Exact Two–Image Structure from Motion

John Oliensis (oliensis@research.nj.nec.com)
NEC Research Institute
4 Independence Way
Princeton, N.J. 08540

Abstract

For two–image structure from motion, we present a simple, *exact* expression for a least–squares image–reprojection error for *finite* motion that depends only on the motion. Optimal estimates of the structure and motion can be computed by minimizing this expression just over the motion parameters. Also, we present a solution to the triangulation problem: an exact, *explicit* expression for the optimal structure estimate given the motion. We identify a new ambiguity in recovering the structure for known motion. We study the exact error's properties experimentally and demonstrate that it often has several local minima for forward or backward motion estimates. Our experiments also show that the "reflected" local minimum of Oliensis and Soatto et al. occurs for large translational motions. Our exact results assume that the camera is calibrated and use a least–squares image–reprojection error that applies most naturally to a spherical imaging surface. We approximately extend our approach to the standard least–squares error in the image plane and uncalibrated cameras. We present an improved version of the Sampson error which gives better results experimentally.

Keywords: structure from motion, two–image structure from motion, least–squares error, triangulation, ambiguity, spherical retina, coplanarity, Sampson error, local minima.

1 Introduction

This paper revisits one of the oldest and most fundamental problems in structure from motion (SFM): estimating the 3D structure of a scene and the camera motion from 2 images. The problem goes back more than 20 years [12][8] and has great practical importance. Most current SFM algorithms use a 2–image technique in their initial stages, and, despite the fact that 2 images contain less information than a larger number, recent research has shown that one can often estimate structure and motion from 2 images with surprising accuracy and robustness [28][18][17][23][22].

All algorithms for 2–image SFM use the same basic strategy. They first compute an initial estimate of the camera motion, either by a linear method [12][9] or, more recently, using RANSAC and the "7–point" algorithm [7][24]. In a second, "polishing" stage, they improve the initial estimate by minimizing an error function with respect to the unknown motion and/or structure. This paper focuses on the second stage.

The only current polishing technique that is optimally accurate, bundle adjustment (e.g., [25][6]), minimizes an error function over *all* unknowns, the structure as well as the motion. Since it minimizes over a large number of variables, it is relatively slow and its behavior is opaque: one cannot easily predict whether it will find the best estimate at the lowest error or confirm that it has done so after it has converged. Another standard technique minimizes an approximation to the reprojection error that depends just on the five motion unknowns. The most popular approximation we call the *weighted–coplanarity* or *WC* error (e.g., [27][28]); [4] has recently dubbed it the *Sampson error*. Minimizing the *WC* error is faster than bundle adjustment, but it gives poor accuracy in some situations.

In this paper, we propose an approach to 2–image SFM which resolves the dilemma between speed/reliability and accuracy. We start from a slight modification of the standard least–squares image–reprojection error: the image–reprojection error that we use applies most naturally to a camera with a spherical imaging surface. (We point out that this error has advantages over the standard image–plane least–squares error even for a standard camera with a planar imaging surface.) Our change in error greatly simplifies the SFM problem. By exploiting it, we present in Section 2 a simple, *exact* expression for the least–squares error for *finite* motion that depends only on the five motion parameters. Minimizing this expression is as fast as minimizing the *WC* error and as optimal as bundle adjustment. Because the minimization is over a small number of parameters, it is relatively transparent: there is a good chance of understanding whether it will succeed in finding the best estimate and of ensuring that it does [15].

As a corollary, Section 3 presents a solution to the stereo triangulation problem: a simple, exact, *explicit* expression for the optimal estimate of the 3D structure given the motion. It is important to ensure that the structure estimate is optimal, since this estimate depends extremely sensitively on the noise. Also, we describe a new ambiguity in recovering the structure for known motion.

Our expression for the structure is optimal for our spherical–image least–squares error. For the standard image–plane error, [5] gave an implicit solution for the optimal structure estimate in terms of the root of a 6th degree polynomial. Recently, for a new version of the image–plane error, Nister computed the corresponding optimal estimate in terms of the root of a 4th degree polynomial [13].

In Section 4, we use our exact theoretical results to study the least–squares error for finite motion experimentally. The results of such a study can be used to improve algorithms [17] or to understand the intrinsic difficulty confronted by any algorithm in searching for the optimal estimate [15]. Due to the simplicity of our expressions, we can easily compute the error over a range of motion estimates. We show that it typically has several local minima for forward (or backward) motion estimates. Some of these pose the real danger that an algorithm may mistakenly converge to them. Also, our experiments show that the error's "reflected" local minimum at sideways estimates, which was analyzed in [17][20][1][21], occurs even for large translations.

Our exact results assume that the camera is calibrated. In Section 5, we describe an approximate extension of our approach to 2 uncalibrated cameras (the projective case). Also, we approximately extend our approach from the spherical–image error to the standard least–squares error in the image plane. Our ambiguity in estimating the structure also occurs for the image–plane error for some motions and, as we discuss, it helps to explain the failures of the *WC* (Sampson) approximation. In the Appendix, we derive an improved version of the WC error and show that it gives a better approximation to the least–squares image–plane error.

Though this paper focuses on the least–squares error, our results can be applied also for robust estimation.

Our exact expression for the least–squares error that depends only on the motion appeared in [19].

1.1 Definitions and Preliminaries

We use MATLAB notation: we represent a column vector by listing entries separated by ';' and a row vector by separating entries by a comma or space. A colon indicates a range of indices. Also, for a set of quantities V_a, we use $\{V\}$ to denote the column vector with entries V_a. Let $\mathbf{1}_N$ be the $N \times N$ identity

matrix.

Unless otherwise stated, we assume the camera is calibrated and, without loss of generality, we take the focal length to be 1. We assume there are 2 images and N_p points tracked over the images. Let the column vector $\mathbf{p}_{im} \equiv (x_{im}; y_{im})$ denote the mth point in the ith image, where $m = \{1, 2, ..N_p\}$ and $i = \{0, 1\}$. Define $\mathbf{P}_m \equiv \begin{pmatrix} X_m; & Y_m; & Z_m \end{pmatrix}$ to be the mth 3D point in the coordinate system of the zeroth image. Z_m is the *depth* of this point. Let $\hat{\mathbf{x}} = (1; 0; 0)$, $\hat{\mathbf{y}} \equiv (0; 1; 0)$, and $\hat{\mathbf{z}} \equiv (0; 0; 1)$.

Let $\mathbf{T} \equiv (T_x; T_y; T_z)$ be the translation from image 0 to image 1, and let $\hat{\mathbf{T}} \equiv \mathbf{T}/|\mathbf{T}|$ be the translation direction. Let R be the rotation between images 0 and 1. We define the motion of a 3D point $\mathbf{P}$ under R and $\mathbf{T}$ by $\mathbf{P} \longrightarrow R(\mathbf{P} - \mathbf{T})$. Let $F(R, \mathbf{T}) \in \Re^{3\times3}$ represent the essential matrix [4]: F acting on a 3D vector $\mathbf{V}$ gives $F\mathbf{V} = R(\mathbf{T} \times \mathbf{V})$.

If $\mathbf{V}$ is a vector, let $\hat{\mathbf{V}}$ denote the unit vector $\mathbf{V}/|\mathbf{V}|$. Let $\underline{\mathbf{V}}$ denote the 2D image point corresponding to the 3D point $\mathbf{V}$, with $\underline{\mathbf{V}} \equiv (V_x; V_y)/V_z$. For 2D vectors $\mathbf{v}$, $\mathbf{v}'$ we use the notation $\mathbf{v} \times \mathbf{v}'$ to mean $v_x v'_y - v_y v'_x$. Define the 3D point $\bar{\mathbf{v}}$ corresponding to $\mathbf{v}$ by $\bar{\mathbf{v}} \equiv \begin{bmatrix} \mathbf{v}; & 1 \end{bmatrix}$. Let $R * \mathbf{v}$ denote the image point obtained after a rotation of the image point $\mathbf{v}$, with $R * \mathbf{v} \equiv \underline{(R\bar{\mathbf{v}})}$. Let $\mathbf{e} \equiv \underline{\mathbf{T}}$ denote the epipole, the image of $\mathbf{T}$ in image 0.

Given a matrix M, let $[[M]]_2$ be the 2×2 submatrix of M consisting of entries from the first 2 rows and columns.

For finite motion, the standard least–squares image–reprojection error in the image plane is

$$E_{LS}(R, \mathbf{T}, \{\mathbf{P}\}) \equiv \sum_{m=1}^{N_p} \left(\left| \underline{\mathbf{P}_m} - \mathbf{p}_{0m} \right|^2 + \left| \underline{R(\mathbf{P}_m - \mathbf{T})} - \mathbf{p}_{1m} \right|^2 \right), \tag{1}$$

where the $\mathbf{p}_{im}$ are the measured image points and our notation indicates that E_{LS} depends on all the $\mathbf{P}_m$ as well as on R, $\mathbf{T}$. We define the *optimal least–squares estimates* of R, $\mathbf{T}$, and the $\mathbf{P}_m$ as the "values" that give the least error E_{LS}. (We put "values" in quotes since the least–squares estimates of the $\mathbf{P}_m$ can occur at infinity or at the points $(0; 0; 0)$ or $\mathbf{T}$, where the error E_{LS} is not defined.)

2 An Exact Expression for the Least–Squares Error

2.1 Introduction

Our goal is to derive an effective least–squares error function that depends just on R, $\mathbf{T}$. One can do this by minimizing E_{LS} over the $\mathbf{P}_m$. Define

$$E\left(R,\mathbf{T}\right) \equiv \inf_{\mathbf{P}_1 \ldots \mathbf{P}_{Np}} E_{LS}, \tag{2}$$

where $\inf$ represents the infimum or greatest lower bound.[1] We also refer to $E\left(R,\mathbf{T}\right)$ as the "least–squares error." Due to the standard scale ambiguity in SFM, one can write E as the function $E\left(R,\hat{\mathbf{T}}\right)$, where $\hat{\mathbf{T}}$ is the estimate of the translation direction, or as $E\left(R,\mathbf{e}\right)$, where $\mathbf{e}$ is the epipole estimate. In principle, since $E\left(R,\hat{\mathbf{T}}\right)$ has the same minimum as E_{LS} at the same values for R, $\hat{\mathbf{T}}$, one can solve for the optimal estimates of R and $\hat{\mathbf{T}}$ by minimizing $E\left(R,\hat{\mathbf{T}}\right)$.

Unfortunately, $E\left(R,\hat{\mathbf{T}}\right)$ has a complex dependence on its arguments [26] and is not useful for minimization. Thus, researchers have proposed algorithms that minimize approximations to $E\left(R,\hat{\mathbf{T}}\right)$. The most popular approximation [28], which we call the *WC* approximation, is

$$E_{\mathrm{WC}}\left(R,\hat{\mathbf{T}}\right) \equiv \sum_{m=1}^{N_p} \frac{\left(\bar{\mathbf{p}}_{1m}^T F \bar{\mathbf{p}}_{0m}\right)^2}{\left|\hat{\mathbf{z}}\times F\bar{\mathbf{p}}_{0m}\right|^2 + \left|\hat{\mathbf{z}}\times F^T\bar{\mathbf{p}}_{1m}\right|^2}. \tag{3}$$

[28] showed that the *WC* approximation becomes exact when $\hat{\mathbf{T}}$ is sideways *and* R is around the z axis, that is, $\mathbf{T}\cdot\hat{\mathbf{z}} = 0$ and $R\hat{\mathbf{z}} = \hat{\mathbf{z}}$. However, for roughly forward estimates of $\hat{\mathbf{T}}$, it is known that E_{WC} gives a poor approximation to E.

One can think of E_{WC} as a first–order approximation to $E\left(R,\hat{\mathbf{T}}\right)$. The Appendix rederives E_{WC} from this point of view and also presents a new, second–order approximation to $E\left(R,\hat{\mathbf{T}}\right)$ which does better than E_{WC} experimentally.

[1] We must use the infimum rather than the minimum since, as mentioned above, the infimum of E_{LS} may occur at an "infinite value" of $\mathbf{P}_m$, or at $\mathbf{P}_m$ equal to one of the camera positions. For simplicity, we sometimes gloss over this distinction below. One *can* take the minimum over R or the translation direction $\hat{\mathbf{T}}$, since these variables live in a compact space.

2.2 The Directional Error

As noted in Section 1, our approach starts from a least–squares error that differs from the standard one in (1). Define the length–3 *unit* vectors $\hat{\mathbf{p}}_{0m} \equiv \overline{\mathbf{p}_{0m}} / \left|\overline{\mathbf{p}_{0m}}\right|$ and $\hat{\mathbf{p}}_{1m} \equiv \overline{\mathbf{p}_{1m}} / \left|\overline{\mathbf{p}_{1m}}\right|$. Rather than defining the least–squares error in the *image plane*, as in (1), we consider the *directional* least–squares error[2]

$$E_{\theta LS}\left(R, \mathbf{T}, \{\mathbf{P}\}\right) \equiv \sum_{m=1}^{N_p} \left(\left| \frac{\mathbf{P}_m}{|\mathbf{P}_m|} \times \hat{\mathbf{p}}_{0m} \right|^2 + \left| \frac{R\left(\mathbf{P}_m - \mathbf{T}\right)}{\left|R\left(\mathbf{P}_m - \mathbf{T}\right)\right|} \times \hat{\mathbf{p}}_{1m} \right|^2 \right) = \sum_{m=1}^{N_p} \left(\sin^2\theta_{0m} + \sin^2\theta_{1m} \right), \tag{4}$$

where θ_{im} is the angle between the expected and observed ray directions for the mth point and ith image. This gives the natural form of the least–squares error on a spherical imaging surface (e.g., [10]), and researchers have tended to use it for omnidirectional cameras or cameras with wide field of view (FOV). We stress that the directional error is useful even for standard cameras with planar imaging surfaces. Because this error is rotationally invariant, it embodies a *better* noise model for physical lenses and the 3D world than the standard error does (see [14] for discussion). With this new starting point, we can minimize the error over the structure exactly, and the result, unlike the *WC* error, represents a physically meaningful error function, with no artifacts stemming from approximations.

2.3 Minimizing $E_{\theta LS}$ with respect to the structure

This section computes an explicit expression for the result of minimizing $E_{\theta LS}$ over the structure. Define the measurement errors

$$\boldsymbol{\delta}_{\theta 0m} \equiv \left(\mathbf{P}_m / |\mathbf{P}_m|\right) \times \hat{\mathbf{p}}_{0m}, \quad \boldsymbol{\delta}_{\theta 1m} \equiv \left(R\left(\mathbf{P}_m - \mathbf{T}\right) / \left|R\left(\mathbf{P}_m - \mathbf{T}\right)\right|\right) \times \hat{\mathbf{p}}_{1m}, \quad \boldsymbol{\Delta}_{\theta m} \equiv \left(\boldsymbol{\delta}_{\theta 0m}; \boldsymbol{\delta}_{\theta 1m}\right), \tag{5}$$

where $\boldsymbol{\Delta}_{\theta m} \in \Re^6$. We write the directional error in terms of these quantities as

$$E_{\theta LS}\left(R, \mathbf{T}, \{\mathbf{P}\}\right) \equiv \sum_{m=1}^{N_p} \left|\boldsymbol{\Delta}_{\theta m}\right|^2,$$

[2] Instead of (4), one could also use $E'_{\theta LS} \equiv \sum_{m=1}^{N_p} w_m \left(\sin^2\theta_{0m} + \sin^2\theta_{1m}\right)$, with a different weighting w_m for each 3D point, without changing any of the subsequent analysis.

where our notation indicates that $E_{\theta LS}$ depends on all the $\mathbf{P}_m$. Let $E_{\theta m}\left(R,\hat{\mathbf{T}}\right) \equiv \inf_{\mathbf{P}_m} |\mathbf{\Delta}_{\theta m}|^2$ and define[1]

$$E_\theta\left(R,\hat{\mathbf{T}}\right) \equiv \inf_{\mathbf{P}_1 \ldots \mathbf{P}_{Np}} E_{\theta LS} = \sum_{m=1}^{N_p} \inf_{\mathbf{P}_m} |\mathbf{\Delta}_{\theta m}|^2 = \sum_{m=1}^{N_p} E_{\theta m}\left(R,\hat{\mathbf{T}}\right),$$

where $E_\theta\left(R,\hat{\mathbf{T}}\right)$ is the analog of $E\left(R,\hat{\mathbf{T}}\right)$ for the image–plane error. We will solve explicitly for $E_{\theta m}\left(R,\hat{\mathbf{T}}\right)$.

