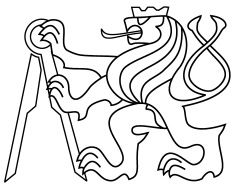
CENTER FOR
MACHINE PERCEPTION



CZECH TECHNICAL
UNIVERSITY IN PRAGUE

BACHELOR THESIS

Robust Focal Length Computation

Oleh Rybkin

rybkiole@fel.cvut.cz

May 22, 2017

Available at
WritethefullURLofthepaperhere,ifavailableintheelectronicform.

Thesis Advisor: Ing. Tomáš Pajdla, PhD.

Acknowledge grants here. Use centering if the text is too short.

Center for Machine Perception, Department of Cybernetics
Faculty of Electrical Engineering, Czech Technical University
Technická 2, 166 27 Prague 6, Czech Republic
fax +420 2 2435 7385, phone +420 2 2435 7637, www: http://cmp.felk.cvut.cz

Robust Focal Length Computation

Oleh Rybkin

May 22, 2017

Text of acknowledgements...

Author's declaration

I declare that I have work out the presented thesis independently and that I have listed all information sources used in accordance with the Methodical Guidelines about Maintaining Ethical Principles for Writing Academic Theses.

Prohlášení autora práce

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

V Praze dne ....

Podpis autora práce

Text of abstract...

Contents

List of Symbols and Abbreviations

x	Scalar value.
$\mathbf{x}$	Column vector.
$\mathtt{X}$	Matrix.
X	Set.
$\mathbb{C}$	Complex numbers field.
$\mathbb{R}$	Real numbers field.
$\mathrm{K}[x_1, x_2, x_3, \ldots, x_n]$	Polynomial ring in variables $x_1, x_2, x_3, \ldots, x_n$ over the field K.
$\mathbb{R}^n$	Linear space of dimension n over real numbers.
$\mathbb{P}^{n-1}$	Projective space of dimension $n-1$ over real numbers.
$\mathbf{x} \otimes \mathbf{y}$	Kronecker (outer) product of vectors $\mathbf{x}$,$\mathbf{y}$.
$[x]_\times$	Cross product matrix, i.e. a matrix multiplication by which represents a cross product with vector x.
$\det \mathtt{X}$	Determinant of the matrix $\mathtt{X}$.
$\operatorname{diag} \mathbf{x}$	Diagonal matrix that have elements of $\mathbf{x}$ at its diagonal.
$\operatorname{vec} \mathtt{X}$	Vector produced by stacking columns of the matrix $\mathtt{X}$.
$\mathtt{I}$	Identity matrix.
$\mathtt{0}_{x,y}$	Zero matrix from $\mathbb{R}^{x \times y}$.
$\mathcal{N}(\mu, \sigma^2$	Gauss distribution with mean μ and standard deviation σ.

List of Figures

1 Introduction

We analyze the problem of computing epipolar geometry of two partially calibrated cameras, where only focal lengths are unknown.

Estimating epipolar geometry with unknown focal lengths is an important issue in practical problems. In laboratory environment it is possible to calibrate the camera beforehand using established procedures [8]. When taking images in the wild, however, it is often impractical or impossible to use these procedures. It is desirable to have an automated procedure to estimate camera calibration from images themselves.

The usual way [23, 22, 20] to estimate camera external and internal parameters is to extract points of interests [2, 18] with tentative correspondences, and use RANSAC [**?**] method for joint estimation of correspondence inlier pairs and camera parameters. In RANSAC, a procedure to compute camera parameters from a (preferably small) number of points is needed. These procedures, that essentially are used as black box in RANSAC, are called minimal problems because it is desirable to find a procedure that would use the theoretical minimum number of points.

The basic procedure to compute the focal lengths and epipolar geometry given correspondences consists of two steps: finding a Fundamental matrix and decomposing the matrix into Calibration matrices and Essential matrix. For the first task, the minimal number of correspondences needed is 7. Hartley and Zisserman [8] describe an algorithm for this which uses 7 correspondences as well as the singularity condition. The algorithm gives three different estimations of the matrix, i.e. one correct and two false ones. Hartley [10] also summarizes an algorithm by Longuet-Higgins [17] which uses 8 correspondences instead. Bougnoux [3] gives a concise formula to compute focal length from the fundamental matrix, if the rest of calibration information, i.e. i.e. the principal point, the skew, the ratio of the pixel dimensions, is known.

