Perturbation Analysis of the 8-Point Algorithm: a Case Study for Wide FoV Cameras

Thiago L. T. da Silveira and Claudio R. Jung
Institute of Informatics, Federal University of Rio Grande do Sul, Brazil
{tltsilveira,crjung}@inf.ufrgs.br

Abstract

This paper presents a perturbation analysis for the estimate of epipolar matrices using the 8-Point Algorithm (8-PA). Our approach explores existing bounds for singular subspaces and relates them to the 8-PA, without assuming any kind of error distribution for the matched features. In particular, if we use unit vectors as homogeneous image coordinates, we show that having a wide spatial distribution of matched features in both views tends to generate lower error bounds for the epipolar matrix error. Our experimental validation indicates that the bounds and the effective errors tend to decrease as the camera Field of View (FoV) increases, and that using the 8-PA for spherical images (that present $360°\times180°$ *FoV) leads to accurate essential matrices. As an additional contribution, we present bounds for the direction of the translation vector extracted from the essential matrix based on singular subspace analysis.*

1. Introduction

Obtaining 3D scene structure and camera poses based on two or more views of the same scene have been widely studied by the computer vision community, and the epipolar geometry encoded by either the *fundamental* or the *essential* matrix is a keystone in the context of Structure from Motion [28]. For simplicity, we will use the name *epipolar matrix* to denote either one of these two matrices, as in [1].

There are several approaches for estimating epipolar matrices [18, 21, 25], but the 8-Point Algorithm (8-PA) proposed by Longuet-Higgins [20] is still very popular, serving as an initial estimate or as the basis for further developments such as feature normalization schemes [15, 24], the imposition of rank constraints [49], the inclusion of additional matrix factorizations [44], the use of robust loss functions to better handle outliers [16], or the inclusion of weights and deep learning strategies for simultaneous keypoint matching and motion recovery [46]. A class of approaches focuses on estimating optimal inlier sets [45], which is particularly important when the number of matched features is scarce and outliers are present. When using wider Field of View (FoV) cameras, however, the spatial region captured by both cameras present more overlap on the image domain, typically leading to more matched features. In particular, spherical images present full $360°\times180°$ FoV and are becoming increasingly popular in the context of multiview 3D reconstruction and/or pose estimation [12, 29, 30, 31, 32].

As noted in [46], the typical pipeline for estimating epipolar matrices consists of finding correspondence points across images, and then applying an outlier removal strategy (e.g. RANSAC [9] or its variants [33]) to retrieve a subset of *potential* inliers. This pruned set of correspondences is then fed to the estimation method itself, such as the 8-PA, and errors in the matching step can degrade the estimated epipolar matrix. They are caused by noisy feature matching (features are correctly matched but not at the exact location) or wrong feature matching, and determining the impact of these errors on the estimated matrix is a relevant research topic. Existing error analysis approaches [5, 36, 42] focus on noisy features, assuming that badly matched features can be effectively removed by the outlier rejection method. However, outliers are typically detected based on some kind of distance to the epipolar line (or curve when using spherical images), such as algebraic or Sampson distances [39]. Hence, wrong matches along (or close to) the epipolar line/curve might be erroneously considered inliers.

This paper presents a perturbation analysis for the 8-PA that does not assume any kind of matching error distribution. We represent the matched features in homogeneous coordinates as 3D unit vectors, which relate to the viewing directions of the rays that connect the 3D world points and the camera centers in the calibrated case. This is a particularly natural choice when using spherical cameras [12, 17], but can also be explored when the pinhole model is assumed. We present error bounds for epipolar matrix estimation as a function of matching errors and shed some light on the impact of the Field of View (FoV) of the cameras on error propagation. In particular, we show that the 8-PA potentially provides good estimates for the essential matrix E

when using spherical images, which present a 360°×180° FoV. Additionally, we present bounds for the translation vector (up to a scale) extracted from the essential matrix.

2. Related work

Estimating the uncertainty when computing epipolar matrices has been studied by several authors [5, 15, 22, 24, 36, 42], focusing on the relationship between correspondence errors and the epipolar matrices themselves or directly on the errors of estimated 3D structure and/or pose.

Weng *et al.* [42] presented an error analysis of the essential matrix estimation based on first order perturbations, relating the covariance matrix of the matching errors with the covariance of the essential matrix. This assumption implies that the variance of the feature detector error must be determined, which is very dependent on the detector itself and the scene, and very sensitive to outliers. Furthermore, they assumed the un-normalized 8-PA, which was later shown by Hartley that leads to numerical instabilities [15].

Mühlich and Mester [24] related the error produced by the 8-PA with the perturbation of eigenvalues and singular values. They assumed that the covariance matrix of the correspondence matching errors is known to obtain a bound for the essential matrix error, and use this bound to propose a new feature normalization scheme. Mair and colleagues [22] extended the analysis in [42] by including the normalization schemes presented in [15] and [24]. Notice that both [24] and [22] present the drawback of assuming known feature matching variances.

Csurka *et al.* [5] presented an error analysis of the fundamental matrix F obtained by a non-linear method. They model F as a random vector, such that the mean of the distribution is the actual matrix and the covariance encodes the uncertainty errors. They assume that outliers were rejected in a previous step so that the analysis is focused only on noisy correspondences. Sur and colleagues [36] follow a similar path but focus on the errors of epipolar matrix estimation using the 8-PA. However, as in [5], they present the uncertainty as a covariance matrix and discard the presence of outliers, which limits the application of their method. Also, they consider the un-normalized version of the 8-PA, as [42].

Hartley [15] evaluated the condition number of the measurement matrix used in the 8-PA, suggesting a normalized version that is numerically more stable. The core of his analysis was that using "raw" homogeneous coordinates (just appending the value 1 to the pixel coordinates) leads to a magnitude imbalance and hence ill-conditioned matrices. When using normalized homogeneous coordinates (unit vectors), as suggested in [13] and used in most approaches that explore spherical cameras [12, 29, 30, 31, 32], the reasoning used by Hartley no longer applies.

Most of approaches reviewed in this section assume a distribution model for the matching errors, and disregard the impact of outliers. They also empirically evaluate the effect of the camera FoV on the errors, without any mathematical formalism. Next, we present our bounds for epipolar matrices based on correspondence errors, and provide a tighter relationship between the distribution of matched features and error propagation, showing that wider FoV cameras tend to lead to better epipolar matrix estimation. In particular, spherical cameras present a 360-degree FoV and potentially allows several matches if descriptors tailored to the spherical domain [4, 48] are used, as noted in [7].

3. Perturbation analysis for epipolar matrices

In this section, we briefly review the formulation behind epipolar matrices, provide some generic bounds for singular spaces, and then present error bounds for epipolar matrices. For the particular case of the essential matrix, we also provide bounds for the translation direction.

3.1. Epipolar matrices and the 8-PA

Let us consider a set of $n \geq 8$ 3D points $\boldsymbol{X}_i$ viewed simultaneously by two cameras $\mathcal{C}_1$ and $\mathcal{C}_2$, which are projected to image points $\boldsymbol{x}_1^i$ and $\boldsymbol{x}_2^i$ (in homogeneous coordinates), respectively. The core concept of the epipolar geometry is that each pair of correspondence points $(\boldsymbol{x}_1^i, \boldsymbol{x}_2^i)$ is related by the epipolar constraint [14]

$$\left(\boldsymbol{x}_2^i\right)^\top \mathcal{E} \boldsymbol{x}_1^i = 0, \tag{1}$$

where $\mathcal{E}_{3\times3}$ is an epipolar matrix that depends on the parameters of the two cameras.