From the rotational invariance of $\delta_{\theta 1m}$, one can rewrite it as $\delta_{\theta 1m} = \left(\left(\mathbf{P}_m - \hat{\mathbf{T}}\right) / \left|\mathbf{P}_m - \hat{\mathbf{T}}\right|\right) \times R^{-1}\hat{\mathbf{p}}_{1m}$. Define the *unrotated point* in image 1 by $\hat{\mathbf{p}}^U_{1m} \equiv R^{-1}\hat{\mathbf{p}}_{1m}$. Then

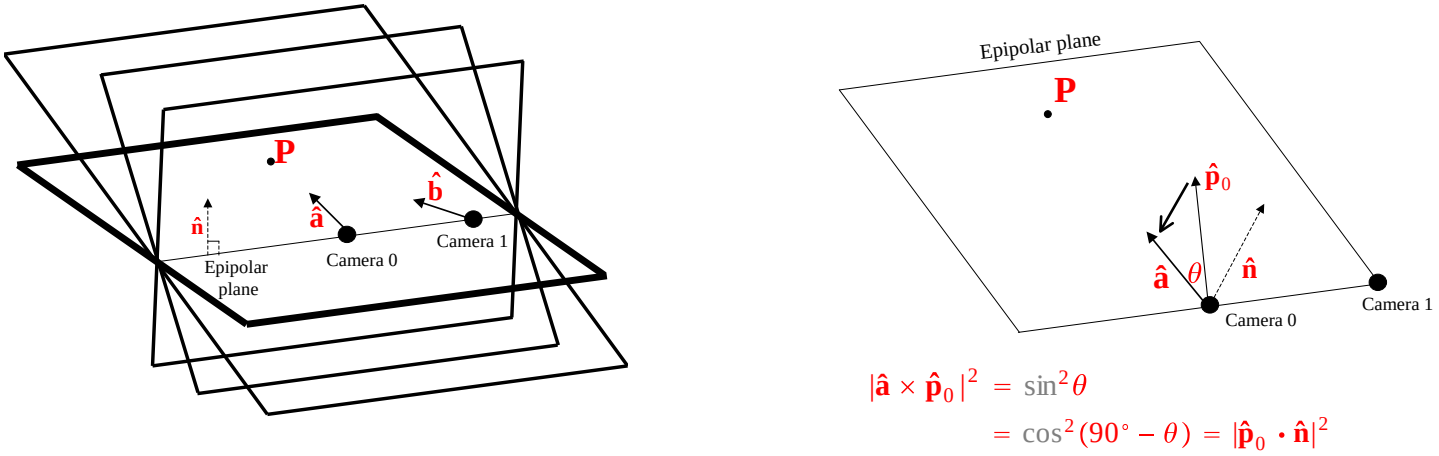
$$\delta_{\theta 1m} = \left(\left(\mathbf{P}_m - \hat{\mathbf{T}}\right) / \left|\mathbf{P}_m - \hat{\mathbf{T}}\right|\right) \times \hat{\mathbf{p}}^U_{1m}.$$

With $\delta_{\theta 1m}$ written this way, the quantities $\delta_{\theta 1m}$, $\delta_{\theta 0m}$, and $\mathbf{\Delta}_{\theta m}$ have the same functional dependence on $\mathbf{P}_m$ regardless of whether the rotation is zero or nonzero: *the rotation has no effect on the minimization over* $\mathbf{P}_m$. The only change for nonzero rotation is that one represents image 1 by the unrotated data $\hat{\mathbf{p}}^U_{1m}$ rather than the original data $\hat{\mathbf{p}}_{1m}$. This simplification is our main reason for using the directional error.

We minimize $|\Delta_{\theta m}|^2$ over $\mathbf{P}_m$ in two steps. First, we fix an epipolar plane passing through the two camera positions (Figure 1a), and we minimize over all $\mathbf{P}_m$ (except $\tilde{\mathbf{0}}$ and $\hat{\mathbf{T}}$) on this plane. Second, we minimize the result of the first minimization over all choices of the epipolar plane. The combination of the two steps clearly gives the same result as minimizing over all $\mathbf{P}_m$. We follow the standard practice of neglecting the constraint that the depths are positive.

The rationale for our two–step approach is the same as in [5]. If one first restricts the minimization over $\mathbf{P}_m$ to an epipolar plane, one can replace this minimization by simpler, *independent* minimizations over the *images* of $\mathbf{P}_m$ in the two cameras, where the images and corresponding rays must also lie in the epipolar plane. This is because, no matter what independent choices we make for the two image rays, they must intersect since they are coplanar, and thus they correspond to a valid choice for $\mathbf{P}_m$. The image rays that give the minimum are the rays in the epipolar plane that lie closest to the image data $\hat{\mathbf{p}}_{0m}$ and $\hat{\mathbf{p}}_{1m}$, and in general they are given by projecting the image data perpendicularly into the epipolar plane (Figure 1b). With these minimizing values, the quantity $|\mathbf{\Delta}_{\theta m}|^2$ becomes a simple function of the

epipolar plane, and the second step that minimizes over the choice of this plane is straightforward. We give detailed derivation below and in the Appendix, which explains how to deal with exceptional cases.



(a) (b)

Figure 1: (a): The minimization over $\mathbf{P}$ is divided in two steps. Step 1: choose an epipolar plane passing through the camera positions, characterized by its normal $\hat{\mathbf{n}}$, and minimize over all $\mathbf{P}$ on this plane. This is equivalent to minimizing *independently* over all images $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ of $\mathbf{P}$ within the epipolar plane. Step 2: minimize the result over all choices of the plane. (b): The optimal estimate of $\hat{\mathbf{a}}$, the image of $\mathbf{P}$ in image 0, after fixing the epipolar plane. It corresponds to the result of minimizing over all $\mathbf{P}$ on this plane. The optimal estimate of $\hat{\mathbf{a}}$ generally equals the perpendicular projection of the measured image ray into the epipolar plane, and it gives $\delta_0^2 = |\hat{\mathbf{a}} \times \hat{\mathbf{p}}_0|^2 = |\hat{\mathbf{n}} \cdot \hat{\mathbf{p}}_0|^2$.

Minimizing over $\mathbf{P}_m$. Fix an epipolar plane with normal $\hat{\mathbf{n}}_m$. Since the plane passes through the camera positions, we have $\hat{\mathbf{n}}_m \cdot \hat{\mathbf{T}} = 0$. Define the unit vectors

$$\hat{\mathbf{a}}_m \equiv \mathbf{P}_m / |\mathbf{P}_m|, \quad \hat{\mathbf{b}}_m \equiv \left(\mathbf{P}_m - \hat{\mathbf{T}}\right) / \left|\mathbf{P}_m - \hat{\mathbf{T}}\right|. \tag{6}$$

These are the ray directions to $\mathbf{P}_m$ in images 0 and 1, and one can think of them as the images of $\mathbf{P}_m$. Since $\mathbf{P}_m \cdot \hat{\mathbf{n}}_m = \hat{\mathbf{T}} \cdot \hat{\mathbf{n}}_m = 0$, clearly $\hat{\mathbf{a}}_m \cdot \hat{\mathbf{n}}_m = 0$ and $\hat{\mathbf{b}}_m \cdot \hat{\mathbf{n}}_m = 0$. Both $\hat{\mathbf{a}}_m$ and $\hat{\mathbf{b}}_m$ lie on the unit circle in the epipolar plane. We refer to this as the *epipolar circle*: it is the equivalent on a spherical

imaging surface of the epipolar line in the standard image plane. One can write the error contribution from the mth point in terms of $\hat{\mathbf{a}}_m$ and $\hat{\mathbf{b}}_m$ as

$$|\mathbf{\Delta}_{\theta m}|^2 = |\hat{\mathbf{a}}_m \times \hat{\mathbf{p}}_{0m}|^2 + \left|\hat{\mathbf{b}}_m \times \hat{\mathbf{p}}_{1m}^U\right|^2 . \tag{7}$$

Note that the error $|\mathbf{\Delta}_{\theta m}|^2$ does not change its value if one takes $\hat{\mathbf{a}}_m \longrightarrow -\hat{\mathbf{a}}_m$ or $\hat{\mathbf{b}}_m \longrightarrow -\hat{\mathbf{b}}_m$.

Claim 1: *One can replace the minimization of $|\mathbf{\Delta}_{\theta m}|^2$ over all $\mathbf{P}_m$ on the epipolar plane by **independent** minimizations*[3] *over all $\hat{\mathbf{a}}_m$ and $\hat{\mathbf{b}}_m$ on the epipolar circle*:

$$\inf_{\mathbf{P}_m} |\Delta_{\theta m}|^2 \Big|_{\mathbf{P}_m \cdot \hat{\mathbf{n}}_m = 0} = \min_{\hat{\mathbf{a}}} |\hat{\mathbf{a}} \times \hat{\mathbf{p}}_{0m}|^2 \Big|_{\hat{\mathbf{a}} \cdot \hat{\mathbf{n}} = 0, |\hat{\mathbf{a}}| = 1} + \min_{\hat{\mathbf{b}}} \left|\hat{\mathbf{b}} \times \hat{\mathbf{p}}_{1m}^U\right|^2 \Big|_{\hat{\mathbf{b}} \cdot \hat{\mathbf{n}} = 0, |\hat{\mathbf{b}}| = 1} . \tag{8}$$

Proof: see Appendix.

The minimizations over $\hat{\mathbf{a}}_m$ and $\hat{\mathbf{b}}_m$ in (8) are trivial (Figure 1b) and give

$$\inf_{\mathbf{P}_m} |\Delta_{\theta m}|^2 \Big|_{\mathbf{P}_m \cdot \hat{\mathbf{n}}_m = 0} = |\hat{\mathbf{n}}_m \cdot \hat{\mathbf{p}}_{0m}|^2 + \left|\hat{\mathbf{n}}_m \cdot \hat{\mathbf{p}}_{1m}^U\right|^2 , \tag{9}$$

which we write as

$$\inf_{\mathbf{P}_m} |\Delta_{\theta m}|^2 \Big|_{\mathbf{P}_m \cdot \hat{\mathbf{n}}_m = 0} = \hat{\mathbf{n}}_m^T S_{\theta m} \hat{\mathbf{n}}_m, \qquad S_{\theta m} \equiv \hat{\mathbf{p}}_{0m} \hat{\mathbf{p}}_{0m}^T + \hat{\mathbf{p}}_{1m}^U \hat{\mathbf{p}}_{1m}^{UT}. \tag{10}$$

Note that (9) and (10) remain true in the exceptional cases $\hat{\mathbf{p}}_{0m} \parallel \hat{\mathbf{n}}_m$ or $\hat{\mathbf{p}}_{1m}^U \parallel \hat{\mathbf{n}}_m$, where the image data has zero projection on the epipolar plane.

From the previous arguments, we have

$$E_{\theta m} \equiv \inf_{\hat{\mathbf{n}}} \hat{\mathbf{n}}^T S_{\theta m} \hat{\mathbf{n}} \Big|_{\hat{\mathbf{n}} \perp \hat{\mathbf{T}}, |\hat{\mathbf{n}}| = 1} . \tag{11}$$

Since $\hat{\mathbf{n}} \cdot \hat{\mathbf{T}} = 0$, one only needs to know how $S_{\theta m}$ acts on the 2D space perpendicular to $\hat{\mathbf{T}}$. In effect, $S_{\theta m}$ becomes a 2×2 matrix. The least eigenvalue of this matrix gives $E_{\theta m}$, and the corresponding eigenvector gives the optimal estimate of $\hat{\mathbf{n}}$. The eigenvalues of a 2×2 matrix are trivial to compute, and the result is:

$$E_{\theta m} = A_{\theta m}/2 - \sqrt{A_{\theta m}^2/4 - B_{\theta m}}, \tag{12}$$

$$A_{\theta m} \equiv \hat{\mathbf{p}}_{0m}^T \left(1 - \hat{\mathbf{T}}\hat{\mathbf{T}}^T\right) \hat{\mathbf{p}}_{0m} + \hat{\mathbf{p}}_{1m}^T R \left(1 - \hat{\mathbf{T}}\hat{\mathbf{T}}^T\right) R^{-1} \hat{\mathbf{p}}_{1m}, \qquad B_{\theta m} \equiv \left(\hat{\mathbf{T}} \cdot \hat{\mathbf{p}}_{0m} \times R^{-1} \hat{\mathbf{p}}_{1m}\right)^2 .$$

[3] One take the minimum over $\hat{\mathbf{a}}$ or $\hat{\mathbf{b}}$ since these live in compact spaces; the infimum is unnecessary.

One can also write the $A_{\theta m}$ and $B_{\theta m}$ in terms of the essential matrix F (see Section 1.1 for its definition): $A_{\theta m} = \hat{\mathbf{p}}_{0m}^T F^T F \hat{\mathbf{p}}_{0m} + \hat{\mathbf{p}}_{1m}^T F F^T \hat{\mathbf{p}}_{1m}$ and $B_{\theta m} \equiv \left(\hat{\mathbf{p}}_{1m}^T F \hat{\mathbf{p}}_{0m}\right)^2$. We define

$$E_{\theta,\mathrm{ex}}\left(R, \hat{\mathbf{T}}\right) \equiv \sum_{m=1}^{N_p} \left(A_{\theta m}/2 - \sqrt{A_{\theta m}^2/4 - B_{\theta m}} \right) = E_\theta\left(R, \hat{\mathbf{T}}\right). \tag{13}$$

One can also apply this result for robust estimation. For instance, one could apply (12) within RANSAC [3], or modify (12) to get a robust error function that flattens for larger values of the error.

3 Optimal Structure Estimate

By minimizing $|\Delta_{\theta m}|^2$ over $\mathbf{P}_m$, we are implicitly computing the optimal estimate of the structure given the motion. In this section, we give an exact, *explicit* expression for the optimal structure estimate for our directional error. Our expression just requires computing square roots. The previous result of [5] for the image–plane error required solving a 6th–degree polynomial. In addition, we identify a configuration of the estimated epipole $\mathbf{e}$ and image observations $\mathbf{p}_{0m}$, $\mathbf{p}_{1m}^U$ that produces a one–parameter ambiguity in recovering the structure. This ambiguity also occurs for the standard image–plane least–squares error, though only for some motions. For convenience, we drop the subscript m in this section and, without loss of generality (due to the scale ambiguity of SFM), we take $|\mathbf{T}| = 1$.

Our discussion is complicated by the need to consider special cases. Section 3.1 begins with the special and trivial case when at least one of the image rays $\hat{\mathbf{p}}_1^U$ and $\hat{\mathbf{p}}_0$ is parallel to $\mathbf{T}$. Section 3.2 describes the typical case when neither $\hat{\mathbf{p}}_1^U$ nor $\hat{\mathbf{p}}_0$ is parallel to $\mathbf{T}$.

3.1 Image rays parallel to $\mathbf{T}$

1): If both $\hat{\mathbf{p}}_1^U$ and $\hat{\mathbf{p}}_0$ are parallel to $\mathbf{T}$, we clearly have $\mathbf{P} = \lambda \mathbf{T}$, where λ is an undetermined constant. 2): Next, assume that $\hat{\mathbf{p}}_0$ is parallel to $\mathbf{T}$ but $\hat{\mathbf{p}}_1^U$ is not parallel to $\mathbf{T}$. Let $\hat{\mathbf{p}}_1^U = \mathbf{T} + \mathbf{v}$, where, by assumption, $\mathbf{v}$ is nonzero and not parallel to $\mathbf{T}$. Consider a point on the ray passing though $\mathbf{T}$ (the position of the second camera) in the direction $\hat{\mathbf{p}}_1^U$, which we write as $\mathbf{P}_\varepsilon = (\mathbf{T} + \varepsilon \mathbf{v}) / (1 - \varepsilon)$, where $\varepsilon > 0$ is a small constant. The error $\delta_{\theta 1}$ for image 1 is zero for all points on this ray, since $\mathbf{P}_\varepsilon - \mathbf{T} = \varepsilon (\mathbf{T} + \mathbf{v}) / (1 - \varepsilon) = \varepsilon \hat{\mathbf{p}}_1^U / (1 - \varepsilon)$, which implies that

$$\left|\delta_{\theta 1}\left(\mathbf{P}_\varepsilon\right)\right|^2 = \left| \left((\mathbf{P}_\varepsilon - \mathbf{T}) / |\mathbf{P}_\varepsilon - \mathbf{T}| \right) \times \hat{\mathbf{p}}_1^U \right|^2 = 0.$$

By choosing ε small enough, we can make $\mathbf{P}_\varepsilon$ and $\mathbf{P}_\varepsilon / |\mathbf{P}_\varepsilon|$ arbitrarily close to $\mathbf{T}$ and thus make $|\delta_{\theta 0}(\mathbf{P}_\varepsilon)|^2$, the error in image 0, arbitrarily small. The total error $|\Delta_\theta(\mathbf{P}_\varepsilon)|^2 \equiv |\delta_{\theta 0}(\mathbf{P}_\varepsilon)|^2 + |\delta_{\theta 1}(\mathbf{P}_\varepsilon)|^2$ goes to zero in the limit $\varepsilon \longrightarrow 0$. We take our estimate of the structure as

$$\lim_{\varepsilon \longrightarrow 0} (\mathbf{T} + \varepsilon \mathbf{v}) / (1 - \varepsilon) = \mathbf{T},$$

since this corresponds to zero error.

By symmetry, the case when $\hat{\mathbf{p}}_1^U$ is parallel to $\mathbf{T}$ but $\hat{\mathbf{p}}_0$ is not parallel to $\mathbf{T}$ reduces to the case just considered. The same arguments show that $\mathbf{P} = (0; 0; 0)$.

3.2 Image rays nonparallel to $\mathbf{T}$

We assume neither $\hat{\mathbf{p}}_0$ nor $\hat{\mathbf{p}}_1^U$ is parallel to $\mathbf{T}$. We compute the optimal estimate of $\mathbf{P}$ in three stages. First, we compute the optimal estimate of the epipolar plane. Second, we use this estimate to compute the optimal estimates within the epipolar plane of the two images of $\mathbf{P}$, namely, $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$. Last, we compute the optimal structure estimate by standard triangulation, by computing the intersection of the two image rays. Without loss of generality, we adopt a coordinate system where $\mathbf{T} = \hat{\mathbf{z}}$, and $\hat{p}_{0,x} > 0$ and $\hat{p}_{0,y} = 0$, see Figure 2.

