\intercal \boldsymbol{J}_2^\dagger\|^2}, \tag{30}$$

where $\boldsymbol{d}_i = \pi^{-1}(\boldsymbol{p}_i)$ and the two Jacobians are evaluated at $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ respectively. In the rest of the paper we refer to this as the *Tangent Sampson* error. The name reflects the fact that after we linearize the unprojection π^{-1}, we are essentially optimizing in the tangent plane to S^2 at the bearing vectors of the undistorted points. Note that the formula is analogous to the original Sampson error (18) (replacing $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ with the image points, and replacing $\boldsymbol{J}_1^\dagger$ and $\boldsymbol{J}_2^\dagger$ with $[1\,0; 0\,1; 0\,0]$). Further, since $\boldsymbol{J}_1$ and $\boldsymbol{J}_2$ are evaluated at the original points of the correspondence, they are independent of $\boldsymbol{E}$ and can be precomputed. Thus, during RANSAC and refinement the computational cost stays approximately the same as for the classical Sampson error.

3.1. Backward Camera Models

While most camera models used in practice are forward-models, i.e. mapping from the world to the image, there also exist backward-models which instead take the image points as input. The division model (8) is perhaps the most popular example here. The derivations above can easily be adapted to allow for backward-models. The pseudoinverses $\boldsymbol{J}_k^\dagger$ are simply replaced with the 3×2 Jacobians from the undistortion map,

$$\mathcal{E}_{TS}^2 = \frac{(\boldsymbol{d}_2^\intercal \boldsymbol{E} \boldsymbol{d}_1)^2}{\|\boldsymbol{d}_2^\intercal \boldsymbol{E} \boldsymbol{J}_1\|^2 + \|\boldsymbol{d}_1^\intercal \boldsymbol{E}^\intercal \boldsymbol{J}_2\|^2}. \tag{31}$$

In particular, for the division model (8) we would have

$$\boldsymbol{J} = \frac{\partial h}{\partial \boldsymbol{d}} \frac{\partial \boldsymbol{d}}{\partial \boldsymbol{x}} = \frac{\partial h}{\partial \boldsymbol{d}} \left[I_{2\times 2};\; 2\lambda \boldsymbol{x}^\intercal\right] \in \mathbb{R}^{3\times 2}, \tag{32}$$

where we need to explicitly include the Jacobian $\frac{\partial h}{\partial \boldsymbol{d}}$ of the normalization function h, since the unprojected $\boldsymbol{d}$ from (8) is not a unit vector.

3.2. Measurement Uncertainty

Following the approach from [3], we can incorporate measurement uncertainty in the keypoint positions into the Tangent Sampson error. We will assume that the keypoint positions are affected by independent Gaussian noise

$$\boldsymbol{p}_1 \sim \mathcal{N}(\bar{\boldsymbol{p}}_1, \boldsymbol{\Sigma}_1), \quad \boldsymbol{p}_2 \sim \mathcal{N}(\bar{\boldsymbol{p}}_2, \boldsymbol{\Sigma}_2) \tag{33}$$

with known covariances. For observations $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$, the maximum-likelihood error for an essential matrix $\boldsymbol{E}$ is

$$\mathcal{E}_{ML}^2(\boldsymbol{z}, \boldsymbol{E}) = \min_{\hat{\boldsymbol{z}}} \quad \frac{1}{2}(\boldsymbol{z} - \hat{\boldsymbol{z}})^\intercal \boldsymbol{\Sigma}^{-1}(\boldsymbol{z} - \hat{\boldsymbol{z}}) \tag{34}$$

$$\text{s.t.} \quad C(\hat{\boldsymbol{z}}) = 0,$$

where C is the constraint function (23) and $\boldsymbol{\Sigma}$ is the joint covariance of $\boldsymbol{z} = [\boldsymbol{p}_1; \boldsymbol{p}_2]$, i.e. $\boldsymbol{\Sigma} = diag(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2)$.