The epipolar constraint provided by Eq. (1) is flexible to a variety of scenarios and camera types. For instance, when using two uncalibrated perspective cameras, the correspondence points are given by

$$\begin{aligned} \boldsymbol{x}_1^i &= K_1 \boldsymbol{X}_i = \begin{bmatrix} x_1^i & y_1^i & z_1^i \end{bmatrix}^\top, \\ \boldsymbol{x}_2^i &= K_2 \left(R\boldsymbol{X}_i + \boldsymbol{t}\right) = \begin{bmatrix} x_2^i & y_2^i & z_2^i \end{bmatrix}^\top, \end{aligned} \tag{2}$$

for $i = 1, 2, ..., n$, where K_1 and K_2 are the intrinsic parameters of the two cameras, R is the rotation matrix of the second camera w.r.t. the first, and $\boldsymbol{t}$ is the translation vector. In this case, $\mathcal{E} = F$ is called the *fundamental* matrix, which presents 7-degrees of freedom (DoF).

When using calibrated cameras (i.e., the intrinsics are known), we can pre-multiply pixel values by K_1^{-1} and K_2^{-1}, respectively. This is equivalent to use the notation of Eq. (2) with $K_1 = K_2 = I_{3\times3}$ (identity matrix), so that the image coordinates represent the viewing directions from each camera center to the corresponding 3D point $\boldsymbol{X}_i$. Similarly, spherical cameras capture the full 360°×180° neighborhood, and image coordinates are typically unit vectors that point from the camera center to the corresponding 3D

point [2]. This is the same formulation of calibrated perspective cameras, and in both cases, the epipolar matrix $\mathcal{E} = E$ is called the *essential* matrix, which presents 5-DoF.

In all these cases, the scale of the two-view coordinates $\boldsymbol{x}_c^i$ (for $c \in \{1, 2\}$) is arbitrary, since they relate to homogeneous coordinates. In the calibrated (or spherical) case, the unit vectors $\boldsymbol{x}_c^i / \|\boldsymbol{x}_c^i\|_2$ represent the normalized ray (in the Euclidean space) instead of pixel coordinates (in the projective space). Since they are generic to different central optical systems (*e.g.* calibrated perspective and spherical), they have been used as the parametrization for the correspondences in recent papers such as [13].

The core of the 8-PA is to express the epipolar constraint provided by Eq. (1) as linear combinations of the matrix entries. For simplicity, we will derive all results assuming the calibrated case, i.e., $\mathcal{E} = E = [e_{ij}]$, for $i, j \in \{1, 2, 3\}$.

If $\boldsymbol{e} = [e_{11} \;\; e_{21} \;\; e_{31} \;\; e_{12} \;\; e_{22} \;\; e_{32} \;\; e_{13} \;\; e_{23} \;\; e_{33}]^\top$ is defined as the row-wise concatenation of e_{ij}, then the epipolar constraint can be written as $A\boldsymbol{e} = \boldsymbol{0}$, where A is an $n{\times}9$ measurement matrix for which the i^{th} row is given by:

$$A_i = \begin{bmatrix} x_1^i \left(\boldsymbol{x}_2^i\right)^\top & y_1^i \left(\boldsymbol{x}_2^i\right)^\top & z_1^i \left(\boldsymbol{x}_2^i\right)^\top \end{bmatrix}. \tag{3}$$

The least squares solution for $\boldsymbol{e}$ subject to $\|\boldsymbol{e}\| = 1$ is provided by the least right singular value of A [15], so that the accuracy of the 8-PA relates to the singular subspaces of A. Next, we present some bounds on singular subspaces, and apply them to the context of epipolar matrix estimation.

3.2. Bounds for singular subspaces

Perturbation bounds aim to quantify how the spectrum changes after adding a small perturbation to a matrix, and they play an important role in SVD and spectral methods analysis [3]. Given an approximately rank-r matrix M and a perturbation matrix P, both of dimension $n{\times}m$, an important problem is to understand how much the (left and/or right) singular spaces of M and $\tilde{M} = M + P$ differ from each other [3]. Consider that the SVD decomposition of matrix M is given by

$$M = \begin{bmatrix} U & U_\perp \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V & V_\perp \end{bmatrix}^\top, \tag{4}$$

where $\begin{bmatrix} U & U_\perp \end{bmatrix}$ and $\begin{bmatrix} V & V_\perp \end{bmatrix}$ are orthogonal matrices of orders n and m, respectively, $\Sigma_1 = \mathrm{diag}(\sigma_1, \ldots, \sigma_r)$ and $\Sigma_2 = \mathrm{diag}(\sigma_{r+1}, \ldots)$ are $r{\times}r$ and $(n{-}r){\times}(m{-}r)$ matrices with null off-diagonal values, respectively. Variables $\sigma_1 \geq \sigma_2 \geq \cdots \geq 0$ are the singular values of M in descending order. Decomposing the perturbed matrix $\tilde{M}$ as

$$\tilde{M} = M + P = \begin{bmatrix} \tilde{U} & \tilde{U}_\perp \end{bmatrix} \begin{bmatrix} \tilde{\Sigma}_1 & 0 \\ 0 & \tilde{\Sigma}_2 \end{bmatrix} \begin{bmatrix} \tilde{V} & \tilde{V}_\perp \end{bmatrix}^\top, \tag{5}$$

produces submatrices having the same structures as $U, U_\perp, V, V_\perp, \Sigma_1$ and Σ_2.

A well known bound for estimating the perturbation influence within the singular subspaces comes from Wedin's $\sin\Theta$ theorem [41], which provides a uniform bound for both the left and right singular spaces in terms of the singular value gap and perturbation level. Precisely, it states that if the gap $\delta = \min(\tilde{\Sigma}_1) - \max(\Sigma_2) > 0$, then:

$$\begin{aligned} &\max\left\{\left\|\sin\Theta(V, \tilde{V})\right\|_2, \left\|\sin\Theta(U, \tilde{U})\right\|_2\right\} \\ &\leq \frac{\max\left\{\left\|P\tilde{V}\right\|_2, \left\|\tilde{U}^\top P\right\|_2\right\}}{\delta} \leq \frac{\|P\|_2}{\delta}, \end{aligned} \tag{6}$$

where $\Theta(M_1, M_2) = \mathrm{diag}\left(\cos(\hat{\sigma}_1)^{-1}, \cos(\hat{\sigma}_2)^{-1}, \ldots, \cos(\hat{\sigma}_r)^{-1}\right)$ are the canonical angles between two $p \times r$ orthogonal columns M_1 and M_2 [35], and $\hat{\sigma}_1 \geq \hat{\sigma}_1 \geq \ldots \geq \hat{\sigma}_r \geq 0$ are the singular values of $M_1^\top M_2$. Despite of the wide application range, Wedin's theorem may not be sufficiently precise for some analysis where left and right singular spaces change in different orders of magnitude after the perturbation.