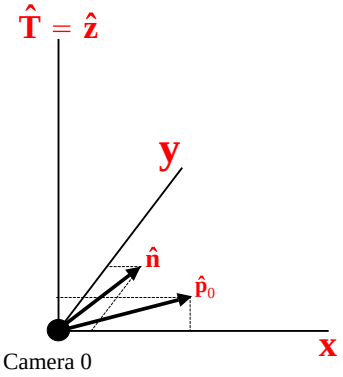


Figure 2: The coordinate system used in Section 3.2, with $\hat{\mathbf{T}} = \hat{\mathbf{z}}$, $\hat{\mathbf{p}}_0 = (p_{0x}; 0; p_{0z})$, and $\hat{\mathbf{n}} = (n_x; n_y; 0)$.

We must again consider several special cases, one of which gives an ambiguity in estimating the

epipolar plane and the structure. We first compute the optimal structure estimate for the generic, non–ambiguous case and then return in Section 3.2.1 to the special ambiguous case.

Determining the optimal epipolar plane. As in Section 2.3, we represent the epipolar plane by its surface normal $\hat{\mathbf{n}}$. Since $\hat{\mathbf{n}} \cdot \mathbf{T} = \mathbf{0}$, the unit normal $\hat{\mathbf{n}}$ lies in the xy plane in our coordinate system, and we represent it as a length–2 vector in this plane, see Figure 2. Let $[[S_\theta]]_2$ be the 2×2 submatrix of S_θ consisting of the x and y entries of S_θ, where we defined S_θ in (10). Since $\hat{p}_{0,y} = 0$ in our coordinate system, we have

$$[[S_\theta]]_2 = \begin{pmatrix} \hat{p}_{0,x}^2 + \left(\hat{p}_{1,x}^U\right)^2 & \hat{p}_{1,x}^U \hat{p}_{1,y}^U \\ \hat{p}_{1,x}^U \hat{p}_{1,y}^U & \left(\hat{p}_{1,y}^U\right)^2 \end{pmatrix}. \tag{14}$$

As discussed following (11), the eigenvector of $[[S_\theta]]_2$ corresponding to the least eigenvalue gives the optimal estimate of $\hat{\mathbf{n}}$. If $[[S_\theta]]_2$ has a unique least eigenvalue, the estimate of $\hat{\mathbf{n}}$ is uniquely determined (up to sign), but if $[[S_\theta]]_2$ has two eigenvalues of equal size, $\hat{\mathbf{n}}$ can lie in *any* direction in the x–y plane.

Optimal estimate of the epipolar plane: generic case. If $[[S_\theta]]_2$ has a unique least eigenvalue, it is straightforward to show that the unique optimal estimate of $\hat{\mathbf{n}}$ (up to sign) is

$$\hat{\mathbf{n}} = \kappa^{-1} \begin{pmatrix} \hat{p}_{0,x}^2 + \left(\hat{p}_{1,x}^U\right)^2 - \left(\hat{p}_{1,y}^U\right)^2 - G \\ 2\hat{p}_{1,x}^U \hat{p}_{1,y}^U \end{pmatrix}, \tag{15}$$

where

$$\begin{aligned} G &\equiv \sqrt{\left(\hat{p}_{0,x}^2 + \left(\hat{p}_{1,x}^U\right)^2 - \left(\hat{p}_{1,y}^U\right)^2\right)^2 + 4\left(\hat{p}_{1,x}^U\right)^2 \left(\hat{p}_{1,y}^U\right)^2} \\ &= \sqrt{\left(\hat{p}_{0,x}^2 - \left(\hat{p}_{1,x}^U\right)^2 - \left(\hat{p}_{1,y}^U\right)^2\right)^2 + 4\left(\hat{p}_{0,x}\right)^2 \left(\hat{p}_{1,x}^U\right)^2}, \end{aligned} \tag{16}$$

and κ is a normalizing factor that enforces $|\hat{\mathbf{n}}| = 1$. Explicitly, we have

$$\kappa^2 = 2G\left(G - \left(\hat{p}_{0,x}^2 + \left(\hat{p}_{1,x}^U\right)^2 - \left(\hat{p}_{1,y}^U\right)^2\right)\right). \tag{17}$$

Though the expression for $\hat{\mathbf{n}}$ in (15) is not defined for $\kappa = 0$, we show in the Appendix

Claim 2. *Assume the matrix $[[S_\theta]]_2$ has a unique least eigenvalue, so $\hat{\mathbf{n}}$ is uniquely determined (up to sign). Then Equation (15) gives the correct results at $\kappa = 0$ after taking the appropriate limits.*

Optimal estimates of $\hat{\mathbf{a}}$, $\hat{\mathbf{b}}$ (generic case). Given the unique optimal estimate of $\hat{\mathbf{n}}$ in (15), we compute the optimal estimates of the images of $\mathbf{P}$, namely $\hat{\mathbf{a}} \equiv \mathbf{P}/|\mathbf{P}|$ and $\hat{\mathbf{b}} \equiv (\mathbf{P}-\mathbf{T})/|\mathbf{P}-\mathbf{T}|$. As discussed at the beginning of Section 2.3, see also Figure 1b, the optimal estimates of $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ generally equal the projections of the image data into the epipolar plane. By definition, they are the values that minimize the error $|\mathbf{\Delta}|^2$ in (5) and (7). If neither $\hat{\mathbf{p}}_0$ nor $\hat{\mathbf{p}}_1^U$ is parallel to $\hat{\mathbf{n}}$, one can easily show from (7) that the minima occur at the projections

$$\hat{\mathbf{a}} = \pm\frac{\hat{\mathbf{p}}_0 - \hat{\mathbf{n}}\,(\hat{\mathbf{n}}\cdot\hat{\mathbf{p}}_0)}{|\hat{\mathbf{p}}_0 - \hat{\mathbf{n}}\,(\hat{\mathbf{n}}\cdot\hat{\mathbf{p}}_0)|}, \quad \hat{\mathbf{b}} = \pm\frac{\hat{\mathbf{p}}_1^U - \hat{\mathbf{n}}\left(\hat{\mathbf{n}}\cdot\hat{\mathbf{p}}_1^U\right)}{\left|\hat{\mathbf{p}}_1^U - \hat{\mathbf{n}}\left(\hat{\mathbf{n}}\cdot\hat{\mathbf{p}}_1^U\right)\right|}. \tag{18}$$

In fact, we show in the Appendix:

Claim 3. *Assume the matrix $[[S_\theta]]_2$ has a unique least eigenvalue, so $\hat{\mathbf{n}}$ is uniquely determined (up to sign). Then, (18) always holds.*

The optimal estimate of $\hat{\mathbf{P}}$ (generic case). Given the optimal estimates for $\hat{\mathbf{a}}$, $\hat{\mathbf{b}}$ in (18), we compute the optimal estimate of $\mathbf{P}$ by standard triangulation:

$$\mathbf{P} = \lambda_0 \left(\begin{pmatrix} 0 \\ 0 \\ \hat{p}_{0,z} \end{pmatrix} + \frac{\hat{p}_{0,x}}{2G} \begin{pmatrix} \hat{p}_{0,x}^2 + \left(\hat{p}_{1,x}^U\right)^2 - \left(\hat{p}_{1,y}^U\right)^2 + G \\ 2\hat{p}_{1,x}^U \hat{p}_{1,y}^U \\ 0 \end{pmatrix} \right), \tag{19}$$

where G and λ_0 are given in (16) and (35), (34). Recall that this form assumes $|\mathbf{T}| = 1$. Our detailed derivation in the Appendix shows that the result (19) also holds in the limit $\hat{p}_{1y}^U \longrightarrow 0$.

Coordinate–free expression for $\mathbf{P}$. One can rewrite (19) in a covariant form, independent of our coordinate system, by replacing $\hat{\mathbf{z}} \longrightarrow \hat{\mathbf{T}}$,

$$\hat{\mathbf{x}} \longrightarrow \left(\hat{\mathbf{p}}_0 - \hat{\mathbf{T}}\left(\hat{\mathbf{T}}\cdot\hat{\mathbf{p}}_0\right)\right) / \left|\hat{\mathbf{p}}_0 - \hat{\mathbf{T}}\left(\hat{\mathbf{T}}\cdot\hat{\mathbf{p}}_0\right)\right|, \quad \hat{\mathbf{y}} \longrightarrow \left(\hat{\mathbf{T}}\times\hat{\mathbf{p}}_0\right) / \left|\hat{\mathbf{T}}\times\hat{\mathbf{p}}_0\right|,$$

which gives, for instance, $\hat{p}_{0,z} \longrightarrow \hat{\mathbf{T}}\cdot\hat{\mathbf{p}}_0$ and $\hat{p}_{0,x} \longrightarrow \sqrt{1-\left(\hat{\mathbf{T}}\cdot\hat{\mathbf{p}}_0\right)^2}$ and

$$\hat{p}_{1,x}^U \longrightarrow \left(\hat{\mathbf{T}}\times\hat{\mathbf{p}}_0\right)\cdot\left(\hat{\mathbf{T}}\times\hat{\mathbf{p}}_1^U\right) / \sqrt{1-\left(\hat{\mathbf{T}}\cdot\hat{\mathbf{p}}_0\right)^2}.$$

3.2.1 The ambiguous case: non–unique $\hat{\mathbf{n}}$

We now analyze the case when $\hat{\mathbf{n}}$ is not unique. As discussed following (14), this occurs when the matrix $[[S_\theta]]_2$ has two equal eigenvalues. The eigenvalues of $[[S_\theta]]_2$ are the solutions to the characteristic equation

$$0 = \det\left([[S_\theta]]_2 - x\mathbf{1}_2\right) = x^2 - x\mathrm{Trace}\left([[S_\theta]]_2\right) + \det\left([[S_\theta]]_2\right),$$

where $\mathbf{1}_2$ is the 2×2 identity matrix. If $[[S_\theta]]_2$ has two eigenvalues of the same size ε, we have

$$x^2 - x\mathrm{Trace}\left([[S_\theta]]_2\right) + \det\left([[S_\theta]]_2\right) = (x - \varepsilon)^2,$$

which implies $\det\left([[S_\theta]]_2\right) = \varepsilon^2$, $\mathrm{Trace}\left([[S_\theta]]_2\right) = 2\varepsilon$, and

$$4\det\left([[S_\theta]]_2\right) = \left(\mathrm{Trace}\left([[S_\theta]]_2\right)\right)^2. \tag{20}$$

When (20) does not hold, the $\hat{\mathbf{n}}$–estimate is unique (up to sign). When the condition (20) holds, $[[S_\theta]]_2$ has equal eigenvalues and $\hat{\mathbf{n}}$ is ambiguous.

We rewrite the condition (20) in terms of the image measurements. From the expression for $[[S_\theta]]_2$ in (14),

$$\mathrm{Trace}\left([[S_\theta]]_2\right) = \hat{p}_{0,x}^2 + \left(\hat{p}_{1,x}^U\right)^2 + \left(\hat{p}_{1,y}^U\right)^2, \qquad \det\left([[S_\theta]]_2\right) = \hat{p}_{0,x}^2\left(\hat{p}_{1,y}^U\right)^2.$$

After we substitute these expressions and take the square root, the condition (20) becomes

$$2\left|\hat{p}_{0,x}\hat{p}_{1,y}^U\right| = \hat{p}_{0,x}^2 + \left(\hat{p}_{1,x}^U\right)^2 + \left(\hat{p}_{1,y}^U\right)^2,$$

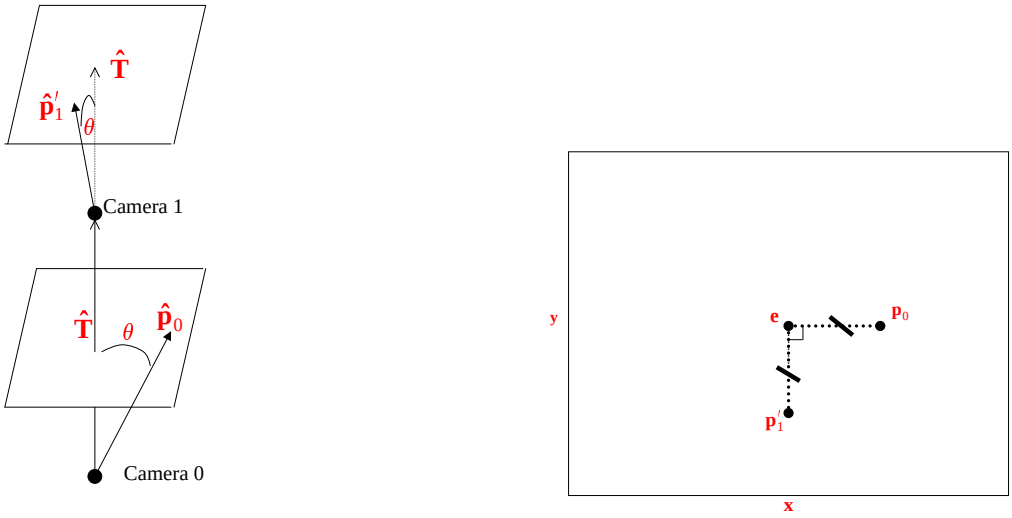
which implies $0 = \left(\hat{p}_{1,x}^U\right)^2 + \left(\left|\hat{p}_{1,y}^U\right| - \left|\hat{p}_{0,x}\right|\right)^2$. Thus, $\hat{\mathbf{n}}$ is uniquely determined (up to sign) unless

$$\hat{p}_{1,x}^U = 0 = \left|\hat{p}_{1,y}^U\right| - \left|\hat{p}_{0,x}\right|, \tag{21}$$

which is the form of the condition (20) that we use below. See Figure 3. One can also write the ambiguity condition (21) in a form that doesn't depend on the coordinate system as $\hat{\mathbf{T}} \times \hat{\mathbf{p}}_0 = \pm\hat{\mathbf{T}} \times \left(\hat{\mathbf{T}} \times \hat{\mathbf{p}}_1^U\right)$.

.

The optimal structure estimate (ambiguous case). When the ambiguity condition (21) holds, one can choose $\hat{\mathbf{n}}$ in any direction in the x–y plane. The estimates of $\hat{\mathbf{a}}$, $\hat{\mathbf{b}}$ in (18) remain valid unless $\hat{p}_{0,x}^2 =$



(a) (b)

Figure 3: The ambiguous configuration for triangulation. In (b), the epipole and image points have been projected into the image plane for camera 0 using the coordinate system of Figure 2b and Section 3.2. The projections form a right isosceles triangle.

$\left(\hat{p}^U_{1,y}\right)^2 = 1$. Neglecting temporarily this exceptional case, the standard triangulation computation gives (see the Appendix for details)

$$\mathbf{P} = \hat{p}^{-1}_{0,z} \left(n_x / \left(n_y s_1 s_2 + n_x\right)\right) \left(n_y^2 \hat{p}_{0,x}; -n_x n_y \hat{p}_{0,x}; \hat{p}_{0,z}\right), \tag{22}$$

where $\hat{\mathbf{n}} = (n_x; n_y; 0)$ and the factors $s_1 \equiv \hat{p}_{0,x}/\hat{p}^U_{1,y}$ and $s_2 \equiv \hat{p}^U_{1,z}/\hat{p}_{0,z}$ equal 1 or -1. Since $n_x^2 + n_y^2 = 1$, there is a one–parameter ambiguity in determining $\mathbf{P}$.