The procedure suffers from a number of known failure cases:

- It is not possible to determine a Fundamental matrix if the correspondences are in singular position.
- The camera pair may have intersecting (or parallel) optical axes. In that case it is impossible to determine the focal lengths.
- The plane defined by the baseline and the optical axis of one camera may be perpendicular to the plane defined by the baseline and optical axis of the other camera. In this case no focal length can be recovered as the Bougnoux formula fails.
- All computed Fundamental matrices may have rank 1 or be complex.
- A computed focal length may be complex.
- There may be no such computed camera configuration where most points in 3D space lie in front of the cameras.

Because of these deficiencies, the commonly held view is that the algebraic approach of using 7pt algorithm and Bougnoux formula is not robust, and sometimes fails entirely, unable to find any valid solution. Hartley [9] argues that in many practical cases this approach cannot be used and other methods [9, 25, 13] were proposed to solve this problem.

We analyze the procedure of computing focal lengths from points correspondences TP: of what? and different degeneracies to show that slight modifications of the procedure allow to alleviate a number of the problems.

2 Basic notions

2.1 Camera geometry

The book [8] describes camera geometry, the main topic of this thesis, in chapter 6. All needed preliminaries are also well described in chapters 1-5 of the book. In the next two sections we very briefly summarize and recapitulate the topic. We follow the notation of [8].

2.1.1 Single camera geometry

The most convenient way of representing points in 3D or 2D space while dealing with camera geometry is using corresponding projective spaces.

Definition 2.1.1. Given a linear vector space $\mathbb{R}^n$, a *projective space* $\mathbb{P}^{n-1}$ could be defined as a factorization of the linear space by a relation $\sim$:

$$\mathbf{x}_1 \sim \mathbf{x}_2 \iff (\exists \lambda \in \mathbb{R}\backslash\{0\}\ \mathbf{x_1} = \lambda \mathbf{x_2}).$$

Zero point is also excluded from projective space.

A point $\mathbf{X}_i = \begin{bmatrix} x & y & z \end{bmatrix}^{\mathrm{T}}$ from 3D linear space $\mathbb{R}^3$ can be though of as a point $\mathbf{X}' = \begin{bmatrix} x & y & z & 1 \end{bmatrix}^{\mathrm{T}}$ from 3D projective space $\mathbb{P}^3$. Similarly, a point $\mathbf{X}' = \begin{bmatrix} x & y & z & f \end{bmatrix}^{\mathrm{T}}$ can be converted back to the point $\mathbf{X} = \begin{bmatrix} \frac{x}{f} & \frac{y}{f} & \frac{z}{f} \end{bmatrix}^{\mathrm{T}}$ unless f is zero. Because of this possible conversion we will frequently refer to the same geometrical point as belonging to projective space $\mathbb{P}^n TP : -1$ as well as linear space $\mathbb{R}^n$ simultaneously.

Definition 2.1.2. A *projective line* is a line of points in projective space $\mathbb{P}^n$. Conveniently, it can be also expressed by a vector from the space $\mathbb{P}^n TP : +1$. For example, a line in projective plane $ax + by + fc = 0$, which corresponds to the line $ax + by + c = 0$ in real plane, can be written as a vector $\mathtt{l} = \begin{bmatrix} a & b & c \end{bmatrix}^{\mathrm{T}}$.

The figure 2.1 shows an essential concept of projecting points from world (3D) space to image (2D) space. This projection can be described as a linear operation in corresponding projective spaces.

The projection operator is given by camera parameters, which can be divided in two groups - extrinsic and intrinsic parameters. We next define essential intrinsic parameters.

Definition 2.1.3. *The principal axis*, also *optical axis*, is the line passing through camera center and perpendicular to the image plane.

Definition 2.1.4. *The principal point* $\mathbf{p} = \begin{bmatrix} p_x & p_y & 1 \end{bmatrix}^{\mathrm{T}}$ is the point lying at the intersection of image plane and principal axis.

Definition 2.1.5. *The focal length* f is the distance from the camera center to the principal point.

Camera intrinsic parameters are can be formed into camera calibration matrix.

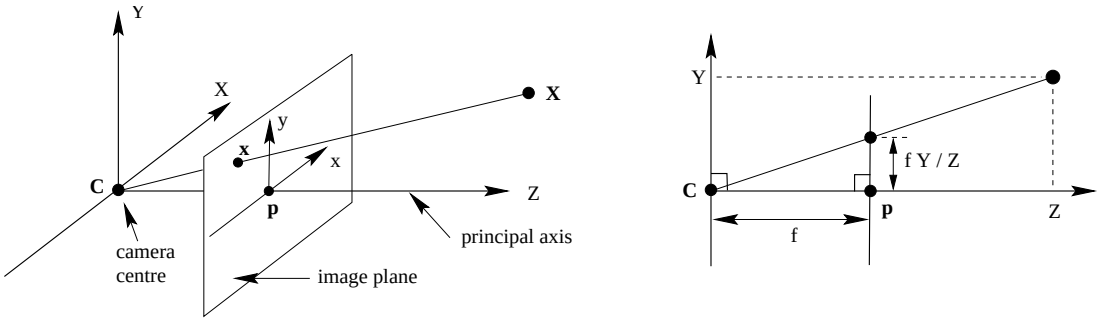


Figure 2.1 Camera model. Courtesy [8]