There is a number of works that present tighter perturbation bounds, but applicable only to problems with known noise properties [26, 40]. On the other hand, Cai and Zhang [3] recently established rate-optimal perturbation bounds for the left and right singular spaces separately without any noise assumption. In short, these bounds are given by:

$$\left\|\sin\Theta(U, \tilde{U})\right\| \leq \min\left(\frac{\xi z_{21} + \zeta z_{12}}{\xi^2 - \zeta^2 - \varsigma}, 1\right) \tag{7}$$

and

$$\left\|\sin\Theta(V, \tilde{V})\right\| \leq \min\left(\frac{\xi z_{12} + \zeta z_{21}}{\xi^2 - \zeta^2 - \varsigma}, 1\right), \tag{8}$$

provided that $\xi^2 > \zeta^2 + \varsigma$, where $\xi = \sigma_{\min}(U^\top \tilde{M} V)$, $\zeta = \|U_\perp^\top \tilde{M} V_\perp\|$, $\varsigma = \min(z_{21}^2, z_{12}^2)$, $z_{12} = \|\mathbb{P}_U P \mathbb{P}_{V_\perp}\|$ and $z_{21} = \|\mathbb{P}_{U_\perp} P \mathbb{P}_V\|$. Here, $\mathbb{P}_D$ is projection operator onto the column space of a matrix D [3], and $\|\cdot\|$ is either the spectral ($\|\cdot\|_2$) or the Frobenius ($\|\cdot\|_F$) matrix norm.

The bounds provided in Eqs. (7) and (8) tackle separately the left and right subspaces and are tighter than Wedin's bound [3]. However, they involve projections of the perturbation matrix onto the noiseless right and left singular subspaces, which are not known in practical applications.

3.3. Perturbation bounds and the 8-PA

Here, we relate generic bounds for singular subspaces and the 8-PA. Ideally, the measurement matrix $A_{n\times 9}$ used in the 8-PA (recall Eq. (3)) presents rank $r = 8$. Using the notation of Eq. (4), its SVD generates an $8{\times}8$ diagonal matrix Σ_1 containing all the non-zero singular values of A, with the corresponding left and right singular vectors provided in U

and V, respectively. Also, Σ_2 should be an $(n-8)\times 1$ null matrix, and in particular $\boldsymbol{e} = V_\perp$ is the least right singular vector that contains the elements of the epipolar matrix.

In practice, feature matching is not exact. Without loss of generality (as done in [24]), let us assume that $\boldsymbol{x}_1^i$ corresponds to the exact feature points in the first image and $\tilde{\boldsymbol{x}}_2^i$ to the noisy correspondences in the second image, leading to an approximate matrix $\tilde{A} = A + P$, where P is the perturbation. Due to matching errors, there is no guarantee that $\tilde{A}$ presents rank 8, so that $\tilde{\Sigma}_2$ may not be null.

Our goal here is to estimate the error between the actual essential matrix $\boldsymbol{e} = V_\perp$ and the estimated one $\tilde{\boldsymbol{e}} = \tilde{V}_\perp$, both expressed in vector form. A natural distance measure is the angular distance between them, computed as

$$\theta = \angle(\boldsymbol{e}, \tilde{\boldsymbol{e}}) = \cos^{-1} \left| \boldsymbol{e}^\top \tilde{\boldsymbol{e}} \right|, \tag{9}$$

which is a particular case of the canonical angles [35].

Furthermore, as analyzed in [8], the canonical angles relate to projection errors. More precisely, if $\mathbb{P}_V = VV^\top$ and $\mathbb{P}_{\tilde{V}} = \tilde{V}\tilde{V}^\top$ are the orthogonal projection matrices onto the subspaces spanned by the columns of V and $\tilde{V}$, respectively, then

$$\|\mathbb{P}_V - \mathbb{P}_{\tilde{V}}\|_2 = \left\| \sin \Theta(V, \tilde{V}) \right\|_2. \tag{10}$$

Also, since $\boldsymbol{e}$ and $\tilde{\boldsymbol{e}}$ are the orthogonal complements of V and $\tilde{V}$, respectively, then

$$\begin{aligned} |\sin\theta| &= \|\mathbb{P}_{\boldsymbol{e}} - \mathbb{P}_{\tilde{\boldsymbol{e}}}\|_2 = \|(I - \mathbb{P}_V) - (I - \mathbb{P}_{\tilde{V}})\|_2 \\ &= \|\mathbb{P}_V - \mathbb{P}_{\tilde{V}}\|_2 = \left\| \sin \Theta(V, \tilde{V}) \right\|_2, \end{aligned} \tag{11}$$

recalling that θ is the angle between $\boldsymbol{e}$ and $\tilde{\boldsymbol{e}}$. Combining Eq. (11) with Wedin's bound provided by Eq. (6), we can conclude that

$$|\sin\theta| \leq \frac{\|P\|_2}{\delta} \leq \frac{\|P\|_F}{\delta}, \tag{12}$$

meaning that the error in the estimate of the essential matrix is proportional to the norm of the perturbation P and inversely scaled by the second least singular value of $\tilde{A}$. Although the spectral norm provides a tighter bound, the Frobenius norm will be used, since it can be expressed only in terms of the matching errors. In fact, we can express the perturbation matrix $P = \tilde{A} - A$ as a function of the matched points $\boldsymbol{x}_1^i$, $\boldsymbol{x}_2^i$ and $\tilde{\boldsymbol{x}}_2^i$. Based on Eq. (3), the i^{th} row of P is given by

$$P_i = \begin{bmatrix} x_1^i \left(\Delta \boldsymbol{x}_2^i\right)^\top & y_1^i \left(\Delta \boldsymbol{x}_2^i\right)^\top & z_1^i \left(\Delta \boldsymbol{x}_2^i\right)^\top \end{bmatrix}, \tag{13}$$

where $\Delta \boldsymbol{x}_2^i = \tilde{\boldsymbol{x}}_2^i - \boldsymbol{x}_2^i$. Hence,

$$\begin{aligned} \|P\|_F &= \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{9} |P_{ij}|^2} = \sqrt{\sum_{i=1}^{n} \|\Delta \boldsymbol{x}_2^i\|^2 \|\boldsymbol{x}_1^i\|^2} \\ &= \sqrt{2 \sum_{i=1}^{n} (1 - \cos \alpha_j)}, \end{aligned} \tag{14}$$

where $\alpha_j = \angle(\tilde{\boldsymbol{x}}_2^i, \boldsymbol{x}_2^i)$ is the angular matching error. Hence, the total perturbation $\|P\|_F$ does not depend on the choice of the feature points $\boldsymbol{x}_1^i$ on the first image, but solely on the matching errors on the second image.

Furthermore, a typical error measure when estimating fundamental/essential matrices is based on the relative error using the Frobenius norm. If E and $\tilde{E}$ denote the matrix forms of the essential matrices related to $\boldsymbol{e}$ and $\tilde{\boldsymbol{e}}$, respectively, then $d(E, \tilde{E}) = \min\{\|E - \tilde{E}\|_F, \|E + \tilde{E}\|_F\}$, can be used to measure the error of the estimated matrix. Hence,

$$\begin{aligned} d(E, \tilde{E}) &= \min\{\|\boldsymbol{e} - \tilde{\boldsymbol{e}}\|_2, \|\boldsymbol{e} + \tilde{\boldsymbol{e}}\|_2\} \\ &= \sqrt{2\,(1 - \cos\theta')}, \end{aligned} \tag{15}$$

where $\theta' = \min\{\theta, \pi - \theta\} = \sin^{-1} |\sin\theta| \in [0, \pi/2]$.