The procedure from Section 2.1 yields in this case

$$\mathcal{E}_{TS}^2 = \frac{C(\boldsymbol{z})^2}{\boldsymbol{J}_C \boldsymbol{\Sigma} \boldsymbol{J}_C^\intercal} = \frac{(\boldsymbol{d}_2^\intercal \boldsymbol{E} \boldsymbol{d}_1)^2}{\|\boldsymbol{d}_2^\intercal \boldsymbol{E} \boldsymbol{J}_1^\dagger\|_{\Sigma_1}^2 + \|\boldsymbol{d}_1^\intercal \boldsymbol{E}^\intercal \boldsymbol{J}_2^\dagger\|_{\Sigma_2}^2}, \tag{35}$$

where $\|\boldsymbol{v}\|_\Sigma^2 = \boldsymbol{v}^T \Sigma \boldsymbol{v}$. See the supplementary material for details on the derivation. This is a generalization of the last expression in [3, Section 3.4] to the arbitrary camera model case. We evaluate how it compares to the original Tangent Sampson error (30) in a synthetic setting in Section 4.4.1.

4. Experimental Evaluation

In the rest of the paper we experimentally evaluate the Tangent Sampson error. We compare against a selection of popular error metrics, summarized in Table 1.

4.1. Implementation Details

We implemented all error formulations in C++ using Eigen [9] as a linear algebra library and Ceres [1] for optimization. Since evaluation is run on Intel hardware, we additionally speed up the errors that do not require iterative methods or explicit camera projection ($\mathcal{E}_{TS}$, $\mathcal{E}_S$, $\mathcal{E}_{SED}$, $\mathcal{E}_{CS}$, and $\mathcal{E}_{ALG}$) by vectorizing their implementations with AVX and FMA instruction sets. In Table 1 we also list the approximate computational cost of evaluating the residual. Tangent Sampson error is significantly faster compared to other metrics which also measure the error in the original image, and has comparable, but slightly higher, runtime to the purely geometric and pinhole-based metrics. For the computational cost we include parts which cannot be precomputed (i.e. parts that depend on the essential matrix). For a comparison with the non-optimized errors (using pure Eigen-implementations), we refer the reader to the supplementary material, where we list the runtime for all formulations without explicit vectorization and some setup details.

		Name		Runtime Rel.	Runtime Abs.
$\mathcal{E}_{ALG}$	(7)	Algebraic error		1.0	1.3 ns
$\mathcal{E}_{CS}$	(6)	Cosine error		1.8	2.4 ns
$\mathcal{E}_{ML}$	(3)	(Pinhole) Reproj. error		28.0	36.7 ns
$\mathcal{E}_{SED}$	(4)	(Pinhole) Sym. Epi. Dist.		1.7	2.3 ns
$\mathcal{E}_{S}$	(5)	(Pinhole) Sampson error		1.4	1.9 ns
$\mathcal{E}_{\pi ML}$	(22)	Reproj. error		886	1160 ns
$\mathcal{E}_{\pi SED}$	(10)	Proj. Sym. Epi. Dist.		92	121 ns
$\mathcal{E}_{TS}$	(30)	Tangent Sampson		2.2	2.8 ns

Table 1. Error metrics compared in the experimental evaluation. The errors are grouped by the domain where the error is computed: geometric errors (), the undistorted image () and the original image (),. Computational cost is shown relative to the algebraic error which is the cheapest.

4.2. Datasets

For the evaluation we consider two datasets. First the meta-calibration dataset from BabelCalib [21]. The dataset is a collection of checkerboard images from a wide variety of cameras, including some with extremely wide field-of-view ($> 270°$). For the evaluation we the split the dataset into three groups depending on their field-of-view; $< 90°$, $90 - 150°$, and $> 150°$. As the original dataset is biased towards higher field-of-view cameras, we also include some additional checkerboard images taken with a standard DSLR (circa 50 - 70°) to even out the distribution. These checkerboard datasets allow us to evaluate on a large number of different camera and lens setups. In Section 4.3 we show that the Tangent Sampson error consistently approximates the true reprojection error, regardless of the camera used, and in Section 4.4 we show that when used for local refinement it yields the most accurate epipolar geometries.