Definition 2.1.6. *A calibration matrix*[1] $\mathtt{K}$ is a matrix of the form

$$\mathtt{K} = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix}.$$

TP: this is not clear. The matrix is not really defined up to scale when you write it with $K_{33} = 1$.

If we assume that the camera center is the world space zero point, principal axis coincide with Z axis, and image space axes x, y are aligned with world space axes X, Y (true for the figure 2.1), we can project 3D points with only camera intrinsics. The projection from $\mathbb{R}^3$ to $\mathbb{P}^2$ image plane is given by $\mathbf{x}_i = \mathtt{K}\mathbf{X}_i$.

Extrinsic parameters of a camera i are described by a rotation matrix in the world space $\mathtt{R}_i$ describing camera orientation and the camera center point $\mathbf{C}_i$ in the world space. Given extrinsic and intrinsic parameters, we can construct the full projection matrix.

Definition 2.1.7. The camera *projection matrix* $\mathtt{P}_i$ can be constructed from camera parameters and projects points $\mathbf{X}_i$ from world space $\mathbb{P}^3$ to the image plane $\mathbb{P}^2$:

$$\mathbf{x}_i = \mathtt{P}\mathbf{X}_i = \mathtt{KR}[\mathtt{I}| - \mathbf{C}]\mathbf{X}_i$$

.

2.1.2 Epipolar geometry

Epipolar geometry, one of the key concepts TP: Not true. The main topic is robust computtion of f. of this thesis, is a geometry of two cameras seeing the same 3D object.

This section is described in depth in chapter 9 of [8]. As a slight deviation from the notation of Hartley and Zisserman, cameras and respective image points are indexed with numbers, e.g., $\mathbf{x}_1$, $\mathbf{x}_2$ and not $\mathbf{x}$, $\mathbf{x}'$ respectively.

Extrinsic parameters of a camera pair can be concisely given by the translation vector between camera centers $\mathbf{t} = \mathbf{C}_1 - \mathbf{C}_2$ and the rotation matrix between cameras' coordinate systems $\mathtt{R} = \mathtt{R}_2\,\mathtt{R}_1^T$. If $\mathtt{R}, \mathbf{t}$ are known, extrinsics $\mathtt{R}_1, \mathtt{R}_2, \mathbf{C}_1, \mathbf{C}_2$ are defined up to a projective transformation.

Geometry of two cameras and a point is usually described in terms of epipoles, the key concept of epipolar geometry.

[1] Throughout this work we assume unity aspect ratio and zero skew.

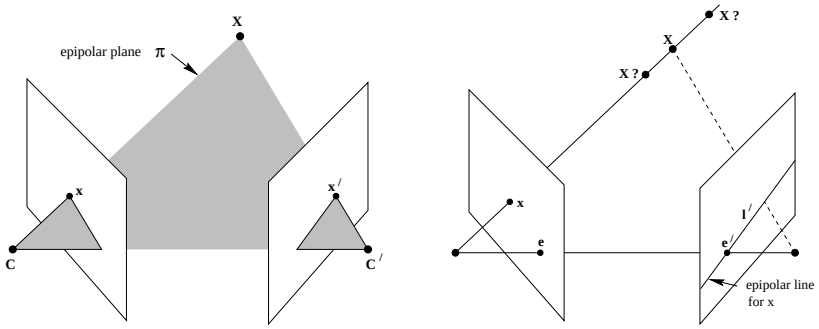


Figure 2.2 Camera pair. Courtesy [8]

Definition 2.1.8. *The epipole* $\mathbf{e}_i$ is the point in image plane that lies at the line joining camera centers. Equivalently, it is the point to which the center of another camera projects.

Definition 2.1.9. *An epipolar line* $\mathbf{l}$ is a line in image plane on which the epipole of that plane lies.

Definition 2.1.10. *An epipolar plane* π is a plane in the world space which contains the line joining camera centers. The plane is associated with two epipolar lines that it also contains.

Definition 2.1.11. *The baseline* is the line in 3D joining two camera centers.

The information about camera epipolar geometry can be described by a single matrix as defined next. If interested, see extended discussion and proofs in [8].