Since Eq. (12) provides bounds for $|\sin\theta|$, we have

$$d(E, \tilde{E}) \leq \sqrt{2 \left[1 - \cos\left(\sin^{-1} \min\left\{1, \frac{\|P\|}{\delta}\right\}\right)\right]}. \tag{16}$$

3.4. Relationship between the gap δ and the spatial distribution of the features

One interesting aspect of the bound presented in inequality (6) is that the denominator δ is fully computable based on the observed matrix $\tilde{A}$, without any knowledge on the noiseless matrix A. More precisely, $\delta = \tilde{\sigma}_8$ is the second least singular value of $\tilde{A}$, which depends on several aspects: the 3D structure of the scene (noting that points along a single plane lead to degeneracy), the locations of selected keypoints, the relative camera poses and the FoV of the cameras. In particular, some authors [13, 45, 47] have empirically studied the effect of the camera FoV. Here, we provide a more formal relationship by relating the gap δ with the spatial distribution of the features, which is highly related to the camera FoV.

Let us consider the singular values of $\tilde{A}$ given by $\tilde{\sigma}_1 \geq \tilde{\sigma}_2 \geq \cdots \geq \tilde{\sigma}_9$, where $\tilde{\sigma}_i = \sqrt{\tilde{\lambda}_i}$ and $\tilde{\lambda}_i$ is one of the first nine eigenvalues of $\tilde{A}^\top \tilde{A}$ (or, equivalently, of $\tilde{A}\tilde{A}^\top$). Hartley [15] showed that when using un-normalized homogeneous coordinates, the entries along the diagonal of $\tilde{A}^\top \tilde{A}$ vary considerably in magnitude and used interlacing properties to find estimates on the eigenvalues and condition number of the matrix. When using unit vectors, however, such

analysis is not possible. Instead, we evaluate the impact of the spatial distribution of the matched features, which is strongly affected by the FoV of the cameras.

Merikoski *et al*. [23] presented several bounds for singular values and eigenvalues based on traces. In particular, they showed that for a square matrix $B_{p\times p}$ with real non-negative eigenvalues (in decreasing order), the second least eigenvalue satisfies the following condition:

$$\lambda_{p-1}(B) \leq \frac{\text{tr}(B)}{p-1} - \sqrt{\frac{1}{(p-1)(p-2)}\left(\text{tr}(B^2) - \frac{\text{tr}(B)^2}{p-1}\right)}. \tag{17}$$

If we consider $B = \tilde{A}^\top \tilde{A}$ (so that $p = 9$), we have that

$$\text{tr}(B) = \|\tilde{A}\|_F^2 = \sum_{i=1}^{n} \|\boldsymbol{x}_1^i\|^2 \|\tilde{\boldsymbol{x}}_2^i\|^2 = n, \tag{18}$$

recalling that n is the number of matched points. Since B is symmetric, we also have that $\text{tr}(B^2) = \text{tr}(B^\top B) = \|B\|_F^2$. Let us also consider $C = \tilde{A}\tilde{A}^\top$, so that $\text{tr}(B^\top B) = \text{tr}(C^\top C) = \|C\|_F^2$. Matrix $C = [c_{ij}]_{n\times n}$ presents an interesting structure, since each element is given as a dot product of rows from $\tilde{A}$:

$$\begin{aligned} c_{ij} &= \tilde{A}_i \tilde{A}_j^\top \\ &= x_1^i x_1^j (\tilde{\boldsymbol{x}}_2^i)^\top \tilde{\boldsymbol{x}}_2^j + y_1^i y_1^j (\tilde{\boldsymbol{x}}_2^i)^\top \tilde{\boldsymbol{x}}_2^j + z_1^i z_1^j (\tilde{\boldsymbol{x}}_2^i)^\top \tilde{\boldsymbol{x}}_2^j \\ &= (\boldsymbol{x}_1^i)^\top \boldsymbol{x}_1^j (\tilde{\boldsymbol{x}}_2^i)^\top \tilde{\boldsymbol{x}}_2^j = (\cos\beta_{ij})(\cos\gamma_{ij}), \end{aligned} \tag{19}$$

where $\beta_{ij} = \angle(\boldsymbol{x}_1^i, \boldsymbol{x}_1^j)$ and $\gamma_{ij} = \angle(\tilde{\boldsymbol{x}}_2^i, \tilde{\boldsymbol{x}}_2^j)$ are the angles between features i and j in the first and second images, respectively. Clearly, both β_{ij} and γ_{ij} are limited by the FoV of the camera: if it is small, the entries c_{ij} tend to be closer to one. Also, we have that

$$\|C\|_F^2 = \sum_i \sum_j c_{ij}^2 = \sum_i \sum_j (\cos^2\beta_{ij})(\cos^2\gamma_{ij}). \tag{20}$$

Recalling that $\tilde{\sigma}_8 = \sqrt{\tilde{\lambda}_8}$, we simplify Eq. (17) to obtain

$$\tilde{\sigma}_8 \leq \sqrt{\frac{n}{8} - \frac{1}{8}\sqrt{\frac{8\|C\|_F^2 - n^2}{7}}}. \tag{21}$$

If all the angles β_{ij} and γ_{ij} are small, $\|C\|_F$ tends to be larger, yielding a smaller value for $\delta = \tilde{\sigma}_8$ and hence more potential sensibility to perturbations. In the limit, we have $\|C\|_F \approx n$, which leads to $\delta \approx 0$. In this case, even small perturbations P can lead to highly degraded estimates for the epipolar matrix. On the other hand, the bound in inequality (21) is at most $\sqrt{n}/8$, which is an "optimistic" upper bound for δ (best case scenario), leading to $\delta = \mathcal{O}(\sqrt{n})$. Also, note that a single outlier can significantly increase the perturbation $\|P\|_F$ according to Eq. (14), so that n must be very large to compensate for the presence of "bad" outliers.

Our analysis can be easily extended to weighted versions of the 8-PA, which is used in Iterative Reweighted Least-Squares schemes (IRLS) [39] or in the loss function of recent deep learning approaches [46]. This involves defining an $n\times n$ diagonal matrix $W = [w_{ij}]$ with the weights for each correspondence pair and minimizing $\|WA\boldsymbol{e}\|^2$. In that case, the perturbation error is $\|WP\|_F = \sqrt{\sum_{i=1}^n w_{ii}^2 \|\tilde{\boldsymbol{x}}_2^i - \boldsymbol{x}_2^i\|^2}$, so that the influence of outliers can be alleviated by choosing small weights for bad matches.

3.5. Perturbation analysis of 5-DoF pose estimation from the essential matrix

For calibrated or spherical cameras, the 5-DoF pose parameters – rotation and the direction of the translation vector – can be extracted from the essential matrix through the SVD [14]. In particular, the direction of the translation vector is given by the least left singular vector $\boldsymbol{t}$ of E, and it is more prone to errors in the essential matrix than the rotation matrix, as noted in [25, 38].