As one varies the normal $\hat{\mathbf{n}}$, the estimate $\mathbf{P}$ in (22) traces out a curve known as a *twisted cubic* [4]. One can write this curve in a more transparent form in our coordinate system as

$$\begin{aligned} X &= \left(\frac{\hat{p}_{0,x}}{\hat{p}_{0,z}}\right) \frac{Z\left(Z-1\right)^2}{2Z^2 - 2Z + 1}, \\ Y &= s_1 s_2 \left(\frac{\hat{p}_{0,x}}{\hat{p}_{0,z}}\right) \frac{Z^2\left(Z-1\right)}{2Z^2 - 2Z + 1}. \end{aligned} \tag{23}$$

Last, we consider the exceptional case when $\hat{p}^U_{1,x} = 0$ and $\hat{p}^2_{0,x} = \left(\hat{p}^U_{1,y}\right)^2 = 1$. One can describe this configuration more simply by noting that the unit rays $\hat{\mathbf{p}}_0$, $\hat{\mathbf{p}}^U_1$, and $\hat{\mathbf{T}}$ form an orthogonal system, with $\hat{\mathbf{p}}_0 = \hat{\mathbf{x}}$, $\hat{\mathbf{p}}^U_1 = \pm\hat{\mathbf{y}}$, and $\hat{\mathbf{T}} = \hat{\mathbf{z}}$ in our coordinate system. As the Appendix shows, the optimal estimate of $\mathbf{P}$ can lie anywhere on the twisted cubic of (23) or, in addition, anywhere on the lines $(X; 0; 1)$ or $(0; Y; 1)$. One can describe these additional lines without reference to a coordinate system as, respectively,

$$\mathbf{T} \cdot \mathbf{P} = 1,\ \hat{\mathbf{p}}^U_1 \cdot \mathbf{P} = 0 \tag{24}$$

and

$$\mathbf{T} \cdot \mathbf{P} = 1,\ \hat{\mathbf{p}}_0 \cdot \mathbf{P} = 0. \tag{25}$$

3.3 Summary

- If neither $\hat{\mathbf{p}}^U_1$ nor $\hat{\mathbf{p}}_0$ is parallel to $\mathbf{T}$ and the ambiguity condition (21) does not hold, $\mathbf{P}$ is given by (19). This is the case normally encountered (the generic case).
- If both $\hat{\mathbf{p}}_0$ and $\hat{\mathbf{p}}^U_1$ are parallel to $\mathbf{T}$, we have $\mathbf{P} = \lambda\mathbf{T}$, where λ is an undetermined constant. There is a one–parameter ambiguity in determining $\mathbf{P}$.
- If $\hat{\mathbf{p}}_0 \parallel \mathbf{T}$ but $\hat{\mathbf{p}}^U_1$ is not parallel to $\mathbf{T}$, we have $\mathbf{P} = \mathbf{T}$. If $\hat{\mathbf{p}}^U_1$ is parallel to $\mathbf{T}$ but $\hat{\mathbf{p}}_0$ is not parallel to $\mathbf{T}$, we have $\mathbf{P} = (0; 0; 0)$.
- When neither $\hat{\mathbf{p}}^U_1$ nor $\hat{\mathbf{p}}_0$ is parallel to $\mathbf{T}$ and the ambiguity condition (21) holds, but $\hat{p}^2_{0,x} = \left(\hat{p}^U_{1,y}\right)^2 \neq 1$, (22) or (23) gives $\mathbf{P}$. There is a one–parameter ambiguity in determining $\mathbf{P}$.
- When neither $\hat{\mathbf{p}}^U_1$ nor $\hat{\mathbf{p}}_0$ is parallel to $\mathbf{T}$ and the ambiguity condition (21) holds and $\hat{p}^2_{0,x} = \left(\hat{p}^U_{1,y}\right)^2 = 1$, either (22), (24), or (25) gives $\mathbf{P}$. The structure $\mathbf{P}$ may occur on any one of three one–dimensional sets.

In coordinate–free notation, the ambiguity condition (21) holds when $\hat{\mathbf{T}} \times \hat{\mathbf{p}}_0 = \pm\hat{\mathbf{T}} \times \left(\hat{\mathbf{T}} \times \hat{\mathbf{p}}^U_1\right)$. Geometrically, it holds when the projections of the rays $\hat{\mathbf{p}}_0$, $\hat{\mathbf{T}}$, and $\hat{\mathbf{p}}^U_1$ into an image plane perpendicular to $\hat{\mathbf{T}}$ form a right isosceles triangle. See Figure 3. The coordinate–free conditions giving the last case, for which $\mathbf{P}$ may occur on any one of three one–dimensional sets, are $\hat{\mathbf{T}} \times \hat{\mathbf{p}}_0 = \pm\hat{\mathbf{T}} \times \left(\hat{\mathbf{T}} \times \hat{\mathbf{p}}^U_1\right)$ *and* $\hat{\mathbf{T}} \cdot \hat{\mathbf{p}}_0 = \hat{\mathbf{T}} \cdot \hat{\mathbf{p}}^U_1 = 0$.

Discussion. When does the ambiguous configuration matter in practice? Consider the image plane perpendicular to $\mathbf{T}$, which gives $\mathbf{e} = (0; 0)$ for the epipole estimate (as in Figure 3b). Let the 2D points

$\mathbf{p}_0$, $\mathbf{p}_1$, and $\mathbf{p}_1^U$ represent the projections of the measurements $\hat{\mathbf{p}}_0$, $\hat{\mathbf{p}}_1$, and $\hat{\mathbf{p}}_1^U = R^{-1}\hat{\mathbf{p}}_1$ in this plane. Let $\mathbf{e}_{\rm true}$ be the projection of the true translation $\mathbf{T}_{\rm true}$ into the image plane, and let the 2D points $\mathbf{p}_{0\rm true}$, $\mathbf{p}_{1\rm true}$, and $\mathbf{p}_{1\rm true}^U$ be the projections of the true rays $\hat{\mathbf{p}}_{0\rm true}$, $\hat{\mathbf{p}}_{1\rm true}^U$, and $\hat{\mathbf{p}}_{1\rm true}^U = R_{\rm true}^{-1}\hat{\mathbf{p}}_{1\rm true}$. Since the noise is invariably small, we have $\mathbf{p}_0 \approx \mathbf{p}_{0\rm true}$ and $\mathbf{p}_1 \approx \mathbf{p}_{1\rm true}$. From the coplanarity of $\hat{\mathbf{p}}_{0\rm true}$, $\hat{\mathbf{p}}_{1\rm true}^U$, and $\mathbf{T}_{\rm true}$, it follows that $\mathbf{e}_{\rm true}$, $\mathbf{p}_{0\rm true}$, and $\mathbf{p}_{1\rm true}^U$ lie on the same epipolar line.

Suppose that the motion estimates are near their true values, with $\mathbf{T} \approx \mathbf{T}_{\rm true}$ and $R \approx R_{\rm true}$. Then we have $\mathbf{p}_0 \approx \mathbf{p}_{0\rm true}$ (small noise), $\mathbf{p}_1^U \approx \mathbf{p}_{1\rm true}^U$ (small noise and $R \approx R_{\rm true}$), and $\mathbf{e} \approx \mathbf{e}_{\rm true}$, and the collinearity of $\mathbf{p}_{0\rm true}$, $\mathbf{p}_{1\rm true}^U$, and $\mathbf{e}_{\rm true}$ implies that the measured points $\mathbf{p}_0$ and $\mathbf{p}_1^U$ are collinear with $\mathbf{e} = (0;0)$ up to small corrections. However, if the points $\mathbf{p}_0$ and $\mathbf{p}_1^U$ lie in the ambiguous configuration, they have equal magnitude and are *perpendicular*. One can resolve this contradiction only if the distance from $\mathbf{e}$ to the measured points is of order the noise, with

$$|\mathbf{p}_{0\rm true} - \mathbf{p}_0|^2 + |\mathbf{p}_{1\rm true}^U - \mathbf{p}_1^U|^2 \approx O\left(|\mathbf{p}_0 - \mathbf{e}|^2 + |\mathbf{p}_1^U - \mathbf{e}|^2\right).$$

See Figure 4a. Thus, for $\mathbf{T} \approx \mathbf{T}_{\rm true}$ and $R \approx R_{\rm true}$, one encounters the ambiguous configuration only when $\mathbf{T}$, $\mathbf{T}_{\rm true}$ lie close to the measurements $\hat{\mathbf{p}}_0$ and $\hat{\mathbf{p}}_1^U$.

However, for motion estimates very different from the true values, the unrotated point $\mathbf{p}_1^U = R^{-1} * \mathbf{p}_1$ can be far from $\mathbf{p}_0$ and $\mathbf{p}_{1\rm true}^U = R_{\rm true}^{-1} * \mathbf{p}_{1\rm true}$, and one can get ambiguous configurations with $\mathbf{e}$ far from $\mathbf{p}_0$ and $\mathbf{p}_1^U$, see Figure 4b. This matters since, as we discuss next, the existence of the ambiguity is associated with a breakdown in the *WC* approximation to the least–squares error.

Least–squares approximations and the triangulation ambiguity. Consider the *WC*–like approximation to the directional error $E_{\theta m}\left(R, \hat{\mathbf{T}}\right)$ given by

$$E_{WC\theta m}\left(R, \hat{\mathbf{T}}\right) \equiv \frac{B_m}{A_m} = \frac{\left(\hat{\mathbf{T}} \cdot \hat{\mathbf{p}}_{0m} \times \hat{\mathbf{p}}_{1m}^U\right)^2}{\hat{\mathbf{p}}_{0m}^T\left(1 - \hat{\mathbf{T}}\hat{\mathbf{T}}^T\right)\hat{\mathbf{p}}_{0m} + \hat{\mathbf{p}}_{1m}^{UT}\left(1 - \hat{\mathbf{T}}\hat{\mathbf{T}}^T\right)\hat{\mathbf{p}}_{1m}^U}.$$

By arguments similar to those of [17], one can show that $E_{WC\theta m}$ gives an excellent approximation to the exact error $E_{\theta m}\left(R, \hat{\mathbf{T}}\right)$ except when the parameter $\rho_m \equiv 4B_m/A_m^2$ approaches its upper limit of 1. Moreover, one can show that $\rho_m \longrightarrow 1$ only when the points $\hat{\mathbf{p}}_{0m}$ and $\hat{\mathbf{p}}_{1m}^U$ approach the ambiguous

(a) (b)

Figure 4: (a): For points in an ambiguous configuration, and for $\mathbf{e} \approx \mathbf{e}_{\rm true}$ and $R \approx R_{\rm true}$, the distance from $\mathbf{e}$ to the points is of the same order as the distance from the measured points to the true ones and, thus, is of the same order as the noise. (b): When $\mathbf{e}$ differs from $\mathbf{e}_{\rm true}$, or R from $R_{\rm true}$, $\mathbf{e}$ can be far from the measured points.

configuration [17]. Thus, the existence of the triangulation ambiguity is associated with the failure of the *WC*–like approximation $E_{WC\theta m}$.

A similar result holds for the *original WC* approximation. If the rotation is around the z axis, there is a triangulation ambiguity for the *image–plane* error just as for the directional error. It is easy to show that the optimal structure estimate is ambiguous whenever the image points lie in the right–isosceles configuration

$$\begin{aligned} (\mathbf{p}_0 - \mathbf{e}) &\perp \left(\mathbf{p}_1^U - \mathbf{e}\right), \\ |\mathbf{p}_0 - \mathbf{e}| &= \left|\mathbf{p}_1^U - \mathbf{e}\right|, \end{aligned}$$

for any finite $\mathbf{e}$. For the same situation, with the rotation around the z–axis, [17] showed that the *WC* error gives a poor approximation only near the right–isosceles configuration.[4] Thus, the ambiguity that we have found connects to the failure of the *WC* approximation.

For rotations out of the image plane, with $R\hat{\mathbf{z}} \neq \hat{\mathbf{z}}$, the optimal structure estimate for the image–plane error is ambiguous only for some motions, see the Appendix. We have not investigated the goodness of the *WC* approximation for out–of–plane rotations, but the fact that the *WC*–like approximation

[4] For z–rotations, the "intermediate" error analyzed in [17] is equivalent to the *WC* error.

$E_{WC\theta m}$ fails near right–isosceles configurations suggests that the *WC* approximation may also have troubles near these configurations even when there is no actual ambiguity.

4 Experiments: the Least–Squares Error

We verified our result for $E_{\theta,\mathrm{ex}}\left(R,\hat{\mathbf{T}}\right)$ in (13) experimentally. For several real and synthetic sequences, we minimized the directional least–squares error $E_{\theta LS}\left(R,\mathbf{T},\{P\}\right)$ over all the $\mathbf{P}_m$ using a generalized optimization routine from MATLAB. The result was always approximately equal to $E_{\theta,\mathrm{ex}}\left(R,\hat{\mathbf{T}}\right)$ and always larger than $E_{\theta,\mathrm{ex}}\left(R,\hat{\mathbf{T}}\right)$. Presumably, the discrepancies between the minimization results and $E_{\theta,\mathrm{ex}}$ were due to our halting the minimization routine before it had fully converged. One advantage of having an explicit form $E_{\theta,\mathrm{ex}}$ for the error is that we don't have to worry about what criterion to use for halting the minimization over the structure (determining the stopping criterion for a minimization routine can be a difficult problem).

Compared to a standard bundle–adjustment minimization of the image–plane least–squares error over the structure and motion (e.g., [25]), minimizing $E_{\theta,\mathrm{ex}}\left(R,\hat{\mathbf{T}}\right)$ over $R,\hat{\mathbf{T}}$ was about 100 times faster. Since bundle adjustment searches for the minimum of a much higher–dimensional and more complicated function than our approach does, it is more likely to get lost in the complexities of this function and either converge more slowly or converge to a false local minimum.

We used our exact form for $E_{\theta,\mathrm{ex}}\left(R,\hat{\mathbf{T}}\right)$ to study the least–squares error experimentally. Since we previously analyzed the error for sideways translations in [17], we focus here on forward/backward motion. That is, we study the error for epipole estimates that lie in or near the image region of tracked points. Note that the true translation can be in any direction. The error we study is the *exact* error for *large* motions. What makes this feasible is that we avoid a costly minimization over the structure but still compute the exact error. Though we focus on the directional error, we expect our qualitative conclusions to apply to the least–squares error in the image plane.

We demonstrate that the error for forward/backward motions is often complex and can have several local minima. [1] recently verified this experimentally for infinitesimal motion, confirming our suggestion in [17], and we show it here for large motions. Below, for simplicity, we use "forward" to include "backward."

We report results for noiseless synthetic sequences generated using the measured ground–truth struc-

ture from two real motion sequences: the UMASS PUMA sequence [11] and the UMASS/Martin–Marietta Rocket–Field sequence [2]. The depths vary from 13–32 for the PUMA structure and from 17–67 for the Rocket–Field structure. We used several translations ranging in magnitude up to 7.2 for the PUMA structure and up to 4.2 for the Rocket–Field structure. For simplicity, we took the *true* rotation to be zero; the directional error is *invariant* to the true rotation. In Figures 5–7, we represent the translation direction by the epipole $\mathbf{e} \equiv (T_x; T_y)\,/T_z$. The figures show

$$\underline{E}_\theta(\mathbf{e}) \equiv \min_R E_{\theta,\mathrm{ex}}(R, \mathbf{e})\,.$$

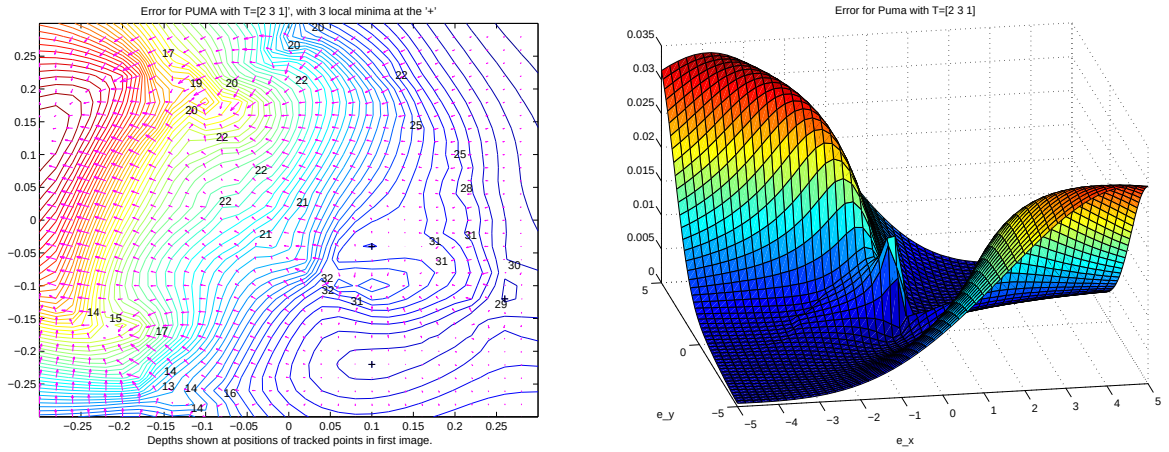
Figures 5–6 show results for the PUMA structure, and Figure 7 shows results for the Rocket–Field structure. For the PUMA structure with $\mathbf{T} = (2; 3; 1)$, the three false local minima have errors of 2.0×10^{-3}, 2.2×10^{-3}, and 3.2×10^{-3}, while the error has a total range of $\left[1.1 \times 10^{-3}, 15.7 \times 10^{-3}\right]$ as $\mathbf{e}$ varies throughout the image region $-0.3 \leq e_x, e_y \leq 0.3$. The local minimum with an error of 2.2×10^{-3} has the largest domain of attraction, with a radius of about 0.07—that is, steepest–descent paths on the error surface always converge to this local minimum if they start from an $\mathbf{e}$ that is within 0.07 of the $\mathbf{e}$ giving the local minimum**.** The number 0.07 corresponds to about 20 pixels in the original image. This false minimum has a large enough domain of attraction and a low enough error to be dangerous for algorithms.

In Figure 6, for the PUMA structure and $\mathbf{T} = (-1; 1; 3)$, the extreme–closeup plot of the error shows that the pictured false minimum has a domain of attraction with radius of about 0.04 or 11 pixels. In Figure 7, the extreme–closeup plot for $\mathbf{T} = (1; 0.5; 2.5)$ shows a false minimum whose domain of attraction has a radius of about 0.01, corresponding to about 7 pixels. This local minimum has an error of 8×10^{-4}, compared to the error's total range of $[2 \times 10^{-4}, 19 \times 10^{-4}]$ for $\mathbf{e}$ varying over the image region ($-0.32 < e_x < 0.32, -0.15 < e_y < 0.2$).

To summarize, for the Rocket structure we found 2 false minima for $\mathbf{T} = (3; 0; 3)$ and 2 for $\mathbf{T} = (1; 0.5; 2.5)$. For the PUMA structure, we found 3, 1, and 2 false minima for, respectively, $\mathbf{T} = (2; 3; 1)$, $\mathbf{T} = (-1; -1; 3)$, and $\mathbf{T} = (4; 0; 6)$.