Definition 2.1.12. *An Essential matrix* $\mathtt{E}$ is a matrix of the form $\mathtt{E} = \mathtt{R}[\mathbf{t}]_\times$. The essential matrix is defined up to overall scaling.

As can be seen from its construction, an essential matrix can only be of rank 2. A weaker version of this constraint can be stated as the rank constraint:

$$\det \mathtt{E} = 0. \tag{2.1}$$

The two non-zero singular values of an essential matrix are equal. This constraint can be stated as the next matrix equation (also called the trace constraint):

$$2\mathtt{E}\mathtt{E}^{\mathrm{T}}\mathtt{E} - \operatorname{trace}(\mathtt{E}\mathtt{E}^{\mathrm{T}})\mathtt{E} = 0. \tag{2.2}$$

The equations 2.1 and 2.2 together are named Demazure polynomials [19].

Essential matrix does not capture intrinsic camera parameters. To that end, we use the fundamental matrix.

Definition 2.1.13. *The Fundamental matrix* $\mathtt{F}$ corresponding to an Essential matrix $\mathtt{E}$ is given by

$$\mathtt{F} = \mathtt{K}_2^{-\mathrm{T}}\mathtt{E}\mathtt{K}_1^{-1}. \tag{2.3}$$

The fundamental matrix is, like the essential matrix, defined up to overall scaling.

The fundamental matrix of a camera pair defines a specific mapping from points in one image to epipolar lines in another image. Suppose that given an image point $\mathbf{x}_1 \in \mathbb{P}^3$ and two cameras $\mathtt{P}_1$, $\mathtt{P}_2$, the line $\mathbf{C}_1\mathbf{x}_1$ in world space is projected by the second camera to the line $\mathbf{l}_2 \in \mathbb{P}^3$ in the second camera image space. Then we can obtain $\mathbf{l}_2$ by $\mathbf{l}_2 = \mathtt{F}\mathbf{x}_1$

For each corresponding image points pair $\mathbf{x}_1, \mathbf{x}_2$, the Fundamental matrix satisfies the so-called Epipolar constraint:

$$\mathbf{x}_2^{\mathrm{T}}\mathtt{F}\mathbf{x}_1 = 0. \tag{2.4}$$

As with essential matrices, a fundamental matrix can only be of rank 2. On the other hand, every matrix of rank 2 is a fundamental matrix [**?**]. A fundamental matrix also satisfies the rank constraint

$$\det \mathtt{F} = 0. \tag{2.5}$$

A version of trace constraint on fundamental matrix can be formulated by substituting equation 2.3 into equation 2.2.

2.2 Computing focal lengths from point correspondences

Given a set of points in one image X_1 and a set of points in another image X_2 that correspond to the same points X in world space, it is possible in some cases to determine the positions of points X and camera intrinsic and extrinsic parameters. This section guides the reader through the process.

Note, that in this part we assume that the point correspondences are given to us by a black box algorithm. In fact, many algorithms and modifications to them were developed to this end, the most notable being the RANSAC family of algorithms [6].

2.2.1 Fundamental matrix computation

First, we need to estimate the fundamental matrix $\mathtt{F}$ of a camera pair. In fact, all the information we need is contained in this matrix. Various algorithms exist that can compute the matrix $\mathtt{F}$ from point correspondences. We present two methods that are, to our best knowledge, used the most. They are relatively simple and can be implemented efficiently.

The algorithms which we will present belong to a family of so-called algebraic solvers, which work by way of constructing a set of constraints on the matrix $\mathtt{F}$ given the correspondences, and then solving the constraints in an algebraically precise way, also giving all possible solutions. If the constraints are represented by polynomial equations TP: delete (which, as we will see, they are), the techniques of algebraic geometry can be used to solve these constraints.

We first define the 7pt algorithm. The algorithm uses (at least) 7 epipolar constraints 2.4 and the rank constraint 2.5. This gives us 8 constraints on a 3×3 matrix. Note, however, that all of the constraints are homogeneous polynomials, and thus the solution to the system will be a finite number[2] of linear one-dimensional spaces of matrices. As a Fundamental matrix is defined up to scaling, this corresponds to a finite number of Fundamental matrices. Moreover, the number of matrices can be expected to be no more than 3, because of the constraints used[3].

[2] We suppose that the constraints are algebraically independent. Equivalently, at least 7 of the correspondence point pairs are in general position.