To evaluate the performance of the error metrics in the context of RANSAC we consider image pairs extracted from Structure-from-Motion datasets, where the poses from the reconstruction are taken as a psuedo-ground truth. As the experiments in Section 4.3-4.4 show that most error metrics behave similarly for low to medium field-of-view cameras ($< 90°$), we focus our evaluation on cameras with larger field-of-view. For this we use two datasets (Grossmunster and Kirchenge) from Larsson et al. [17], which contain 373 and 369 images taken with a fisheye-lens.

4.3. Comparison with True Reprojection Error

We first evaluate how different error metrics compare with the true maximum-likelihood estimate, $\mathcal{E}_{ML}$ (3) for pinhole cameras and $\mathcal{E}_{\pi ML}$ (22) for non-pinhole cameras. To compute the ML-estimate we perform two-view triangulation followed by non-linear optimization of the 3D point, minimizing the squared reprojection error (using the known camera intrinsics). Note that this approach is orders of magnitude slower compared to the metrics we evaluate (see Table 1) and cannot reasonably be used in practical scenarios.

For the experiment we consider the collection of checkerboard images from [21]. We use the calibration toolbox from BabelCalib [21] to estimate the extrinsic and intrinsic parameters for each image. For each pair we collect the 2D-2D correspondences from corresponding corners. As the corner detector [10] is extremely accurate on these images (some having around 0.2 px error during calibration), we add Gaussian noise with 1 px standard deviation, to better differentiate the methods (as all methods give zero error for exact data).

Since it is not meaningful to directly compare the residual values, we instead compare their ordering. For this we make use of *Kendall's τ rank correlation coefficient* [15]. It is defined for two rankings $(p_1, \dots, p_n)$ and $(q_1, \dots, q_n)$ as

$$\tau(p, q) = \frac{|\{\text{concordant pairs}\}| - |\{\text{discordant pairs}\}|}{\binom{n}{2}}. \quad (36)$$

A pair of indices $(i, j) \in \{1, \dots, n\}^2$ where $i > j$ is called *concordant* if and only if

$$(p_i > p_j \wedge q_i > q_j) \vee (p_i < p_j \wedge q_i < q_j). \quad (37)$$

Otherwise, the pair is *discordant*. Note that due to the normalization we have $\tau(p, q) \in [-1, 1]$, where identical rankings yield $\tau = 1$, and reversed $\tau = -1$.

The results are shown in Fig. 3, split into three groups based on the camera field-of-view. For low to medium field-of-view cameras ($< 90°$) we can see that all metrics except the algebraic residual behave similar to the true reprojection error. As the field-of-view increases, there is a more clear difference between the methods, with the pinhole-based errors performing the worst. Note however, that for all cameras the proposed Tangent Sampson error offers the best approximation of the true reprojection error.

In addition to Kendall-Tau distance, we compare the errors qualitatively. For this, we show the values of an error for a fixed correspondence on a pair of images from a panoramic and fisheye camera in Fig. 4. There, we keep the point on the fisheye image fixed, and vary the point on the panoramic image to show different error values coded by different colors. One can see that $\mathcal{E}_{TS}$ error models $\mathcal{E}_{\pi ML}$ closely in the region where the latter is well-defined. $\mathcal{E}_{TS}$ is also one of the two errors that do not degenerate around the epipole, and, at the same time, it correctly models the spread of values of $\mathcal{E}_{\pi ML}$ around the epipolar curve. Note that the shape of true reprojection error varies significantly depending on how close the correspondence is to the epipole. In the supplementary material we show more qualitative results. In the supplementary material we also show a comparison with [16], which proposed an optimal method for computing the πML error (22) for the division model (8).