Note that the error is most complex for forward–motion estimates. For sideways–motion estimates, the error is relatively smooth, and [15][17] have shown that one can understand its behavior analytically. The error surfaces depicted in Figures 5 and 6 confirm that the "reflected" or "rubbery" local minimum [1][17] occurs for sideways motions, and that it occurs when the true translation is large. The "reflected" minima are far from the FOV on the other side of the image center from $\mathbf{e}_{\mathrm{true}}$, lying roughly on the line

.



(a) Closeup (b)

Figure 5: Contour and surface plots of the error $\underline{E_{\boldsymbol{\theta}}}$ as a function of the epipole $\mathbf{e}$, for the structure derived from the PUMA sequence and $\mathbf{T} = (2; 3; 1)$. The surface plot shows the error in low resolution over a large region of $\mathbf{e}$, while the contour plot shows a high–resolution closeup of the image region. The arrows indicate the gradient direction. The error has 3 false local minima marked by '+'. Depths are shown at the positions of the tracked points in the zeroth image. Two of the local minima are at $e_x \approx 0.1$, and the third is near the point with depth 29 at the far right.

joining the image center and $\mathbf{e}_{\text{true}}$.

5 The Image–Plane Error and Uncalibrated Cameras

This section extends our approach to approximate the least–squares error in the *image plane*. We also give an approximation to the least–squares error for uncalibrated cameras (the projective case).

The exact image–plane error $E_{LS}\left(R, \mathbf{T}, \{\mathbf{P}\}\right)$ was defined in (1). We approximate it by

$$E_{LS,\text{approx}}\left(R, \mathbf{T}, \{\mathbf{P}\}\right) \equiv \sum_{m-1}^{N_p} \left(\left| \underline{\mathbf{P}_m} - \mathbf{p}_{0m} \right|^2 + \left| \underline{(\mathbf{P}_m - \mathbf{T})} - \mathbf{p}_{1m}^U \right|^2 \right), \tag{26}$$

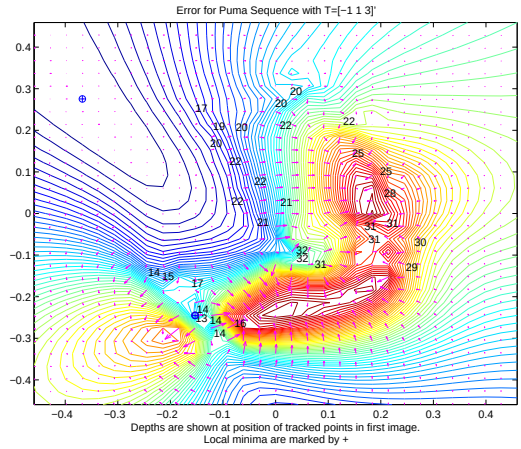
where, by analogy with the discussion in Section 2.3, we define $\mathbf{p}_{1m}^U \equiv R^{-1} * \mathbf{p}_{1m}$. One can compute the infimum of $E_{LS,\text{approx}}$ over the $\mathbf{P}_m$ exactly [17]:

$$E_{\text{ap}}\left(R, \hat{\mathbf{T}}\right) \equiv \inf_{\{\mathbf{P}\}} E_{LS,\text{approx}}\left(R, \mathbf{T}, \{\mathbf{P}\}\right) = \sum_{m=1}^{N_p} \left(\alpha_m/2 - \sqrt{\alpha_m^2/4 - \beta_m} \right), \tag{27}$$

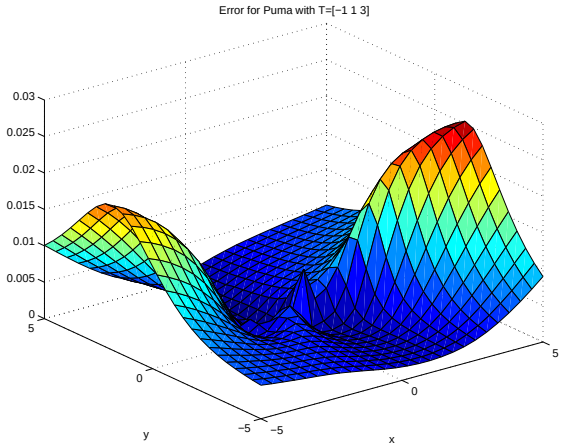
$$\alpha_m \equiv \left|\mathbf{p}_{0m} - \mathbf{e}\right|^2 + \left|\mathbf{p}_{1m}^U - \mathbf{e}\right|^2, \quad \beta_m \equiv \left|(\mathbf{p}_{0m} - \mathbf{e}) \times (\mathbf{p}_{1m}^U - \mathbf{e})\right|^2.$$

The function $E_{\text{ap}}\left(R, \hat{\mathbf{T}}\right)$ gives our initial approximation to the image–plane error $E\left(R, \hat{\mathbf{T}}\right)$ defined in (2). All quantities in (27) are in the image plane.

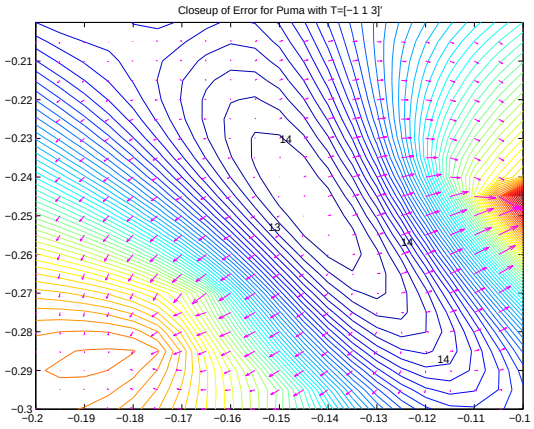
.



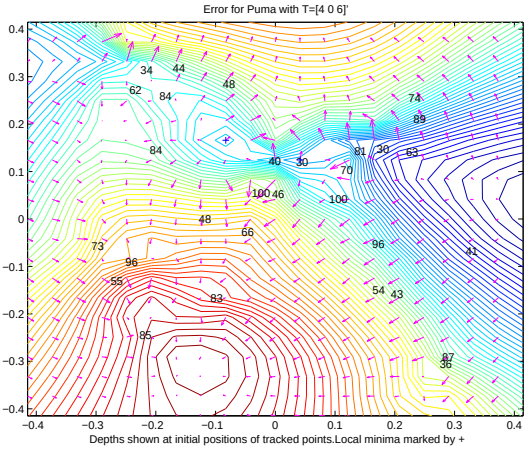
(-1,1,3) closeup



(-1,1,3)



(-1,1,3) extreme closeup



(4,0,6)

Figure 6: Contour and surface plots of the error $\underline{E_{\boldsymbol{\theta}}}$ for the PUMA sequence. The error has 1 false local minimum for $\mathbf{T} = (-1, 1, 3)^T$ and 2 for $\mathbf{T} = (4, 0, 6)^T$. As in Figure 5, the false minima are marked by '+' and depths are shown at the positions of the tracked points in the zeroth image.

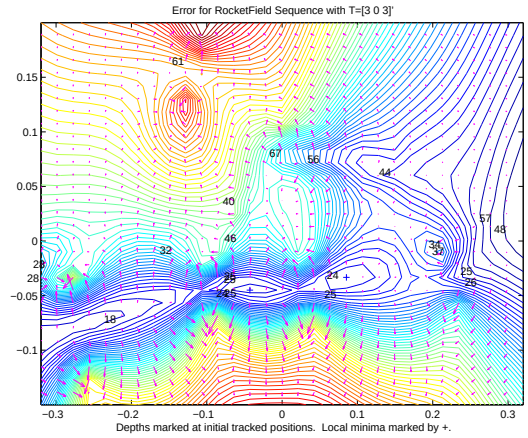
If the rotation R is around the z–axis, with $R\hat{\mathbf{z}} = \hat{\mathbf{z}}$, then

$$\left|\underline{R\left(\mathbf{P}-\mathbf{T}\right)} - \mathbf{p}_{1m}\right| = \mid \left[[R]\right]_2 \left(\underline{\mathbf{P}-\mathbf{T}}\right) - \mathbf{p}_{1m} \mid = \left|\left(\underline{\mathbf{P}-\mathbf{T}}\right) - \left[[R]\right]_2^{-1} \mathbf{p}_{1m}\right| = \left|\left(\underline{\mathbf{P}-\mathbf{T}}\right) - R^{-1} * \mathbf{p}_{1m}\right|. \tag{28}$$

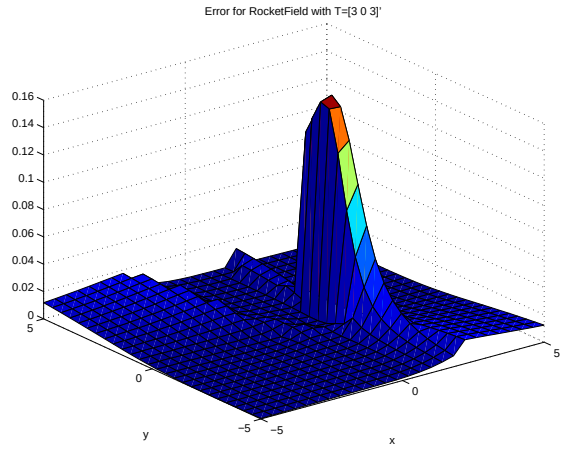
This implies that our approximation $E_{LS,\text{approx}}$ exactly equals the true error E_{LS}, and $E_{\text{ap}}\left(R, \hat{\mathbf{T}}\right)$ exactly equals $E\left(R, \hat{\mathbf{T}}\right)$, for *any* direction of the translation. In contrast, the *WC* error becomes exact only when R is a z–rotation *and* $\mathbf{T} \cdot \hat{\mathbf{z}} = 0$.

The fact that E_{ap} becomes exact for z–rotations implies that it, unlike the *WC* error, gives a good approximation to the true error for forward motion estimates. For a forward estimate with $\mathbf{T} \sim \hat{\mathbf{z}}$, the

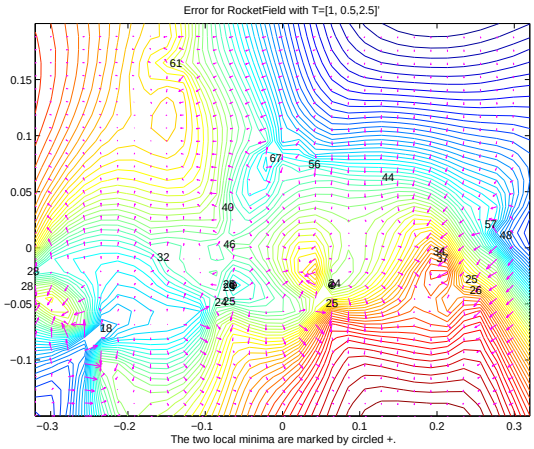
.



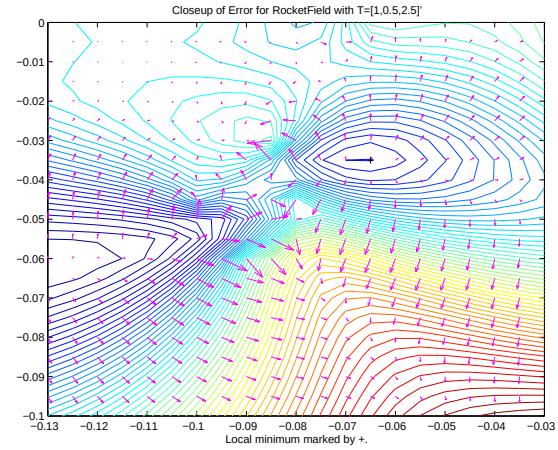
(3,0,3) closeup



(3,0,3)



(1,0.5,2.5) closeup



(1,0.5,2.5) extreme closeup

Figure 7: Contour and surface plots of the error $\underline{E}_{\boldsymbol{\theta}}$ as a function of $\mathbf{e}$, with the structure derived from the Rocket–Field sequence. The error $\underline{E}_{\boldsymbol{\theta}}$ has 2 forward local minima for $\mathbf{T} = (3, 0, 3)^T$ and 2 for $T = (1, 0.5, 2.5)^T$. As in Figure 5, the local minima are marked by '+' and depths are shown at the positions of the tracked points in the zeroth image.

corresponding best rotation estimate cannot differ much from a z–rotation, since otherwise it would remove the image points from the camera's field of view. Since $E_{\rm ap}$ is exact for z–rotations, we expect it to give a good approximation for rotations close to z–rotations. Thus, it should give a good approximation for forward translation estimates (and the best corresponding rotation estimates).

We now describe a modification of $E_{\rm ap}$ that gives a good approximation over a larger region of rotations around the z–rotations. Instead of unrotating image 1 by the full rotation R, via our definition of p_1^U, we unrotate by part of the rotation. One can write any rotation R as $R = R_{z1}R_xR_{z0}$, where the R_{zi} are "in–plane" rotations around the z axis and R_x is an "out–of–plane" rotation about the x axis.

Only the out–of–plane rotation R_x causes problems for our approximation. We have the exact equality

$$\left|\underline{\mathbf{P}_m} - \mathbf{p}_{0m}\right|^2 + \left|\underline{R\left(\mathbf{P}_m - \mathbf{T}\right)} - \mathbf{p}_{1m}\right|^2 = \left|\underline{R_x^{-1/2}\mathbf{P}_m^*} - R_{z0}\mathbf{p}_{0m}\right|^2 + \left|\underline{R_x^{1/2}\left(\mathbf{P}_m^* - \mathbf{T}^*\right)} - R_{z1}^{-1}\mathbf{p}_{1m}\right|^2,$$

where $\mathbf{P}_m^* \equiv R_x^{1/2} R_{z0} \mathbf{P}_m$ and $\mathbf{T}^* \equiv R_x^{1/2} R_{z0} \mathbf{T}$. This gives

$$E\left(R, \hat{\mathbf{T}}\right) = \sum_{m=1}^{N_p} \inf_{\mathbf{P}_m^*} \left(\left|\underline{R_x^{-1/2}\mathbf{P}_m^*} - R_{z0}\mathbf{p}_{0m}\right|^2 + \left|\underline{R_x^{1/2}\left(\mathbf{P}_m^* - \mathbf{T}^*\right)} - R_{z1}^{-1}\mathbf{p}_{1m}\right|^2 \right).$$

We approximate this by

$$\sum_{m=1}^{N_p} \inf_{\mathbf{P}_m^*} \left(\left|\underline{\mathbf{P}_m^*} - R_x^{1/2} * R_{z0}\mathbf{p}_{0m}\right|^2 + \left|\underline{\left(\mathbf{P}_m^* - \mathbf{T}^*\right)} - R_x^{-1/2} * R_{z1}^{-1}\mathbf{p}_{1m}\right|^2 \right),$$

which gives a better approximation than E_{ap} since $R_x^{1/2}$ is half as large as the original out–of–plane R_x. To do even better, we instead use

$$\sum_{m=1}^{N_p} \inf_{\mathbf{P}_m^*} \left(\left|C_{0m}\left(\underline{\mathbf{P}_m^*} - R_x^{1/2} * R_{z0}\mathbf{p}_{0m}\right)\right|^2 + \left|C_{1m}\left(\underline{\left(\mathbf{P}_m^* - \mathbf{T}^*\right)} - R_x^{-1/2} * R_{z1}^{-1}\mathbf{p}_{1m}\right)\right|^2 \right), \quad (29)$$

where the $C_{im} \in \Re^{2\times 2}$ are matrices that correct for the effects of the unrotation by $R_x^{1/2}$or $R_x^{-1/2}$ up to second order in the noise. In general, one can no longer compute (29) exactly. Our final approximation $E_{\mathrm{mod}}(R, \hat{\mathbf{T}})$ is given by approximating (29) up to second order in the noise. This approximation is still exact if R is purely a z–rotation, and it gives a better approximation to the exact result for small out–of–plane rotations.

We experimentally compared our new approximation to the *WC* approximation for forward/backward motion on four real image pairs. We computed the errors on the grid $e_x = -0.5 : 0.1 : 0.5$, $e_y = -0.5 : 0.1 : 0.5$. The image pairs came from the PUMA, Rocket–field and Castle sequences (available from CMU), and one of our indoor sequences. We computed the angular discrepancies between the length–121 vector of the exact least–squares errors and the corresponding vectors for our approximation and the *WC* error. Respectively, our approximation and *WC* gave discrepancies of 0.007° (ours) vs. 0.025° (*WC*), 0.38° (ours) vs. 0.285° (*WC*), 0.0037° (ours) vs. 0.0238° (*WC*), and 0.02° (ours) vs. 0.57° (*WC*), respectively for the four pairs. We found similar results on synthetic image pairs, whether the true motion was sideways or forward. For instance, for an image pair generated using the Rocket–field

structure, we found errors of 0.02° (ours) and 0.79° (*WC*) for sideways motion, and 0.012° (ours) and 0.072° (*WC*) for forward motion.