Let us consider a true rank-2 essential matrix E with $\|E\|_F = 1$, and let $\boldsymbol{t}$ and $\tilde{\boldsymbol{t}}$ denote the least left singular values of E and its estimate $\tilde{E}$, respectively, and assume that the direction ambiguity was solved (*e.g.* by using cheirality constraints [37]). Using the notation of Eqs. (4) and (5), the SVD of E generates a 2×2 diagonal matrix $\Sigma_1 = \frac{1}{\sqrt{2}} I_2$ containing the two equal singular values of E, and an 1×1 null matrix Σ_2 (consider an analogous notation for the SVD of $\tilde{E}$). The least left singular vectors of E and $\tilde{E}$ are given by $\boldsymbol{t} = U_\perp$ and $\tilde{\boldsymbol{t}} = \tilde{U}_\perp$, respectively. Note that the gap between $\tilde{\Sigma}_1$ and Σ_2 is given by $\delta_E = \min\{\tilde{\Sigma}_1\} - \max\{\Sigma_2\} = \sigma_2(\tilde{E})$, so that Wedin's theorem gives

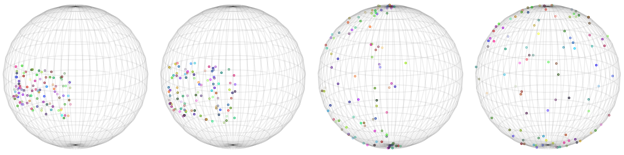
$$|\sin\omega| \leq \frac{1}{\sigma_2(\tilde{E})}\|E - \tilde{E}\|_F = \frac{1}{\sigma_2(\tilde{E})}\|\boldsymbol{e} - \tilde{\boldsymbol{e}}\|, \tag{22}$$

where $\omega = \angle(\tilde{\boldsymbol{t}}, \boldsymbol{t})$ is the angle between the actual and the estimated translation values, and $\sigma_2(\tilde{E})$ is the second least singular value of $\tilde{E}$. For small perturbations, we expect $\sigma_2(\tilde{E}) \approx \sigma_2(E) = 1/\sqrt{2}$. More precisely, Weyl's bound [43] relates the q^{th} pair of singular values $\sigma_q(\tilde{E})$ and $\sigma_q(E)$ through

$$\left|\sigma_q(\tilde{E}) - \sigma_q(E)\right| \leq \|E - \tilde{E}\|_2 \leq \|E - \tilde{E}\|_F, \tag{23}$$

so that a looser version of the bound in Eq. (22) can be expressed solely based on the difference between $\boldsymbol{e}$ and $\tilde{\boldsymbol{e}}$:

$$|\sin\omega| \leq \frac{\sqrt{2}\|\boldsymbol{e} - \tilde{\boldsymbol{e}}\|}{1 - \sqrt{2}\|\boldsymbol{e} - \tilde{\boldsymbol{e}}\|}. \tag{24}$$



(a) 54.4°×37.8° (b) 65.5°×46.4° (c) 195°×195° (d) 360°×180°

Figure 1: Unitary feature projections in different FoVs.

4. Experimental results

4.1. Synthetic feature matching

In our experimental setup, we first present results using a set of synthetic 3D points projected to calibrated cameras with known parameters, and add artificial noise to the feature locations on the second view. Since the feature are 3D unit vectors (i.e., on the unit sphere), we add von Mises-Fisher (vMF) noise, as done in [12]. We consider the FoVs of typical perspective cameras with $54.4°\times37.8°$ and $65.5°\times46.4°$ [27], a $195°$ fisheye wide-angle camera [19] and a full-spherical camera [2].

Noise is controlled by parameter κ in the vMF distribution. Here, we select $\kappa \in \{500; 1,000; 2,000; 10,000\}$ corresponding to average angular matching errors equivalent to $3.21°$, $2.27°$, $1.60°$ and $0.72°$, respectively. For each combination of FoV and noise level, we generate 1,000 experiments, each one containing 100 3D points randomly selected within a 5-10m radius (constrained to the camera FoV), simulating a large indoor environment. The second camera was randomly placed within a $[-1, 1]^3$ cube with arbitrary rotation. For the sake of illustration, the spherical projection of one set of features using the four selected FoVs is depicted in Figure 1.

Table 1 presents the average sine error between $\boldsymbol{e}$ and $\tilde{\boldsymbol{e}}$ for each combination, as well as the Wedin's bound (Eq. (12)) and Cai and Zhangs' bound (Eq. (8))[1]. Both bounds decrease as the noise level decreases and the FoV increases, as expected. However, on average they showed to be quite loose bounds when compared to the actual errors. Cai and Zhangs' bound tends to be tighter than the Wedin's, but it is important to recall that it is not computable on practical applications, and thus we will focus only on Wedin's bound hereafter in our analysis.

Although Wedin's bound showed to be loose, it still provides useful insights into the essential matrix accuracy as a function of the camera FoVs. Figure 2 illustrates the averaged results for the singular gap δ, the sine error and Wedin's perturbation bound by using an extensive combination of the horizontal and vertical FoVs (abbreviated as HFoV and VFoV, respectively). A total of 100 simulations per HFoV×VFoV was performed using the same setup explained before, and Wedin's bound was truncated in value 1. Also, we vary the noise levels, setting $\kappa = 500$, $\kappa = 10,000$ and $\kappa = 1,000,000$ which correspond to an average angular error of $3.21°$, $0.72°$ and $0.071°$, respectively. The selected range for κ encompasses the tolerances of $0.5625°$ and $2°$ for considering corresponding points as true matches as argued in [48] and [11], respectively.

Figure 2 shows that the maximum δ occurs around the full 360 degrees FoV (first row), presenting a practically stable value regardless of the tested noise level. In fact, the variance on the full FoV for all noise levels is around 2.8×10^{-5}. Interestingly, the "optimal" VFoV is slightly below $180°$. We believe that this happens because using the full VFoV leads to a circular domain, which might increase the number of neighboring features. Moreover, as expected, when the noise level increases the sine error also increases (second row), and Wedin's bound behaves similarly (third row). It is also possible to see that for higher noise levels combined with narrower FoVs, Wedin's bound turns to be useless because its value is even greater than the trivial bound 1. Last but not least, we found in our experiments that the Spearman's correlation [34] between the the gap δ and the *diagonal* FoV (DFoV) is around 0.775, 0.863 and 0.877 for $\kappa = \{500; 10,000; 1,000,000\}$, respectively (p-value $\ll 0.01$), indicating a strong relationship.

The second part of the analysis consists of estimating the accuracy of the 5-DoF pose extracted from the essential matrix (here, we assume the calibrated/spherical case). Our evaluation metric for the 2-DoF translation vector is the angular error [45] given by

$$\varepsilon_{\mathbf{t}} = \cos^{-1}(\mathbf{t}^\top \tilde{\mathbf{t}}). \tag{25}$$

For the sake of illustration, we also show the rotation matrix error, given as the angles between the actual matrix R and the estimate $\tilde{R}$ [38], which is defined as

$$\varepsilon_R = \cos^{-1}\left(\frac{\mathrm{tr}(R^\top \tilde{R}) - 1}{2}\right). \tag{26}$$

Figure 3 presents the average translation and rotation errors as a function of the angular matching error. The distribution of the 3D points, minimum and maximum values for κ and the FoVs are the same as in the experiment related to Table 1. Note that the rotation error is much smaller than the translation error not only for narrow FoVs, as noted in [25, 38], but also for wider FoV cameras. It is also evident that the translation error decreases as the FoV increases, since the essential matrix is estimated more accurately. For the wider FoVs ($195°\times195°$ and $360°\times180°$) we also present Wedin's bound for the angle between the actual and the estimated 2-DoF translation vector, as given in Eq. (22). For narrower FoVs the bound is greater than the trivial value, and hence not shown.