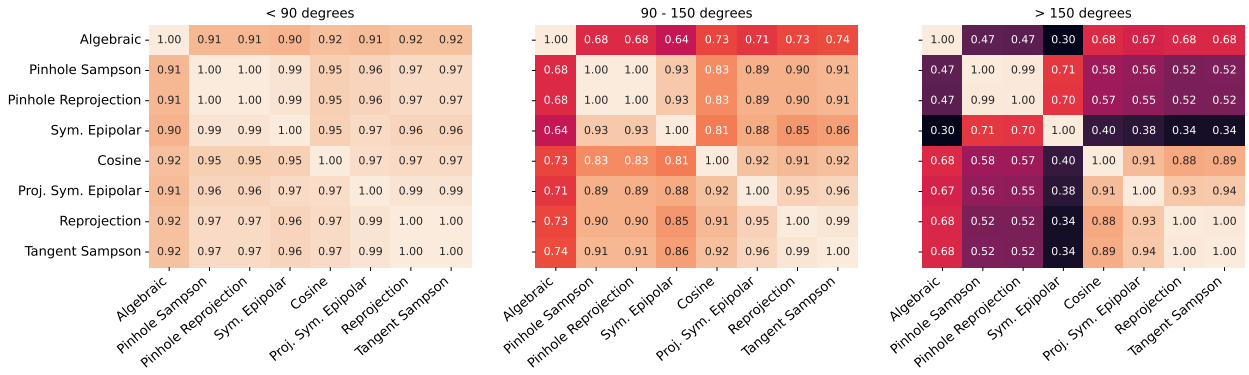

Figure 3. The figure shows the Kendall τ coefficient comparing different error metrics for the checkerboard data. For all cameras the proposed Tangent Sampson error correlates well with the true reprojection error πML.

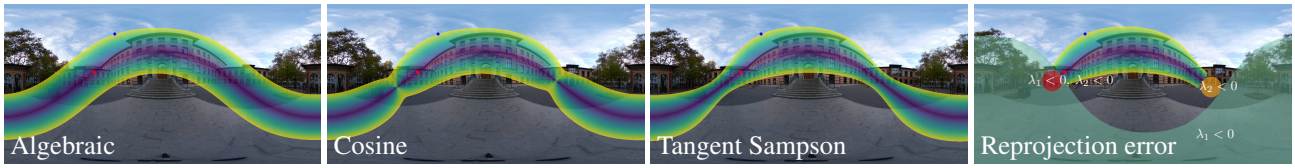

Figure 4. **Qualitative comparison of different error functions.** The thresholds for colormaps were selected to be the error at the point marked by the blue plus. The shape of the true reprojection error varies heavily depending on how close the correspondence is to the epipoles. In the supplementary material we show more examples. We also show $\mathcal{E}_{\pi SED}$ and errors computed on the undistorted image in the supplementary. The geometric error is only shown in the region where the 3D point after midpoint triangulation does not correspond to negative depths of bearing vectors (areas where this does not hold for one or both depths are highlighted).

4.4. Accuracy in Pose Refinement

Next we evaluate the error metrics on how they perform when used for pose-refinement. For the experiment we again consider the image pairs extracted from the checkerboard images. For each image pair we perform local optimization minimizing the proposed error metrics over all correspondences. To initialize the optimization we take the ground truth relative pose and randomly perturb both the rotation and translation by 1 degree. Additionally, point positions were perturbed with Gaussian noise with unit standard deviation along each dimension. For the two ML metrics we triangulate 3D points which are optimized jointly with the camera pose. We experience that this sometimes has problems with worse convergence. The results are shown in Table 2. Again we see that for low to medium field-of-view cameras, the difference between the error metrics is small. The Tangent Sampson error consistently provides the most accurate camera poses after refinement.