For the projective case, one can write the least–squares image–plane error in a form similar to (1) as

$$E_{\text{proj}}\left(H, \mathbf{E}, \{\mathbf{P}\}\right) \equiv \sum_{m=1}^{N_p} \left(\left| \underline{\mathbf{P}_m} - \mathbf{p}_{0m} \right|^2 + \left| \underline{H\left(\mathbf{P}_m - \mathbf{E}\right)} - \mathbf{p}_{1m} \right|^2 \right),$$

where $H \in \Re^{3\times 3}$ is a homography and $\mathbf{E} \in \Re^3$ is the epipole in homogeneous coordinates. Redefining $\mathbf{p}_{1m}^U \equiv \underline{H^{-1}\bar{\mathbf{p}}_{1m}}$, one can approximate E_{proj} by (26) as before, and the infimum of (26) over the $\mathbf{P}_m$ gives a form similar to (27). These approximations become *exact* whenever H corresponds to a z–rotation plus a shift in the image center and a change in the focal length—i.e., they are exact for a 4–parameter subgroup of the full homography group.[5] The homography H has only 5 free parameters [16][4]. Thus, as in the calibrated case, there is just a single parameter of the inter–image transformation that makes our result approximate rather than exact. This suggests that one could modify the analog of (27) along the lines of (29) to give a good approximation for H not too far from the 4–parameter subgroup mentioned above.

One could also use the directional least–squares error for the projective case, with

$$E_{\text{proj}\theta}\left(H, \mathbf{E}, \{\mathbf{P}\}\right) \equiv \sum_{m=1}^{N_p} \left(\left| \frac{\mathbf{P}_m}{|\mathbf{P}_m|} \times \hat{\mathbf{p}}_{0m} \right|^2 + \left| \frac{H\left(\mathbf{P}_m - \mathbf{E}\right)}{|H\left(\mathbf{P}_m - \mathbf{E}\right)|} \times \hat{\mathbf{p}}_{1m} \right|^2 \right).$$

One can approximate $\inf_{\{\mathbf{P}\}} E_{\text{proj}\theta}\left(H, \mathbf{E}, \{\mathbf{P}\}\right)$ as before by (12) with $R \longrightarrow H$ and $\mathbf{T} \longrightarrow \mathbf{E}$, and the approximation is exact when H is a rotation. Using the QR decomposition, one can write any homography as $H = RU$, where R is a rotation and U is upper triangular. (12) is exact for H with U equal to the identity and gives a good approximation for H with U close to the identity.

6 Conclusion

We presented a new algorithm for 2–image structure from motion that gives fast and optimal reconstructions from point features. We showed that the directional least–squares image–reprojection error can be written exactly as a simple function of the motion alone. We used this result to study the directional error experimentally and demonstrated that local minima occur frequently for forward or backward motion estimates (i.e., with the estimated epipole close to the field of view). Our results also confirm that

[5] For the focal length change one must modify the approximations by multiplying $\left(\mathbf{p}_{1m}^U - \mathbf{e}\right)$ by a scale.

the "reflected" local minimum at sideways estimates (see [17][1]) occurs when the true translation has large magnitude.

We derived an exact explicit expression for the optimal estimate of the structure given the motion and pointed out a new ambiguity in recovering the structure for known motion. Our expression is optimal for the directional image–reprojection error. The ambiguity also occurs for the image–plane least–squares error, though only for special motions. The ambiguity connects to the failure of the *WC* (Sampson) approximation, the standard first–order approximation to the least–squares error in the image plane. In the Appendix, we derive an improved form of the *WC* error that gives better results experimentally.

7 Appendix

7.1 Optimal Structure Estimate

For brevity, we drop the subscript m in this section.

Proof of Claim 1. From (6), any choice of $\mathbf{P}$ on the epipolar plane (excluding $\tilde{\mathbf{0}}$ or $\hat{\mathbf{T}}$) gives a unique pair of $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ on the epipolar circle. If one could show that any choice of the pair $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ corresponds to a unique $\mathbf{P}$, the claim would follow immediately, since the set of all $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ would be equivalent to the set of all $\mathbf{P}$ (excluding $\tilde{\mathbf{0}}$ or $\hat{\mathbf{T}}$). Unfortunately, the equivalence does not hold exactly. What one *can* show is that, for all choices of $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ apart from a set of measure zero, one can find a $\mathbf{P}$ that reproduces these choices via (6) *up to sign*, i.e., up to multiplications by -1. This will be enough to prove the claim, since, as noted in Section 2.3, the quantity $|\Delta_\theta|^2$ does not depend on factors of -1 in $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$.

For the moment, we exclude the exceptional cases $\hat{\mathbf{a}} = \pm\hat{\mathbf{T}}$ or $\hat{\mathbf{b}} = \pm\hat{\mathbf{T}}$ or $\hat{\mathbf{b}} = \pm\hat{\mathbf{a}}$. These are one–dimensional sets which have measure zero in the space of all pairs $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$. Choose any other values for $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ on the epipolar circle. To reproduce the chosen $\hat{\mathbf{a}}$ up to a sign factor, we take $\mathbf{P} = \lambda\hat{\mathbf{a}}$ where λ is a constant. We have $|\mathbf{P}| = |\lambda|$. By choosing positive values for λ, it is clear from the definition of $\hat{\mathbf{b}}$ that one can reproduce any value of $\hat{\mathbf{b}}$ on the arc of the unit circle between $\hat{\mathbf{a}}$ and $-\hat{\mathbf{T}}$. By choosing negative values, one can reproduce any value of $\hat{\mathbf{b}}$ on the arc between $-\hat{\mathbf{a}}$ and $-\hat{\mathbf{T}}$. Together, these two arcs covers the whole half–circle from $-\hat{\mathbf{a}}$ to $\hat{\mathbf{a}}$, apart from the exceptional points $-\hat{\mathbf{T}}$, $\hat{\mathbf{a}}$, and $-\hat{\mathbf{a}}$. This means that one can reproduce any $\hat{\mathbf{b}}$, except (possibly) for $\pm\hat{\mathbf{T}}$ or $\pm\hat{\mathbf{a}}$, up to a factor of -1. Modulo the exceptional cases, there exists a $\mathbf{P}$ that reproduces both $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ up to sign.

We now show that the infimum over $\mathbf{P}$ gives the same result as the infimum over $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$. We have

$$\inf_{\mathbf{P}} |\Delta_\theta|^2 \Big|_{\mathbf{P}\cdot\hat{\mathbf{n}}=0} = \inf_{\mathbf{P}} \left(|\hat{\mathbf{a}}(\mathbf{P}) \times \hat{\mathbf{p}}_0|^2 + \left| \hat{\mathbf{b}}(\mathbf{P}) \times \hat{\mathbf{p}}_1^U \right|^2 \right) \Big|_{\mathbf{P}\cdot\hat{\mathbf{n}}=0} .$$

Clearly, taking the infimum over all $\mathbf{P}$ is equivalent to taking the infimum over the constrained set of all pairs $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ which come from some $\mathbf{P}$.

Since $|\mathbf{\Delta}_\theta|^2$ does not depend on sign factors in $\hat{\mathbf{a}}$ or $\hat{\mathbf{b}}$, one can extend the latter infimum to one over all pairs $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ which differ at most by factors of -1 from the $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ that come from some $\mathbf{P}$. The extended infimum still gives the same result as the infimum over $\mathbf{P}$. As we showed above, this extended set consists of all pairs $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ except for the measure–zero set where $\hat{\mathbf{a}} = \pm\hat{\mathbf{T}}$ or $\hat{\mathbf{b}} = \pm\hat{\mathbf{T}}$ or $\hat{\mathbf{b}} = \pm\hat{\mathbf{a}}$.

Since $|\Delta_\theta|^2$ is continuous in $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$, its infimum over all $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ *excluding* the exceptional values $\hat{\mathbf{a}} = \pm\hat{\mathbf{T}}$ and $\hat{\mathbf{b}} = \pm\hat{\mathbf{T}}$ and $\hat{\mathbf{b}} = \pm\hat{\mathbf{a}}$ is the same as its minimum over all $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ *including* these values. (Consider a sequence of nonexceptional values for $\hat{\mathbf{a}}, \hat{\mathbf{b}}$ such that the sequence converges to one of the exceptional pairs, e.g., $\hat{\mathbf{a}} = \hat{\mathbf{T}}$ and $\hat{\mathbf{b}} \neq \hat{\mathbf{T}}$. The continuity of $|\Delta_\theta|^2$ in $\hat{\mathbf{a}}$, $\hat{\mathbf{b}}$ implies that the limit of the sequence equals the value of $|\hat{\mathbf{a}} \times \hat{\mathbf{p}}_0|^2 + \left|\hat{\mathbf{b}} \times \hat{\mathbf{p}}_1^U\right|^2$ at the exceptional pair.) The extended infimum still gives the same result as the infimum over $\mathbf{P}$, which demonstrates the claim. ■

Proof of Claim 2. First, we show that $\kappa = 0$ only when $\hat{p}_{1,x}^U = 0$ or $\hat{p}_{1,y}^U = 0$, and then we show that (15) converges to the correct values for $\hat{\mathbf{n}}$ in the limits $\hat{p}_{1,x}^U \longrightarrow 0$ or $\hat{p}_{1,y}^U \longrightarrow 0$. (We do not consider $\hat{p}_{1,x}^U = \hat{p}_{1,y}^U = 0$, since this is the case $\hat{\mathbf{p}}_1^U \parallel \mathbf{T}$ dealt with in Section 3.1.)

Since $|\hat{p}_{0,x}| > 0$, the second line of (16) implies that $G = 0$ only when $\hat{p}_{1,x}^U = 0$ *and* $\hat{p}_{0,x}^2 = \left(\hat{p}_{1,y}^U\right)^2$, which are exactly the conditions (21) that make $\hat{\mathbf{n}}$ non–unique. Since we are assuming that $\hat{\mathbf{n}}$ *is* unique and, thus, that (21) does not hold, we can take $G > 0$. From $G > 0$ and the expressions for G and κ in (16) and (17), we have $\kappa = 0$ only when $4\left(\hat{p}_{1,x}^U\right)^2\left(\hat{p}_{1,y}^U\right)^2 = 0$. This implies that, as claimed, $\kappa = 0$ only when $\hat{p}_{1,x}^U = 0$ or $\hat{p}_{1,y}^U = 0$.

First, we consider the case $\hat{p}_{1,y}^U = 0$ and $\hat{p}_{1,x}^U \neq 0$. For small $\hat{p}_{1,y}^U$, (15) gives

$$\hat{\mathbf{n}} \approx \frac{1}{2\left|\hat{p}_{1,x}^U \hat{p}_{1,y}^U\right|} \begin{pmatrix} -2(\hat{p}_{1,y}^U)^2(\hat{p}_{1,x}^U)^2 / \left(\hat{p}_{0,x}^2 + (\hat{p}_{1,x}^U)^2\right) \\ 2\hat{p}_{1,x}^U \hat{p}_{1,y}^U \end{pmatrix}$$

to lowest order, which implies that $\hat{\mathbf{n}} \longrightarrow (0;1)$ (up to sign) as $\hat{p}_{1,y}^U \longrightarrow 0$. One can confirm directly

from the expression for $[[S_\theta]]_2$ in (14) that its least eigenvector[6] equals $(0;1)$ (up to sign) at $\hat{p}^U_{1,y} = 0$.

Similarly, for small $\hat{p}^U_{1,x}$, (15) gives

$$\hat{\mathbf{n}} \approx \frac{1}{\kappa} \begin{bmatrix} \hat{p}^2_{0,x} - \left(\hat{p}^U_{1,y}\right)^2 - \left|\hat{p}^2_{0,x} - \left(\hat{p}^U_{1,y}\right)^2\right| + O\left((\hat{p}^U_{1,x})^2\right) \\ 2\hat{p}^U_{1,x}\hat{p}^U_{1,y} \end{bmatrix}$$

to lowest order. If $\hat{p}^2_{0,x} > \left(\hat{p}^U_{1,y}\right)^2$, this implies that $\hat{\mathbf{n}} \longrightarrow (0;\pm 1)$ as $\hat{p}^U_{1,x} \longrightarrow 0$, while, if $\hat{p}^2_{0,x} < \left(\hat{p}^U_{1,y}\right)^2$, it implies that $\hat{\mathbf{n}} \longrightarrow (\pm 1;0)$ as $\hat{p}^U_{1,x} \longrightarrow 0$. Again, one can confirm directly from the expression for $[[S_\theta]]_2$ in (14) that the least eigenvector of $[[S_\theta]]_2$ does have this behavior. ■

Proof of Claim 3. We show that $\hat{\mathbf{p}}_0$ and $\hat{\mathbf{p}}^U_1$ are never parallel to $\hat{\mathbf{n}}$ as long as $\hat{\mathbf{n}}$ is uniquely determined, i.e., the ambiguity condition (21) does not hold. First, suppose $\hat{p}^U_{1,x} \neq 0$ and $\hat{p}^U_{1,y} \neq 0$. Then our expression for $\hat{\mathbf{n}}$ in (15) implies that $|n_y| > 0$, and it follows that $\hat{\mathbf{p}}_0$ is not parallel to $\hat{\mathbf{n}}$. If $\hat{\mathbf{p}}^U_1$ were parallel to $\hat{\mathbf{n}}$, (15) would imply

$$\hat{p}^2_{0,x} + \left(\hat{p}^U_{1,x}\right)^2 - \left(\hat{p}^U_{1,y}\right)^2 - G = 2\left(\hat{p}^U_{1,x}\right)^2,$$

which one can rewrite as

$$\hat{p}^2_{0,x} - \left(\hat{p}^U_{1,x}\right)^2 - \left(\hat{p}^U_{1,y}\right)^2 = \sqrt{\left(\hat{p}^2_{0,x} - \left(\hat{p}^U_{1,x}\right)^2 - \left(\hat{p}^U_{1,y}\right)^2\right)^2 + 4\left(\hat{p}_{0,x}\right)^2\left(\hat{p}^U_{1,x}\right)^2}. \tag{30}$$

This holds only if $\hat{p}^U_{1,x} = 0$, counter to our assumption, so $\hat{\mathbf{p}}^U_1$ is not parallel to $\hat{\mathbf{n}}$.

Second, if $\hat{p}^U_{1,y} = 0$ and $\hat{p}^U_{1,x} \neq 0$, we have $\hat{\mathbf{n}} = (0;\pm 1)$, and it is clear that neither $\hat{\mathbf{p}}_0$ nor $\hat{\mathbf{p}}^U_1$ is parallel to $\hat{\mathbf{n}}$.

Third, suppose $\hat{p}^U_{1,x} = 0$ and $\hat{p}^U_{1,y} \neq 0$. If $\hat{p}^2_{0,x} > \left(\hat{p}^U_{1,y}\right)^2$, then $\hat{\mathbf{n}} = (0;\pm 1)$ and $\hat{\mathbf{p}}_0$ is not parallel to $\hat{\mathbf{n}}$. Also, $\hat{\mathbf{p}}^U_1$ is not parallel to $\hat{\mathbf{n}}$, since if it were we would have $\left(\hat{p}^U_{1,y}\right)^2 = 1$, counter to our assumption that $\hat{p}^2_{0,x} > \left(\hat{p}^U_{1,y}\right)^2$. (Recall that $\hat{\mathbf{p}}_0$ is a unit vector.) If $\hat{p}^2_{0,x} < \left(\hat{p}^U_{1,y}\right)^2$, similar arguments show that neither $\hat{\mathbf{p}}_0$ nor $\hat{\mathbf{p}}^U_1$ is parallel to $\hat{\mathbf{n}}$. As claimed, $\hat{\mathbf{p}}_0$ and $\hat{\mathbf{p}}^U_1$ are never parallel to $\hat{\mathbf{n}}$ when the condition (21) does not hold. This implies (18). ■

[6]The "least eigenvector" is the one with the least eigenvalue.

Computing the structure by triangulation: the derivation of (19) for the generic case. From the definitions of $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$, the structure satisfies $\mathbf{P} = \lambda_0 \hat{\mathbf{a}}$ and $\mathbf{P} = \mathbf{T} + \boldsymbol{\lambda}_1 \hat{\mathbf{b}}$ for some constants λ_0 and λ_1. Substituting the optimal estimates for $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ from (18) and equating the two expressions for $\mathbf{P}$, we get

$$\mathbf{P} = \lambda_0 \left(\mathbf{p}_0 - \hat{\mathbf{n}} \left(\hat{\mathbf{n}} \cdot \mathbf{p}_0\right)\right) = \mathbf{T} + \boldsymbol{\lambda}_1 \left(\hat{\mathbf{p}}_1^U - \hat{\mathbf{n}} \left(\hat{\mathbf{n}} \cdot \mathbf{p}_1^U\right)\right). \tag{31}$$

This equation expresses the usual triangulation constraint that $\mathbf{P}$ occurs at the intersection of the rays from the 2 images. Note that $\mathbf{P}$ may occur at infinity. We can include this case by allowing λ_0 and λ_1 to become infinite, as we discuss below.