[1] Since these bounds produce trivial values (≥ 1) for narrow FoVs, we only show the results for wider FoVs.

κ	54.4°×37.8°	65.5°×46.4°	195°×195°	360°×180°	195°×195°	360°×180°	195°×195°	360°×180°
	Sine error				Wedin's bound [41]		Cai and Zhangs' bound [3]	
500	0.782 ± 0.230	0.778 ± 0.226	0.340 ± 0.252	0.085 ± 0.045	0.868 ± 0.088	0.607 ± 0.076	0.891 ± 0.073	0.671 ± 0.148
1,000	0.781 ± 0.221	0.756 ± 0.241	0.190 ± 0.180	0.054 ± 0.027	0.769 ± 0.136	0.446 ± 0.054	0.705 ± 0.174	0.335 ± 0.082
2,000	0.780 ± 0.223	0.756 ± 0.234	0.103 ± 0.090	0.036 ± 0.017	0.666 ± 0.165	0.326 ± 0.039	0.543 ± 0.205	0.177 ± 0.036
10,000	0.679 ± 0.248	0.563 ± 0.253	0.033 ± 0.025	0.015 ± 0.007	0.356 ± 0.117	0.149 ± 0.017	0.223 ± 0.145	0.053 ± 0.011

Table 1: Impact of the variation in FoV and noise level when computing the perturbation levels.

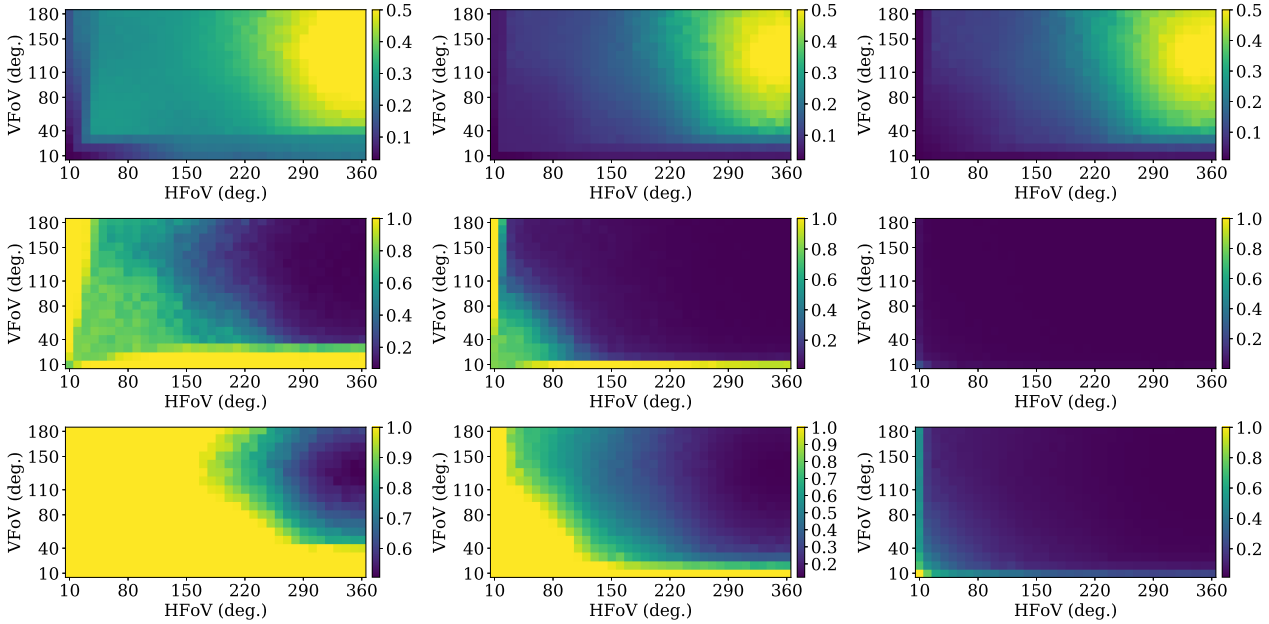
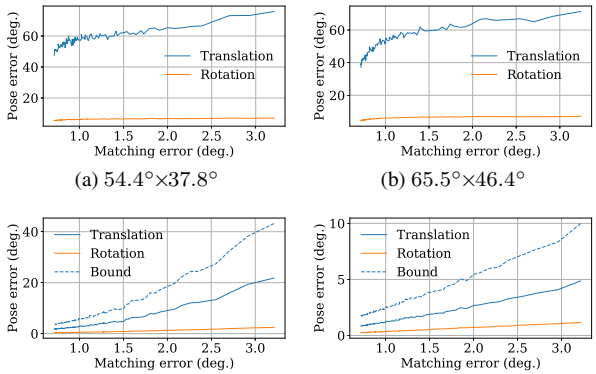


Figure 2: Average results for the delta value, the sine error and the Wedin's bound (in the rows) for different noise levels (in the columns) and FoVs. From the left to the right, $\kappa = 500$, $\kappa = 10,000$ and $\kappa = 1,000,000$.



(a) 54.4°×37.8° (b) 65.5°×46.4°

(c) 195°×195° (d) 360°×180°

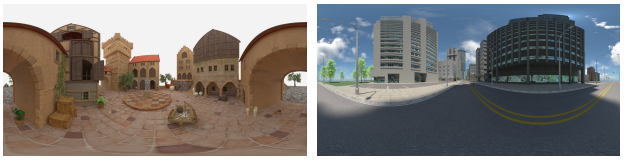
Figure 3: 5-DoF pose error for different noise levels.

To evaluate the impact of outliers, we corrupted a set of $n \in [10; 3,500]$ noisy matchings ($\kappa = 16,250$) with a single outlier. The actual sine error and Wedin's bound decay proportionally to $\mathcal{O}(1/\sqrt{n})$, leading to Pearson's correlations $\rho = 0.9932$ and $\rho = 0.9973$ (p-value $\ll 0.01$), respectively. This corroborates our findings in Section 3.4.

4.2. Real feature matching

Although the vMF noise model is suitable for the chosen features, matching noise in real images is highly dependent on the feature extractor and the local appearance of the images. Since we are not aware of existing datasets with wide FoV cameras (*e.g.*, spherical) and ground truth data w.r.t. the essential matrix, we use realistic computer generated scenes as done in [10, 11, 47]. We rendered non-aligned and non-rectified pairs of spherical images using the Blender Cycles models *Urban Canyon* and *Indoor*, made available by [47], and the *Classroom*, recently used in [6]. We also considered scene captures of the *Medieval Port* model along with the 6-DoF pose ground-truth from [10]. The former and the latter datasets are outdoors, and the other two are indoors. All images are rendered at a 1280×640 resolution in equirectangular format, and Figure 4 illustrates a single spherical view from each dataset.