Convergence Properties. The previous experiment evaluated the accuracy of the camera poses obtained when initialized close to the correct pose. We also experimented with varying the initial pose and measuring how often the optimization converged to the correct pose. Figure 5 shows the percentage of poses (across all pairs) that are within 5 degrees average rotation/translation error for varying errors in the initial pose. For large initial errors ($> 25°$), the proposed Tangent-Sampson error performs slightly better compared to the other metrics.

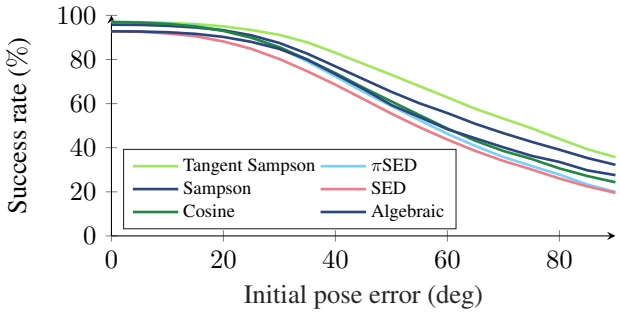

Figure 5. **Convergence properties.** Successful trials (average rot. and trans. error $< 5°$) for varying initial pose errors (in degrees).

4.4.1 Effect of Measurement Uncertainty

Experiments in [3] already show that taking point covariances into account in the cost function provides benefits for fundamental matrix estimation. Here, we conduct another synthetic test to show that this effect persists when estimating the essential matrix with highly nonlinear cameras. We generate 2000 instances with pairs of spherical cameras (1000×2000px) with a random relative pose. For each pair we project 100 random 3D-points uniformly drawn from the cube $[-10, 10]^3$, and add Gaussian noise with covariance $\mathbf{\Lambda}$.

Covariances are generated as $\mathbf{\Lambda} = \alpha \boldsymbol{Q} \text{diag}(\beta, 1-\beta) \boldsymbol{Q}^\intercal$,

	$< 90°$ (5332 pairs)		$90 - 150°$ (21927 pairs)		$> 150°$ (11860 pairs)	
	Rotation	Translation	Rotation	Translation	Rotation	Translation
$\mathcal{E}_{ALG}$	0.379 / 0.314	1.137 / 0.328	1.528 / 0.378	3.938 / 0.588	1.582 / 0.506	3.489 / 0.655
$\mathcal{E}_{S}$	0.376 / **0.312**	**1.056** / 0.327	1.464 / 0.347	3.520 / 0.575	1.634 / 0.567	3.812 / 0.778
$\mathcal{E}_{SED}$	0.376 / **0.312**	1.065 / 0.327	1.475 / 0.352	3.554 / 0.578	2.596 / 0.986	5.667 / 1.342
$\mathcal{E}_{COS}$	0.376 / 0.313	1.066 / **0.326**	1.454 / 0.341	3.459 / 0.558	1.447 / 0.454	3.235 / 0.612
$\mathcal{E}_{ML}$	0.376 / **0.312**	1.103 / 0.327	1.472 / 0.358	3.733 / 0.575	1.630 / 0.563	3.771 / 0.766
$\mathcal{E}_{\pi ML}$	**0.375** / **0.312**	1.075 / 0.327	1.451 / 0.350	3.778 / **0.557**	**1.418** / 0.446	**3.143** / **0.597**
$\mathcal{E}_{\pi SED}$	0.376 / 0.313	1.074 / 0.327	1.458 / 0.343	3.490 / 0.559	1.473 / 0.450	3.249 / 0.607
$\mathcal{E}_{TS}$	0.376 / **0.312**	1.057 / 0.327	**1.443** / **0.340**	**3.432** / **0.557**	1.465 / **0.445**	3.226 / 0.598

Table 2. Evaluation of camera pose refinement. Table shows the mean and median errors (in degrees) of the rotation and translation after refinement. The data is split into three groups based on the camera field-of-view. Best error is highlighted in bold. Second-best is underlined. For low-distortion cameras the choice of error makes almost no difference on the quality of pose.