Equation (31) is equivalent to the two equations

$$\begin{aligned} \lambda_0 \hat{\mathbf{T}} \cdot \left(\hat{\mathbf{n}} \times \hat{\mathbf{p}}_0\right) &= \lambda_1 \hat{\mathbf{T}} \cdot \left(\hat{\mathbf{n}} \times \hat{\mathbf{p}}_1^U\right), \\ \lambda_0 \hat{\mathbf{T}} \cdot \hat{\mathbf{p}}_0 &= 1 + \lambda_1 \hat{\mathbf{T}} \cdot \hat{\mathbf{p}}_1^U, \end{aligned} \tag{32}$$

which we derive by projecting (31) in the directions $\hat{\mathbf{T}} \times \hat{\mathbf{n}}$ and $\hat{\mathbf{T}}$. The "1" in the second line follows from our assumption that $|\mathbf{T}| = 1$. We rewrite (32) in our coordinate system as

$$\lambda_0 \left(-n_y \hat{p}_{0,x}\right) = \lambda_1 \left(n_x \hat{p}_{1,y}^U - n_y \hat{p}_{1,x}^U\right), \quad \lambda_0 \hat{p}_{0,z} = 1 + \lambda_1 \hat{p}_{1,z}^U,$$

which has the solutions

$$\begin{aligned} \lambda_0 &= \left(n_x \hat{p}_{1,y}^U - n_y \hat{p}_{1,x}^U\right)/D, \quad \lambda_1 = -n_y \hat{p}_{0,x}/D, \\ D &\equiv \hat{p}_{0,z} \left(n_x \hat{p}_{1,y}^U - n_y \hat{p}_{1,x}^U\right) + n_y \hat{p}_{0,x} \hat{p}_{1,z}^U. \end{aligned} \tag{33}$$

The denominator D in (33) becomes zero exactly when the rays $\hat{\mathbf{p}}_1^U - \hat{\mathbf{n}} \left(\hat{\mathbf{n}} \cdot \mathbf{p}_1^U\right)$ and $\mathbf{p}_0 - \hat{\mathbf{n}} \left(\hat{\mathbf{n}} \cdot \mathbf{p}_0\right)$ are parallel and "intersect" at an infinite $\mathbf{P}$, corresponding to infinite values for λ_0 and λ_1. Thus, (33) gives the correct limit for this situation.

Substituting our expression for $\hat{\mathbf{n}}$ from (15) into the solutions for λ_0 and λ_1 gives

$$\lambda_0 = \left(\hat{p}_{0,x}^2 - \left(\hat{p}_{1,x}^U\right)^2 - \left(\hat{p}_{1,y}^U\right)^2 - G\right) \hat{p}_{1,y}^U / \left(D' \hat{p}_{1,y}^U\right), \quad \lambda_1 = -\left(2 \hat{p}_{1,x}^U \hat{p}_{1,y}^U\right) \hat{p}_{0,x} / \left(D' \hat{p}_{1,y}^U\right),$$

$$D' \equiv 2 \hat{p}_{1,x}^U \hat{p}_{1,z}^U \hat{p}_{0,x} + \hat{p}_{0,z} \left(\hat{p}_{0,x}^2 - \left(\hat{p}_{1,x}^U\right)^2 - \left(\hat{p}_{1,y}^U\right)^2 - G\right), \tag{34}$$

where the normalizing factor κ in (15) cancels since the equations (33) are homogeneous in $\hat{\mathbf{n}}$. If $\hat{p}_{1,y}^U \neq 0$, one can rewrite the solutions as

$$\lambda_0 = \left(\hat{p}_{0,x}^2 - \left(\hat{p}_{1,x}^U\right)^2 - \left(\hat{p}_{1,y}^U\right)^2 - G\right)/D', \quad \lambda_1 = -2 \hat{p}_{1,x}^U \hat{p}_{0,x}/D'. \tag{35}$$

Since (15) correctly gives $\hat{\mathbf{n}} = (0; \pm 1)$ in the limit $\hat{p}^U_{1y} \longrightarrow 0$, it is not surprising that the solutions (35) are also correct in this limit. For $\hat{p}^U_{1y} \longrightarrow 0$, the equations (35) become

$$\lambda_0 = -\hat{p}^U_{1,x} / \left(\hat{p}_{0,x} \hat{p}^U_{1,z} - \hat{p}_{0,z} \hat{p}^U_{1,x} \right), \qquad \lambda_1 = -\hat{p}_{0,x} / \left(\hat{p}_{0,x} \hat{p}^U_{1,z} - \hat{p}_{0,z} \hat{p}^U_{1,x} \right).$$

One can verify directly that substituting $\hat{\mathbf{n}} = (0; 1)$ into (33) (and dividing through by $\hat{p}^U_{1,x} \neq 0$) yields these equations.

This gives all the ingredients needed to compute the structure $\mathbf{P}$. From (31), we have $\mathbf{P} = \lambda_0 \left(\hat{\mathbf{p}}_0 - \hat{\mathbf{n}} \left(\hat{\mathbf{n}} \cdot \hat{\mathbf{p}}_0 \right) \right)$. Writing $\hat{\mathbf{p}}_0$ and $\hat{\mathbf{n}}$ in our coordinate system using all 3 coordinates, we get (19) after some algebra.

Triangulation in the ambiguous case: deriving (22). We assume the ambiguous condition $\hat{p}^U_{1,x} = 0 = \hat{p}^2_{0,x} - \left(\hat{p}^U_{1,y} \right)^2$ and $\hat{p}^2_{0,x} \neq 1$. Since $\hat{\mathbf{n}}$ is ambiguous, we rewrite (33) as

$$\begin{aligned} \lambda_0 &= n_x \hat{p}^U_{1,y} / \left(n_x \hat{p}_{0,z} \hat{p}^U_{1,y} + n_y \hat{p}_{0,x} \hat{p}^U_{1,z} \right) = \hat{p}^{-1}_{0,z} n_x / \left(n_y s_1 s_2 + n_x \right), \\ \lambda_1 &= -n_y \hat{p}_{0,x} / \left(n_y \hat{p}_{0,x} \hat{p}^U_{1,z} + n_x \hat{p}^U_{1,y} \hat{p}_{0,z} \right), \end{aligned}$$

where $s_1 \equiv \hat{p}_{0,x} / \hat{p}^U_{1,y}$ and $s_2 \equiv \hat{p}^U_{1,z} / \hat{p}_{0,z}$ are sign factors. Then $\mathbf{P} = \boldsymbol{\lambda}_0 \left(\hat{\mathbf{p}}_0 - \hat{\mathbf{n}} \left(\hat{\mathbf{n}} \cdot \hat{\mathbf{p}}_0 \right) \right)$ gives (22).

The exceptional case $\hat{p}^U_{1,x} = 0$ and $\hat{p}^2_{0,x} = \left(\hat{p}^U_{1,y} \right)^2 = 1$. If $\hat{\mathbf{n}}$ is not parallel to $\hat{\mathbf{p}}_0$ or $\hat{\mathbf{p}}^U_1$, (18) remains valid and we get (22) as before. Suppose $\hat{\mathbf{n}} \parallel \hat{\mathbf{p}}_0$. As before, $\hat{\mathbf{b}} = \pm \left(\hat{\mathbf{p}}^U_1 - \hat{\mathbf{n}} \hat{\mathbf{n}} \cdot \hat{\mathbf{p}}^U_1 \right) = \pm \hat{\mathbf{p}}^U_1 = \pm \hat{\mathbf{y}}$, while one can choose $\hat{\mathbf{a}}$ in *any* direction in the y–z plane perpendicular to $\hat{\mathbf{p}}_0 = \hat{\mathbf{x}}$. Solving $\mathbf{P} = \lambda_0 \hat{\mathbf{a}} = \mathbf{T} + \boldsymbol{\lambda}_1 \hat{\mathbf{b}}$ as before for the intersection of the rays, we find $\lambda_0 = \hat{a}_z^{-1}$ and

$$\mathbf{P} = (0; \hat{a}_y / \hat{a}_z; 1). \tag{36}$$

Similarly, if $\hat{\mathbf{n}} \parallel \hat{\mathbf{p}}^U_1$, we find

$$\mathbf{P} = \left(\hat{b}_x / \hat{b} z; 0; 1 \right) \tag{37}$$

The $\hat{a}_z \longrightarrow 0$ limit of (36) and the $\hat{b}_z \longrightarrow 0$ limit of (37) are also valid solutions.

7.2 Ambiguous triangulation for the image–plane error

We analyze when the optimal structure estimate becomes ambiguous for the least–squares error in the *image plane*. We take the motion as known. Consider a single 3D point $\mathbf{P}_m$ and its images $\mathbf{p}_{0m}$, $\mathbf{p}_{1m}$.

For the image–plane error defined in (1), the optimal estimate of $\mathbf{P}_m$ is given by

$$\arg\min_{\mathbf{P}} \left(\left| \underline{\mathbf{P}} - \mathbf{p}_{0m} \right|^2 + \left| \underline{R\left(\mathbf{P} - \mathbf{T}\right)} - \mathbf{p}_{1m} \right|^2 \right). \tag{38}$$

Let F be the essential matrix. One can replace the minimization in (38) by a constrained minimization

$$\min_{\mathbf{v}_0, \mathbf{v}_1} \left(\left| \mathbf{v}_0 - \mathbf{p}_{0m} \right|^2 + \left| \mathbf{v}_1 - \mathbf{p}_{1m} \right|^2 \right) \Big|_{\bar{\mathbf{v}}_1 \cdot F \bar{\mathbf{v}}_0 = 0} \tag{39}$$

over image points $\mathbf{v}_0$ and $\mathbf{v}_1$ satisfying the coplanarity constraint $\bar{\mathbf{v}}_1 \cdot F \bar{\mathbf{v}}_0 = 0$ [4][5]. The values of $\mathbf{v}_0$ and $\mathbf{v}_1$ minimizing (39) determine the optimal estimate of $\mathbf{P}_m$, and the estimate for $\mathbf{P}_m$ is ambiguous if the minimizing values $\mathbf{v}_0$ and $\mathbf{v}_1$ are. Thus, we analyze when the minimizing $\mathbf{v}_0$ and $\mathbf{v}_1$ are ambiguous.

Define $\mathbf{e}_1 \equiv R * \mathbf{e}$ to be the epipole in image 1. We temporarily assume that neither $\mathbf{e}$ nor $\mathbf{e}_1$ is at infinity. To simplify the problem, we shift the images so that $\mathbf{e}$ and $\mathbf{e}_1$ are at the origins of image 0 and image 1, respectively, and we rotate the images around their centers to place $\mathbf{p}_{0m}$ on the x–axis and $\mathbf{p}_{1m}$ on the y–axis. One can do this because the image–plane distances in (39) are invariant to rotations and translations in the image plane. After the transform, one can write the minimization (39) in exactly the same form in terms of the transformed images and using a transformed essential matrix F^* [4][5]. Let U_0, $U_1 \in \Re^{2\times 2}$ be rotations in the image plane, and let $\mathbf{s}_0$, $\mathbf{s}_1 \in \Re^2$ be image–plane translations. Consider the minimization (39) for the transformed image points $\mathbf{p}^*_{0m} \equiv U_0 \left(\mathbf{p}_{0m} - \mathbf{s}_0 \right)$ and $\mathbf{p}^*_{1m} \equiv U_1 \left(\mathbf{p}_{1m} - \mathbf{s}_0 \right)$ as constrained by the transformed matrix F^*:

$$\min_{\mathbf{v}^*_0, \mathbf{v}^*_1} \left(\left| \mathbf{v}^*_0 - \mathbf{p}^*_{0m} \right|^2 + \left| \mathbf{v}^*_1 - \mathbf{p}^*_{1m} \right|^2 \right) \Big|_{\bar{\mathbf{v}}^*_1 \cdot F^* \bar{\mathbf{v}}^*_0 = 0}. \tag{40}$$

If F^* is given by

$$F^* = \begin{pmatrix} U_1 & (0;0) \\ \mathbf{s}_1^T & 1 \end{pmatrix} F \begin{pmatrix} U_0^{-1} & \mathbf{s}_0 \\ (0,0) & 1 \end{pmatrix}, \tag{41}$$

one can show that the minimization in (40) is equivalent to the original [4].

The epipoles satisfy $F\bar{\mathbf{e}} = \mathbf{0}$ and $\bar{\mathbf{e}}_1^T F = \mathbf{0}$ [4]. After the transform of the images, since the transformed epipoles $\mathbf{e}^*$ and $\mathbf{e}^*_1$ are now at the image origins, we have

$$F^* \left(0;0;1\right) = \left(0,0,1\right) F^* = 0.$$

This implies that the transformed essential matrix F^* has the simple form

$$F^* = \begin{bmatrix} F^*_{1,1} & F^*_{1,2} & 0 \\ F^*_{2,1} & F^*_{2,2} & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

We transformed the images to put F^* in this form.

We now follow the same strategy as for the directional error and in [5]. We first fix the epipolar plane and minimize (40) over all points $\mathbf{v}_0^*$ and $\mathbf{v}_1^*$ on this plane. Then, we minimize over the choice of the epipolar plane.

Fixing the epipolar *plane* is equivalent to fixing an epipolar *line* in image 0 that passes through $\mathbf{e}^* = (0; 0)$ [4]. Minimizing over all $\mathbf{v}_0^*$ on the epipolar plane is equivalent to minimizing over all $\mathbf{v}_0^*$ on this line, which we refer as L_0. We parameterize our choice of L_0 by its normal

$$\hat{\mathbf{n}}_0 = \begin{pmatrix} \cos\theta; & \sin\theta \end{pmatrix}.$$

We can write any point on this line as $\mathbf{v}_0^* = \lambda\,(-\sin\theta; \cos\theta)$, where $\lambda \in \Re^1$. For any such $\mathbf{v}_0^*$, the transformed coplanarity constraint requires the corresponding point $\mathbf{v}_1^*$ to satisfy $\left(\mathbf{v}_1^{*T}, 1\right) F^* \,(-\sin\theta; \cos\theta; 1)$, which we can rewrite as

$$\mathbf{v}_1^* \cdot \begin{pmatrix} -F_{1,1}^* \sin\theta + F_{1,2}^* \cos\theta \\ -F_{2,1}^* \sin\theta + F_{2,2}^* \cos\theta \end{pmatrix} = 0$$

using the special form of F^*. Thus, $\mathbf{v}_1^*$ must lie on an epipolar line in image 1 that passes through $\mathbf{e}_1^* = (0; 0)$. We refer to this line as L_1. Its normal $\hat{\mathbf{n}}_1$ is proportional to $\begin{pmatrix} -F_{1,1}^* \sin\theta + F_{1,2}^* \cos\theta; & -F_{2,1}^* \sin\theta + F_{2,2}^* \cos\theta \end{pmatrix}$. By construction, any two points $\mathbf{v}_0^*$ on L_0 and $\mathbf{v}_1^*$ on L_1 satisfy the coplanarity constraint for F^*. Thus, one can minimize over all $\mathbf{P}$ on the epipolar plane by minimizing separately and *independently* over all $\mathbf{v}_0^*$ on L_0 and $\mathbf{v}_1^*$ on L_1.[7]

It is easy to see that the result of this minimization is $|\hat{\mathbf{n}}_0 \cdot \mathbf{p}_{0m}^*|^2 + |\hat{\mathbf{n}}_1 \cdot \mathbf{p}_{1m}^*|^2$, the sum of the squared distances from the image points to the respective epipolar lines. We write this explicitly as

$$p_{0x}^{*2} \cos^2\theta + p_{1y}^{*2} \frac{\left(-F_{2,1}^* \sin\theta + F_{2,2}^* \cos\theta\right)^2}{\left(-F_{1,1}^* \sin\theta + F_{1,2}^* \cos\theta\right)^2 + \left(-F_{2,1}^* \sin\theta + F_{2,2}^* \cos\theta\right)^2}, \tag{42}$$

using the facts that $\mathbf{p}_{0m}^*$ is on the x axis and $\mathbf{p}_{1m}^*$ on the y axis. If there is an ambiguity similar to that for the directional error, any choice of the epipolar plane must give the same least–squares error. This means that (42) must be constant as a function of θ. We analyze when this happens.

If we rewrite (42) as a fraction, the numerator is fourth–order in $\sin\theta$ and $\cos\theta$ while the denominator is only quadratic. (42) can be a constant only if the fourth–order dependence cancels. Since the

[7] As usual, we neglect the positive depth constraint.

fourth–order contribution is $p_{0x}^{*2}\cos^2\theta$ times the denominator in (42), it is easy to see that this requires the denominator to be constant. This happens only if

$$\begin{aligned} F_{1,1}^{*2} + F_{2,1}^{*2} - F_{1,2}^{*2} - F_{2,2}^{*2} &= 0, \\ F_{2,2}^{*}F_{2,1}^{*} + F_{1,1}^{*}F_{1,2}^{*} &= 0, \end{aligned}$$

which imply $F_{2,2}^{*} = sF_{1,1}^{*}$ and $F_{1,2}^{*} = -sF_{2,1}^{*}$, where $s = \pm 1$. After substituting these values, (42) becomes

$$p_{0x}^{*2}\cos^2\theta + \frac{p_{1y}^{*2}}{F_{1,1}^{*2} + F_{2,1}^{*2}}\left(-F_{2,1}^{*}\sin\theta + gF_{1,1}^{*}\cos\theta\right)^2,$$

and it is easy to see this is constant only when $F_{1,1}^{*} = 0$ and $p_{0x}^{*2} = p_{1y}^{*2}$. Thus, the ambiguity occurs only for

$$F^{*} = \begin{bmatrix} 0 & -sF_{2,1}^{*} & 0 \\ F_{2,1}^{*} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{43}$$

and $p_{0x}^{2} = p_{1y}^{2}$.