To obtain the required correspondences for the 8-PA, we



(a) *Medieval Port* [10] (b) *Urban Canyon* [47]

(c) *Indoor* [47] (d) *Classroom* [6]

Figure 4: Datasets used for validation.

used the spherical ORB (SPHORB) [48], which is suited for spherical images, faster than spherical SIFT (SSIFT) [4] and with publicly available code (https://github.com/tdsuper/SPHORB). Given the correspondence pairs, we robustly estimate the epipolar matrix by using RANSAC, and consider a feature pair as inlier if its symmetric projected distance [29] is smaller than 10^{-2}. We accept a model if it has at least 70% of inliers.

Tables 2 and 3 present the average translation and rotation errors (Eqs. (25) and (26)), as well as the epipolar error (Eq. (15)) and the δ value for the four datasets and the four different FoVs (set as in the synthetic matching experiments). For each experiment, a total of 1,000 pairs of images was randomly selected, and the FoV was restricted so that the narrow FoV cameras are pointing out to some location aligned to the scene's horizon. Besides the results for 8-PA, we also show the pose errors after applying the non-linear 5-DoF pose refinement (NLR) based on the projected distance that is, among other three options, pointed out as the best performing in [29].

In the experiment shown in Table 2, the number of matchings is restricted to be the same as the one of the smaller FoV, so that the main observable variable on the results is the spreading of the features. The average number of keypoints in this experiment was 132.8 ± 83.9. Note that the results in synthetic images with real feature matching corroborate the results from Section 4.1, i.e., the wider the FoV the smaller the epipolar and 5-DoF pose errors, and larger the value of δ.

Table 3 shows the metrics using all available matches for each tested FoV (more matches are expected for wider FoVs). The average number of correspondences in this new test indeed increased with the FoV: 134.8 ± 86.0, 166.8 ± 115.1, 722.9 ± 681.3 and 1288.4 ± 1189.0, respectively. Our results indicate that increasing the number of ("good") features indeed helps to improve even more the 8-PA results, especially for wider FoVs. Also, as noted

Metric	$54.4^\circ \times 37.8^\circ$	$65.5^\circ \times 46.4^\circ$	$195^\circ \times 195^\circ$	$360^\circ \times 180^\circ$
$d(E, \tilde{E})$	0.849±0.687	0.815±0.698	0.134±0.287	0.100±0.260
$\varepsilon_\mathbf{t}$ (8-PA)	39.727±36.301	37.445±35.618	11.941±37.707	8.936±27.426
$\varepsilon_\mathbf{t}$ (NLR)	38.497±37.456	35.770±37.816	11.521±37.899	8.214±27.770
ε_R (8-PA)	8.383±20.913	7.629±18.385	1.923±7.802	1.684±8.952
ε_R (NLR)	8.276±20.852	7.235±17.511	1.500±7.022	1.453±8.705
δ	0.033±0.017	0.038±0.022	0.150±0.087	0.257±0.157

Table 2: Results for synthetic imagery for different FoVs when the number of keypoints is limited.

Metric	$54.4^\circ \times 37.8^\circ$	$65.5^\circ \times 46.4^\circ$	$195^\circ \times 195^\circ$	$360^\circ \times 180^\circ$
$d(E, \tilde{E})$	0.818±0.667	0.777±0.675	0.053±0.611	0.038±0.052
$\varepsilon_\mathbf{t}$ (8-PA)	37.280±33.974	35.323±33.994	8.480±34.474	3.427±15.588
$\varepsilon_\mathbf{t}$ (NLR)	35.936±34.534	33.728±36.002	8.421±34.469	3.157±15.566
ε_R (8-PA)	6.053±10.479	5.417±7.015	0.2865±0.5924	0.083±0.077
ε_R (NLR)	6.203±12.603	5.126±7.209	0.208±0.519	0.065±0.066
δ	0.034±0.017	0.043±0.023	0.388±0.188	0.851±0.365

Table 3: Results for synthetic imagery for different FoVs with free number of keypoints.

by [47], wider FoVs greatly improve pose and 3D estimation based on non-linear bundle adjustment algorithms since features are more likely to be visible in more than two captures of temporally aligned image sets.

5. Conclusions

We present a perturbation analysis for epipolar matrix estimation using the well-known 8-PA by exploring singular subspace analysis. We show that the bound is inversely proportional to the second least singular value of the observation matrix, which is strongly affected by the spatial distribution of the matched features. In particular, the features extracted when using narrow FoV images are spatially concentrated, leading to larger bounds (and according to our experiments, also larger errors in the estimate of the epipolar matrix). On the other hand, cameras with wider FoV (in the limit case spherical images) present a much better spatial distribution of features, leading to smaller bounds and smaller effective errors in the estimated matrix. This suggests that expensive non-linear approaches for pose (or pose plus 3D) refinement might be relaxed or even suppressed when using spherical cameras.

In the future, we intend to approximate tighter bounds (as [3]) using only observable data, better explore the singular gaps of the epipolar matrices for the 2-DoF translation vector, and extend our bounds for the rotation matrix case.

Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 - and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). The authors also thank Christiano Gava, Didier Stricker and colleagues for giving access to the *Medieval Port* dataset.

References

[1] Sameer Agarwal, Hon-Leung Lee, Bernd Sturmfels, and Rekha R Thomas. On the existence of epipolar matrices. *International Journal of Computer Vision*, 121(3):403–415, 2017.

[2] Torii Akihiko, Imiya Atsushi, and Naoya Ohnishi. Two-and three-view geometry for spherical cameras. *Proc. of the Sixth Workshop on Omnidirectional Vision, Camera Networks and Non- classical Cameras*, 105:29–34, 2005.

[3] T Tony Cai, Anru Zhang, et al. Rate-optimal perturbation bounds for singular subspaces with applications to high-dimensional statistics. *The Annals of Statistics*, 46(1):60–89, 2018.

[4] Javier Cruz-Mota, Iva Bogdanova, Benoît Paquier, Michel Bierlaire, and Jean Philippe Thiran. Scale invariant feature transform on the sphere: Theory and applications. *International Journal of Computer Vision*, 98(2):217–241, 2012.

[5] Gabriella Csurka, Cyril Zeller, Zhengyou Zhang, and Olivier D Faugeras. Characterizing the uncertainty of the fundamental matrix. *Computer vision and image understanding*, 68(1):18–36, 1997.

[6] Thiago L.T. da Silveira, Lorenzo P. Dalaqua, and Claudio R. Jung. Indoor Depth Estimation from Single Spherical Images. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pages 2935–2939. IEEE, oct 2018.

[7] Thiago L.T. da Silveira and Claudio R. Jung. Evaluation of Keypoint Extraction and Matching for Pose Estimation Using Pairs of Spherical Images. *2017 30th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 374–381, 2017.

[8] Zlatko Drmac. On principal angles between subspaces of euclidean space. *SIAM Journal on Matrix Analysis and Applications*, 22(1):173–194, 2000.

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

[10] Christiano Couto Gava, Didier Stricker, and Soichiro Yokota. Dense Scene Reconstruction from Spherical Light Fields. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pages 4178–4182. IEEE, oct 2018.

[11] Hao Guan and William A P Smith. BRISKS: Binary Features for Spherical Images on a Geodesic Grid. In *Conference on Computer Vision and Pattern Recognition*, 2017.