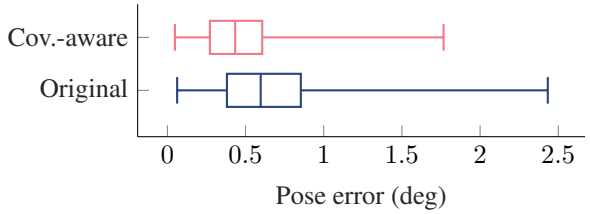

Figure 6. **Measurement Uncertainty.** The distribution of pose errors (maximum of rotation and translation error) when using covariance-aware Tangent Sampson error for pose refinement versus non-covariance-aware in synthetic experiments. Boxes show quartiles of the respective distribution of errors and whiskers extend to the minimal and maximal value.

where $\boldsymbol{Q} \sim U(\mathrm{SO}(2))$ is a random rotation on the plane, $\beta \sim U([0, 0.5])$ is a random eccentricity parameter, and α is the overall scale of the noise. To emulate keypoint detection on a multi-scale image pyramid, all generated points were split into 4 equal sets, with α within each set assigned to 1, 2, 4, and 8, respectively. Points and covariances are then passed to a Levenberg-Marquardt optimization of the relative pose using (35) as the residual function, with true pose serving as the initial estimate. Figure 6 demonstrates how the covariance-aware formulation (35) compares to the baseline (30). Covariance information is useful in the case when points are affected by noise of unequal scales.

4.5. Evaluation in Robust Estimators

Finally we evaluate the methods in the context of RANSAC [7]. We implement a standard LO-RANSAC [18] estimator which uses the various error metrics both for model scoring (using MSAC scoring [24]) and refinement (using non-linear optimization on the inlier-set). For the experiment we consider image pairs from two Structure-from-Motion reconstructions (see Section 4.2). From each dataset we extract all image pairs which contain between 50 and 200 correct matches (according to the Structure-from-Motion ground-truth). For each pair we estimate the relative pose using RANSAC and compute the error in the pose, as well as percentage of successful runs. For each method we tuned the inlier threshold on a separately extracted validation set. For more details on the experiment setup see the supplementary material. The results are presented in Table 3. Using Tangent Sampson error consistently gives slightly better performance compared to other error metrics.

	Grossmunster			Kirchenge		
	$\epsilon_{R,t}$	$< 5°$	RT	$\epsilon_{R,t}$	$< 5°$	RT
$\mathcal{E}_{ALG}$	0.761	75.0	**39.6**	0.318	84.8	**21.6**
$\mathcal{E}_{COS}$	0.561	81.5	50.4	0.254	91.6	25.5
$\mathcal{E}_{SED}$	1.161	68.9	82.0	0.355	86.4	44.0
$\mathcal{E}_{S}$	0.924	74.5	51.0	0.300	89.2	25.3
$\mathcal{E}_{\pi SED}$	0.561	81.3	211.2	0.255	91.6	87.4
$\mathcal{E}_{TS}$	**0.550**	**82.2**	46.4	**0.248**	**91.9**	24.2

Table 3. Evaluation in robust estimators. Table shows the median pose error (maximum of rotation and translation error), percentage of pairs for which the estimation was successful (pose error less than 5 degrees), and runtime (in ms). Best value in the column is highlighted in bold. Second-best is underlined.

5. Conclusions

In this paper we have presented a simple generalization to the classical Sampson error which allows for using arbitrary camera models. In experiments on a wide variety of camera and lens configurations, we have shown that the proposed error metric outperforms the alternative metrics that are used in practice. While the improvements are sometimes minor (e.g. for low-distortion pinhole cameras), they are consistent and come at essentially no additional cost.

Acknowledgments: This project received funding from the strategic research program ELLIIT. MT was supported by the ESOP scholarship from the ETH Foundation.