We now analyze what motions permit the original essential matrix F to be transformed to the form (43). From the definition of the epipole $\mathbf{e}$, we have $\mathbf{T} \sim \bar{\mathbf{e}}$, and one can write $F = R\bar{\mathbf{e}}_{\times}$, where $\bar{\mathbf{e}}_{\times}$ is the skew–symmetric matrix that acts on a vector $\mathbf{V}$ as $\bar{\mathbf{e}}_{\times}\mathbf{V} = \bar{\mathbf{e}} \times \mathbf{V}$. If one shifts the images by the $\mathbf{s}_i$ to put the epipoles at the origins of the images, it is easy to show that F transforms (up to a scale factor) to

$$\begin{bmatrix} A & (0;0) \\ (0,0) & 0 \end{bmatrix}, \quad A \equiv \begin{bmatrix} R_{1,2} - R_{1,3}e_y, & -R_{1,1} + R_{1,3}e_x \\ R_{2,2} - R_{2,3}e_y, & -R_{2,1} + R_{2,3}e_x \end{bmatrix}.$$

After the full transforms in (41), we have

$$[[F^{*}]]_2 = U_1 A U_0^{-1}. \tag{44}$$

The expression for $[[F^{*}]]_2$ in (43) is an image–plane rotation up to scale (plus a reflection if $s = -1$). Since the U_i are rotations, (43) implies that, up to scale, the original matrix A must also be an image–plane rotation (or a rotation plus a reflection, if $s = -1$). This implies

$$\begin{aligned} R_{1,2} - R_{1,3}e_y &= s\left(-R_{2,1} + R_{2,3}e_x\right) \\ R_{2,2} - R_{2,3}e_y &= s\left(R_{1,1} - R_{1,3}e_x\right). \end{aligned}$$

Solving this gives $\mathbf{e} = s\,[R_{3,1}; R_{3,2}] \,/\, (1 + sR_{3,3})$ or $\mathbf{T} \sim \left(\hat{\mathbf{z}} \pm R^{-1}\hat{\mathbf{z}}\right)$, where we have used the fact that R is a rotation. Such a motion makes the matrix A a rotation or rotation–plus–reflection up to scale.

Thus, for $\mathbf{T} \sim \left(\hat{\mathbf{z}} \pm R^{-1}\hat{\mathbf{z}}\right)$, the optimal structure estimate for the image–plane error can have a one–parameter ambiguity similar to that for the directional error. The ambiguity actually occurs if the image points $\mathbf{p}_{0m}$, $\mathbf{p}_{1m}$ satisfy $|\mathbf{p}_{0m} - \mathbf{e}| = |\mathbf{p}_{1m} - \mathbf{e}_1|$, corresponding to $|\mathbf{p}^*_{0m}| = |\mathbf{p}^*_{1m}|$, *and* if the angle between the vectors $(\mathbf{p}_{1m} - \mathbf{e}_1)$ and $\Pi_s\,(\mathbf{p}_{0m} - \mathbf{e})$ equals $\tan^{-1}\,(A_{21}/A_{11})$, where Π_1 is the identity and Π_{-1} is a reflection around the x axis. (When the condition on the angle is satisfied, the rotations U_0 and U_1 make F^* of the form (43).)

One can easily show that there is no ambiguity when $\mathbf{e}$ or $\mathbf{e}_1$ is at infinity.

Our analysis says nothing about *discrete* ambiguities, which occur when 2 or 3 roots of the 6th degree polynomial in [5] give equal least–squares errors and equally good structure estimates.

7.3 Improving the *WC* Approximation to the Image–Plane Error

In this section, we derive the *WC* error and extend it to a second–order approximation of the true error.

As for the directional error, we define the measurement errors $\delta_{0m} \equiv \underline{\mathbf{P}_m} - \mathbf{p}_{0m}, \quad \delta_{1m} \equiv \underline{R\,(\mathbf{P}_m - \mathbf{T})} - \mathbf{p}_{1m}$. Also, we define $\mathbf{\Delta}_m \equiv (\boldsymbol{\delta}_{0m}; \boldsymbol{\delta}_{1m}) \in \Re^4$ and $E_m\,(R, \mathbf{T}) \equiv \inf_{\mathbf{P}_m} |\mathbf{\Delta}_m|^2$,

$$E\,(R, \mathbf{T}) \equiv \inf_{\mathbf{P}_1 \ldots \mathbf{P}_{Np}} E_{LS} = \inf_{\mathbf{P}_1 \ldots \mathbf{P}_{Np}} \sum_{m=1}^{N_p} |\mathbf{\Delta}_m|^2 = \sum_{m=1}^{N_p} \inf_{\mathbf{P}_m} |\mathbf{\Delta}_m|^2 = \sum_{m=1}^{N_p} E_m\,(R, \mathbf{T})\,.$$

As in (39), we replace the infimization over the $\mathbf{P}_m$ by a constrained minimization and write

$$E_m = \min_{\mathbf{v}_0, \mathbf{v}_1} \left(|\mathbf{v}_0 - \mathbf{p}_{0m}|^2 + |\mathbf{v}_1 - \mathbf{p}_{1m}|^2 \right) \Big|_{\bar{\mathbf{v}}_1 \cdot F \bar{\mathbf{v}}_0 = 0} = \inf_{\boldsymbol{\delta}_{0m}, \boldsymbol{\delta}_{1m}} \left(|\boldsymbol{\delta}_{0m}|^2 + |\boldsymbol{\delta}_{1m}|^2 \right) \Big|_{\left(\overline{\boldsymbol{\delta}_{1m} + \mathbf{p}_{1m}}\right) \cdot F \left(\overline{\boldsymbol{\delta}_{0m} + \mathbf{p}_{0m}}\right) = 0}, \tag{45}$$

where now $\boldsymbol{\delta}_{1m} \equiv \mathbf{v}_1 - \mathbf{p}_{1m}$ and $\boldsymbol{\delta}_{0m} \equiv \mathbf{v}_0 - \mathbf{p}_{0m}$. We rewrite the coplanarity constraint $\left(\overline{\boldsymbol{\delta}_{1m} + \mathbf{p}_{1m}}\right) \cdot F\left(\overline{\boldsymbol{\delta}_{0m} + \mathbf{p}_{0m}}\right)$ in terms of $\mathbf{\Delta}_m$ as

$$0 = C_m + \mathbf{L}_m \cdot \mathbf{\Delta}_m + \mathbf{\Delta}_m^T Q \mathbf{\Delta}_m, \tag{46}$$

where the scalar $C_m \equiv \overline{\mathbf{p}_{1m}} \cdot (F\overline{\mathbf{p}_{0m}})$ is the deviation from coplanarity for the *observed* points $\mathbf{p}_{0m}$ and $\mathbf{p}_{1m}$, the linear coefficient is $\mathbf{L}_m \equiv \left(\left[F^T\overline{\mathbf{p}_{1m}}\right]_2; [F\overline{\mathbf{p}_{0m}}]_2\right) \in \Re^4$, and the quadratic coefficient is

$$Q \equiv \frac{1}{2}\begin{bmatrix} 0 & [[F]]_2^T \\ [[F]]_2 & 0 \end{bmatrix} \in \Re^{4\times 4}.$$

Following [29], we minimize E_m by finding the stationary points with respect to $\mathbf{\Delta}_m$ and λ_m of

$$|\mathbf{\Delta}_m|^2 + \lambda_m \left(C_m + \mathbf{L}_m \cdot \mathbf{\Delta}_m + \mathbf{\Delta}_m^T Q \mathbf{\Delta}_m\right), \tag{47}$$

where we impose the coplanarity constraint (46) via the Lagrangian parameter λ_m. Differentiating (47) with respect to Δ_m gives

$$\mathbf{\Delta}_m = -\left(1 + \lambda_m Q\right)^{-1} \lambda_m \mathbf{L}_m / 2, \tag{48}$$

which implies

$$E_m = \lambda_m^2 \mathbf{L}_m^T \left(1 + \lambda_m Q\right)^{-2} \mathbf{L}_m / 4. \tag{49}$$

Differentiating (47) with respect to λ_m gives the constraint (46). Substituting the expression (48) for Δ_m into the constraint equation, we get

$$0 = C_m - \mathbf{L}_m^T \left(1 + \lambda_m Q\right)^{-1} \lambda_m \mathbf{L}_m / 2 + \lambda_m^2 \mathbf{L}_m^T \left(1 + \lambda_m Q\right)^{-1} Q \left(1 + \lambda_m Q\right)^{-1} \mathbf{L}_m / 4. \tag{50}$$

This equation determines λ_m, and one can substitute the result into (48) and (49) to compute the final forms of Δ_m and E_m.

At this stage, we introduce approximations. Since $C_m \approx O\left(\eta\right) \ll 1$, where $O\left(\eta\right)$ represents the size of the noise, we expand the above expressions in powers of C_m. To lowest order in C_m, one can solve the constraint equation (50) by taking

$$\lambda_m = 2C_m / \left|\mathbf{L}_m\right|^2 + O\left(C_m^2\right). \tag{51}$$

One can verify that substituting this expression for λ_m into (50) makes the right side of order $O\left(C_m^2\right)$.

We write the zeroth–order solution for λ_m as $\lambda_m^{(0)} \equiv 2C_m / \left|\mathbf{L}_m\right|^2$. Substituting $\lambda_m^{(0)}$ into (49) yields

$$E_m \approx \left(\lambda_m^{(0)}\right)^2 \left|\mathbf{L}_m\right|^2 / 4 = C_m^2 / \left|\mathbf{L}_m\right|^2 \equiv E_m^{(0)}$$

to lowest order. $E_m^{(0)}$ is our desired first–order approximation and is the standard *WC* error used, e.g., in [29].

We obtain the second–order correction $\lambda_m^{(1)}$ by solving the constraint (50) to the next order. Since (51) implies $\lambda_m \approx O\left(C_m\right)$, we expand the constraint (50) up to $O\left(\lambda_m^2\right)$ as

$$0 \approx C_m - \lambda_m \left|\mathbf{L}_m\right|^2 / 2 + \lambda_m^2 \mathbf{L}_m^T Q \mathbf{L}_m / 2 + \lambda_m^2 \mathbf{L}_m^T Q \mathbf{L}_m / 4. \tag{52}$$

Substituting $\lambda_m \longrightarrow 2C_m/\left|\mathbf{L}_m\right|^2 + \lambda_m^{(1)} + O(C_m^3)$ into (52) gives

$$0 = -\lambda_m^{(1)}\left(\left|\mathbf{L}_m\right|^2/2 + O\left(C_m\right)\right) + (3/4)\left(\lambda_m^{(0)}\right)^2 \mathbf{L}_m^T Q \mathbf{L}_m + O(C_m^3),$$

which implies that we can take

$$\lambda_m^{(1)} = \left(2/\left|\mathbf{L}_m\right|^2\right)(3/4)\left(2C_m/\left|\mathbf{L}_m\right|^2\right)^2 \mathbf{L}_m^T Q \mathbf{L}_m = 6C_m^2 \mathbf{L}_m^T Q \mathbf{L}_m/\left|\mathbf{L}_m\right|^6$$

up to $O\left(C_m^3\right)$ corrections. In turn, this implies that, up to $O\left(C_m^4\right)$ corrections, $E_m - E_m^{(0)} \approx$

$$2\lambda_m^{(0)}\lambda_m^{(1)}\left|\mathbf{L}_m\right|^2/4 - 2\left(\lambda_m^{(0)}\right)^3 \mathbf{L}_m^T Q \mathbf{L}_m/4 = 2C_m^3 \mathbf{L}_m^T Q \mathbf{L}_m/\left|\mathbf{L}_m\right|^6 \equiv E_m^{(1)}.$$

Our second–order approximation to the image–plane least–squares error is

$$E_{\text{imp}} \equiv E_m^{(0)} + E_m^{(1)}.$$

The first–order *WC* error approximates the true error up to a relative factor of $E_m^{(1)}/E_m^{(0)} \approx O\left(C\mathbf{L}^T Q\mathbf{L}/\left|\mathbf{L}\right|^4\right)$ $\approx O\left(\eta\mathbf{L}^T Q\mathbf{L}/\left|\mathbf{L}\right|^4\right) \approx O\left(\eta\right)$, where η represents the size of the noise. Our second–order approximation E_{imp} approximates the true error up to a relative factor of $O\left(\eta^2\right)$.

We verified experimentally that the improved approximation E_{imp} almost always does better than E_{WC}. For the real–image experiments in Section 5, for example, the two approximations gave discrepancies from the exact error of 0.023° (improved) vs. 0.025° (*WC*), 0.225° (improved) vs. 0.285° (*WC*), 0.0204° (improved) vs. 0.0238° (*WC*), and 0.33° (improved) vs. 0.57° (*WC*). Recall that these results are for forward/backward motion estimates, for which the *WC* approximation does relatively badly. For the same image pairs in the sideways–motion region, the discrepancies were 0.002° (improved) vs. 0.005° (*WC*), 0.003° (improved) vs. 0.075° (*WC*), $\left(3.0 \times 10^{-4}\right)^\circ$ (improved) vs. $\left(4.0 \times 10^{-4}\right)^\circ$ (*WC*), and $\left(3.9 \times 10^{-6}\right)^\circ$ (improved) vs. $\left(4.9 \times 10^{-4}\right)^\circ$ (*WC*). These discrepancies represent the results of comparing the approximations to the exact error over a grid $e_x = -5:1:5$, $e_y = -5:1:5$.

References

[1] A. Chiuso, R. Brockett, and S. Soatto, "Optimal Structure from Motion: Local Ambiguities and Global Estimates," **IJCV** 39:3 195-228, 2000.

[2] R. Dutta, R. Manmatha, L.R. Williams, and E.M. Riseman, "A data set for quantitative motion analysis," *CVPR* 159-164, 1989.

[3] M. Fischler and R. Bolles, “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography,” **Comp. Assoc. Comp. Mach.** 24:6 381–395, 1981.

[4] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge, 2000.

[5] R. I. Hartley and P. Sturm, “Triangulation,” **CVIU** 68:2 146–157, 1997.

[6] R. Hartley, “Euclidean Reconstruction from Uncalibrated Views,” *Second Workshop on Invariants*, 187–202, 1993.

[7] R. Hartley, “Projective reconstruction and invariants from multiple images,”**PAMI** 16:10 1036–1041, 1994.

[8] B.K.P. Horn, “Relative orientation,” **IJCV** 4 59–78, 1990.

[9] A.D. Jepson and D.J. Heeger, “Linear subspace methods for recovering translational direction,” in *Spatial Vision in Humans and Robots*, Cambridge, 39–62, 1993.

[10] K. Kanatani, *Geometric Computation for Machine Vision*, Oxford Press, 1993.

[11] R. Kumar and A.R. Hanson, “Sensitivity of the Pose Refinement Problem to Accurate Estimation of Camera Parameters,” *ICCV* 365-369, 1990.

[12] H. C. Longuet–Higgins, “A computer algorithm for reconstructing a scene from two projections,” **Nature** 293 133–135, 1981.

[13] D. Nister, “Automatic dense reconstruction from uncalibrated video sequences,” KTH PhD thesis, 2001.

[14] J. Oliensis, “Exact two–image structure from motion,” NECI TR (extended version), 2001.

[15] J. Oliensis, “The Error Surface for Structure from Motion,” NEC TR, 2001.

[16] J. Oliensis and Y. Genc, “Fast and Accurate Algorithms for Projective Multi–Image Structure from Motion,” **PAMI** 23:6 546–559, 2001.

[17] J. Oliensis, “A New Structure from Motion Ambiguity,” *CVPR* 185–191, 1999, and **PAMI** 22:7 685–700, 2000.

[18] J. Oliensis, “A Multi-frame Structure from Motion Algorithm Under Perspective,” **IJCV** 34:2/3 163–192, 1999.

[19] J. Oliensis and Y. Genc, “New Algorithms for Two-Frame Structure from Motion,” *ICCV* 737–744, 1999.

[20] S. Soatto and R. Brocket, “Optimal Structure from Motion: Local Ambiguities and Global Estimates,” *CVPR* 282–288, 1998.

[21] S. Soatto and R. Brocket, “Optimal and Suboptimal Structure from Motion,” Harvard University TR, 1997.

[22] S. Srinivasan, “Extracting Structure from Optical Flow Using the Fast Error Search Technique,” **IJCV** 37:3 203–230, 2000.

[23] S. Srinivasan, “Fast Partial Search Solution to the 3D SFM Problem,” *ICCV* 528–535, 1999.

[24] P. Torr, D. Murray, “The Development and Comparison of Robust Methods for Estimating the Fundamental Matrix,” **IJCV** 24:3 271-300, 1997.

[25] B. Triggs, P. McLauchlan, R. Hartley, and A. Fitzgibbon, “Bundle Adjustment—A Modern Synthesis,” *Workshop on Vision Algorithms*, 1999.

[26] R. Vidal, Y. Ma, S. Shu, S. Sastry, “Optimal Motion Estimation from Multiview Normalized Epipolar Constraint,” *ICCV* I 34-41, 2001.

[27] J. Weng, T. S. Huang, N. Ahuja, "Motion and structure from two perspective views: algorithms, error analysis, and error estimation," **PAMI** 11:5 451–476, 1989.

[28] Zhengyou Zhang, "On the optimization criteria for two–frame structure from motion," **PAMI** 20:7 717-729, 1998.

[29] Zhengyou Zhang, "Understanding the Relationship Between the Optimization Criteria in Two–View Motion Analysis," *ICCV* 772–777, 1998.