[12] Hao Guan and William A. P. Smith. Structure-From-Motion in Spherical Video Using the von Mises-Fisher Distribution. *IEEE Transactions on Image Processing*, 26(2):711–723, feb 2017.

[13] Simon James Hadfield, Karel Lebeda, and Richard Bowden. HARD-PnP: PnP Optimization Using a Hybrid Approximate Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8828(c):1–1, 2018.

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

[15] Richard I Hartley. In defence of the 8-point algorithm. In *Computer Vision, 1995. Proceedings., Fifth International Conference on*, pages 1064–1070. IEEE, 1995.

[16] Daiki Ikami, Toshihiko Yamasaki, and Kiyoharu Aizawa. Fast and robust estimation for unit-norm constrained linear fitting problems. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8147–8155, June 2018.

[17] Jingwei Huang, Zhili Chen, Duygu Ceylan, and Hailin Jin. 6-DOF VR Videos with a Single 360-Camera. In *Proceedings of the IEEE Virtual Reality*. IEEE, 2017.

[18] John Lim, Nick Barnes, and Hongdong Li. Estimating relative camera motion from the antipodal-epipolar constraint. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(10):1907–1914, 2010.

[19] I-chan Lo, Kuang-tsu Shih, and Homer H Chen. Image Stitching for Dual Fisheye Cameras. *2018 25th IEEE International Conference on Image Processing (ICIP)*, pages 3164–3168, 2018.

[20] H Christopher Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293(5828):133, 1981.

[21] Quan-Tuan Luong and Olivier D Faugeras. The fundamental matrix: Theory, algorithms, and stability analysis. *International Journal of Computer Vision*, 17(1):43–75, jan 1996.

[22] Elmar Mair, Michael Suppa, and Darius Burschka. Error propagation in monocular navigation for zinf compared to eightpoint algorithm. In *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*, pages 4220–4227. IEEE, 2013.

[23] Jorma Kaarlo Merikoski, Humberto Sarria, and Pablo Tarazaga. Bounds for singular values using traces. *Linear Algebra and its Applications*, 210:227–254, 1994.

[24] Matthias Mühlich and Rudolf Mester. The role of total least squares in motion analysis. In *European Conference on Computer Vision*, pages 305–321. Springer, 1998.

[25] David Nistér. An efficient solution to the five-point relative pose problem. *IEEE transactions on pattern analysis and machine intelligence*, 26(6):756–770, 2004.

[26] Sean O'Rourke, Van Vu, and Ke Wang. Random perturbation of low rank matrices: Improving classical bounds. *arXiv preprint arXiv:1311.2657*, 2013.

[27] Thomas Wesley Osborne, Todor Georgiev Georgiev, and Sergiu Radu Goma. Wide field of view array camera for hemispheric and spherical imaging, 2014. Qualcomm Inc. US9819863B2.

[28] Onur Ozyesil, Vladislav Voroninski, Ronen Basri, and Amit Singer. A Survey of Structure from Motion. pages 305–364, 2017.

[29] Alain Pagani and Didier Stricker. Structure from Motion using full spherical panoramic cameras. In *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, pages 375–382. IEEE, nov 2011.

[30] Sarthak Pathak, Alessandro Moro, Hiromitsu Fujii, Atsushi Yamashita, and Hajime Asama. 3D reconstruction of structures using spherical cameras with small motion. In *2016 16th International Conference on Control, Automation and*

Systems (ICCAS), number Iccas, pages 117–122. IEEE, oct 2016.

[31] Sarthak Pathak, Alessandro Moro, Hiromitsu Fujii, Atsushi Yamashita, and Hajime Asama. Distortion-Robust Spherical Camera Motion Estimation via Dense Optical Flow. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pages 3358–3362. IEEE, oct 2018.

[32] Sarthak Pathak, Alessandro Moro, Atsushi Yamashita, and Hajime Asama. Optical Flow-Based Epipolar Estimation of Spherical Image Pairs for 3D Reconstruction. *SICE Journal of Control, Measurement, and System Integration*, 10(5):476–485, 2017.

[33] Rahul Raguram, Jan-Michael Frahm, and Marc Pollefeys. A comparative analysis of ransac techniques leading to adaptive real-time random sample consensus. In *European Conference on Computer Vision*, pages 500–513. Springer, 2008.

[34] C. Spearman. The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1):72–101, 1904.

[35] Gilbert W Stewart. Matrix perturbation theory. 1990.

[36] Frédéric Sur, Nicolas Noury, and Marie-Odile Berger. Computing the uncertainty of the 8 point algorithm for fundamental matrix estimation. In *19th British Machine Vision Conference-BMVC 2008*, page 10, 2008.

[37] Richard Szeliski. *Computer Vision: Algorithms and Applications*. Springer-Verlag, Berlin, Heidelberg, 1st edition, 2010.

[38] Tina Yu Tian, Carlo Tomasi, and David J Heeger. Comparison of approaches to egomotion computation. In *Computer Vision and Pattern Recognition, 1996. Proceedings CVPR'96, 1996 IEEE Computer Society Conference on*, pages 315–320. IEEE, 1996.

[39] Philip HS Torr and David W Murray. The development and comparison of robust methods for estimating the fundamental matrix. *International journal of computer vision*, 24(3):271–300, 1997.

[40] Rongrong Wang. Singular vector perturbation under gaussian noise. *SIAM Journal on Matrix Analysis and Applications*, 36(1):158–177, 2015.

[41] Per-Åke Wedin. Perturbation bounds in connection with singular value decomposition. *BIT Numerical Mathematics*, 12(1):99–111, 1972.

[42] Juyang Weng, Thomas S. Huang, and Narendra Ahuja. Motion and structure from two perspective views: Algorithms, error analysis, and error estimation. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (5):451–476, 1989.

[43] Hermann Weyl. Das asymptotische verteilungsgesetz der eigenwerte linearer partieller differentialgleichungen (mit einer anwendung auf die theorie der hohlraumstrahlung). *Mathematische Annalen*, 71(4):441–479, 1912.

[44] FC Wu, ZY Hu, and FQ Duan. 8-point algorithm revisited: Factorized 8-point algorithm. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 1, pages 488–494. IEEE, 2005.

[45] Jiaolong Yang, Hongdong Li, and Yunde Jia. Optimal Essential Matrix Estimation via Inlier-Set Maximization. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 8689 LNCS, pages 111–126. 2014.

[46] Kwang Moo Yi, Eduard Trulls, Yuki Ono, Vincent Lepetit, Mathieu Salzmann, and Pascal Fua. Learning to find good correspondences. In *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2666–2674, 2018.

[47] Zichao Zhang, Henri Rebecq, Christian Forster, and Davide Scaramuzza. Benefit of large field-of-view cameras for visual odometry. *Proceedings - IEEE International Conference on Robotics and Automation*, 2016-June:801–808, 2016.

[48] Qiang Zhao, Wei Feng, Liang Wan, and Jiawan Zhang. SPHORB: A Fast and Robust Binary Feature on the Sphere. *International Journal of Computer Vision*, 113(2):143–159, 2014.

[49] Yinqiang Zheng, Shigeki Sugimoto, and Masatoshi Okutomi. A practical rank-constrained eight-point algorithm for fundamental matrix estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1546–1553, 2013.