References

[1] Sameer Agarwal, Keir Mierle, and Others. Ceres solver. http://ceres-solver.org. 5

[2] João Pedro Barreto and Kostas Daniilidis. Fundamental matrix for cameras with radial distortion. In *International Conference on Computer Vision (ICCV)*, 2005. 3

[3] Wojciech Chojnacki, Michael J. Brooks, Anton Van Den Hengel, and Darren Gawley. On the fitting of surfaces to data with covariances. *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, 2000. 3, 5, 7

[4] Ondřej Chum, Tomáš Pajdla, and Peter Sturm. The geometric error for homographies. *Computer Vision and Image Understanding (CVIU)*, 2005. 3

[5] David Claus and Andrew W Fitzgibbon. A rational function lens distortion model for general cameras. In *Computer Vision and Pattern Recognition (CVPR)*, 2005. 3

[6] Asen L Dontchev and R Tyrrell Rockafellar. *Implicit functions and solution mappings*, volume 543. Springer, 2009. 4

[7] Martin A Fischler and Robert C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981. 8

[8] Andrew W Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In *Computer Vision and Pattern Recognition (CVPR)*, 2001. 3

[9] Gaël Guennebaud, Benoît Jacob, et al. Eigen v3. http://eigen.tuxfamily.org, 2010. 5

[10] Hyowon Ha, Michal Perdoch, Hatem Alismail, In So Kweon, and Yaser Sheikh. Deltille grids for geometric camera calibration. In *International Conference on Computer Vision (ICCV)*, 2017. 6

[11] Matthew Harker and Paul O'Leary. First order geometric distance (the myth of sampsonus). In *British Machine Vision Conference (BMVC)*, 2006. 3

[12] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*. Cambridge university press, 2003. 1

[13] Richard I Hartley and Peter Sturm. Triangulation. *Computer vision and image understanding*, 68(2):146–157, 1997. 2, 3

[14] Juho Kannala and Sami S Brandt. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. 2006. 4

[15] Maurice G Kendall. A new measure of rank correlation. *Biometrika*, 30(1/2):81–93, 1938. 6

[16] Zuzana Kukelova and Viktor Larsson. Radial distortion triangulation. In *Computer Vision and Pattern Recognition (CVPR)*, 2019. 3, 6

[17] Viktor Larsson, Nicolas Zobernig, Kasim Taskin, and Marc Pollefeys. Calibration-free structure-from-motion with calibrated radial trifocal tensors. In *European Conference on Computer Vision (ECCV)*, 2020. 6

[18] Karel Lebeda, Jirı Matas, and Ondrej Chum. Fixing the locally optimized ransac–full experimental evaluation. In *British Machine Vision Conference (BMVC)*, 2012. 8

[19] Seong Hun Lee and Javier Civera. Geometric interpretations of the normalized epipolar error. *arXiv preprint arXiv:2008.01254*, 2020. 2

[20] Peter Lindstrom. Triangulation made easy. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1554–1561. IEEE, 2010. 2, 3

[21] Yaroslava Lochman, Kostiantyn Liepieshov, Jianhui Chen, Michal Perdoch, Christopher Zach, and James Pritts. Babelcalib: A universal approach to calibrating central cameras. In *International Conference on Computer Vision (ICCV)*, 2021. 6

[22] Quan-Tuan Luong and Olivier D Faugeras. The fundamental matrix: Theory, algorithms, and stability analysis. *International Journal of Computer Vision (IJCV)*, 1996. 2, 3

[23] Paul D Sampson. Fitting conic sections to "very scattered" data: An iterative refinement of the bookstein algorithm. *Computer graphics and image processing*, 1982. 2, 3

[24] Philip HS Torr and Andrew Zisserman. Mlesac: A new robust estimator with application to estimating image geometry. *Computer Vision and Image Understanding (CVIU)*, 2000. 8

[25] Zhengyou Zhang. On the epipolar geometry between two images with lens distortion. In *International Conference on Pattern Recognition (ICPR)*, 1996. 3