[3] This number is calculated as the product of degrees of the polynomials $= 3 \cdot 1^7$

The algorithm uses SVD procedure which serves as implicit optimization of epipolar constraints. In case that precisely 7 epipolar constraints are given, SVD can be replaced with Gauss-Jordan elimination, and the $\mathbf{f}_1$, $\mathbf{f}_2$ vectors will be a basis of right null space of the matrix $\mathtt{B}$. Equivalently, the matrices $\mathtt{F}_1$, $\mathtt{F}_2$ will be a basis of a space of matrices that satisfy the 7 epipolar constraints.

Data: list of $n \geq 7$ right image points $\mathbf{x}_{1,i}$, list of the corresponding left image points $\mathbf{x}_{2,i}$
Result: Fundamental matrix $\mathtt{F}$
begin
  Populate matrix $\mathtt{B} \in \mathbb{R}^{n\times 9}$ with columns $\mathbf{b}_i$: $\mathbf{b}_i \leftarrow \mathrm{vec}(\mathbf{x}_{2,i} \otimes \mathbf{x}_{1,i})$;
  TP: udefined $\otimes$
  Take the right singular vectors $\mathbf{f}_1, \mathbf{f}_2$ corresponding to the two smallest singular values and transform them back to matrices $\mathtt{F}_1, \mathtt{F}_2 \in \mathbb{R}^{3\times 3}$;
  if $\det(\mathtt{F}_2) = 0$ **then**
    return $\mathtt{F}_2$;
  else
    Solve the 3rd degree polynomial in x: $\det(x\mathtt{F}_1 + \mathtt{F}_2) = 0$, and choose the real roots x_i.;
    return $x_i\mathtt{F}_1 + \mathtt{F}_2$;
  end
end

Algorithm 1: 7pt

Note, that we call this algorithm 7pt because the minimal number of correspondences it needs is 7. Nevertheless, any number bigger than 7 can also be used.

The second algorithm doesn't use the rank constraint 2.5. Therefore, generically we expect the result to be a matrix of rank 3 and not a valid Fundamental matrix. The matrix, however, 'better' satisfies the epipolar constraints 2.4. We will compare the performance of that algorithm to 7pt in later parts of the work. Note that this algorithm gives precisely one real solution.

Data: list of $n \geq 8$ right image points $\mathbf{x}_{1,i}$, list of the corresponding left image points $\mathbf{x}_{2,i}$
Result: Fundamental matrix $\mathtt{F}$
begin
  Populate matrix $\mathtt{B} \in \mathbb{R}^{n\times 9}$ with columns $\mathbf{b}_i$: $\mathbf{b}_i \leftarrow \mathrm{vec}(\mathbf{x}_{2,i} \otimes \mathbf{x}_{1,i})$;
  Take the right singular vector $\mathbf{f}$ corresponding to the smallest singular value and transform it back to matrix $\mathtt{F} \in \mathbb{R}^{3\times 3}$;
  return $\mathtt{F}$;
end

Algorithm 2: 8pt

Note, that we call this algorithm 8pt because the minimal number of correspondences it needs is 8. Nevertheless, any number bigger than 8 can also be used.

2.2.2 Essential matrix computation

If the two focal lengths and the principal points are known TP: all other calibration has to be known too, we can estimate essential matrix $\mathtt{E}$ instead of $\mathtt{F}$. This can be done by transforming points so that effects of intrinsic camera parameters are removed:

$$\tilde{\mathbf{x}}_i = \mathtt{K}_i^{-1}\mathbf{x}_i.$$

Then the epipolar constraint 2.4 is satisfied by the essential matrix with respect to the transformed points:

$$\tilde{\mathbf{x}}_2^{\mathrm{T}}\mathtt{E}\tilde{\mathbf{x}}_1 = \mathbf{x}_2^{\mathrm{T}}\mathtt{K}_2^{-\mathrm{T}}\mathtt{E}\mathtt{K}_1^{-1}\mathbf{x}_1 = \mathbf{x}_2^{\mathrm{T}}\mathtt{F}\mathbf{x}_1 = 0.$$

The algebraic solver for essential matrix uses the rank constraint 2.1, the trace constraint 2.2, and needs 5 point correspondences. Note that the trace constraint is a matrix equation, and therefore actually stands as more than one constraint, which explains why counting degrees of freedom seemingly fails in this case.

The construction of the solver is considerably more involved than the previous algorithms, therefore we will only refer to the papers that fully describe it.

Nister et al. [21] TP: thre were othe ralgortihms presented before, e.g. by Nister, ... show in detail how to construct an algebraic solver for this problem. The paper makes extensive use of theory of algebraic geometry [5][4]. It is very instructing to read through the construction of this solver.

We will, however, use another solver, described by Kukelova et al. in [16]. The code can be found online at "http://cmp.felk.cvut.cz/mini/"[4]. This solver, contrary to one of Stewenius et al., makes it possible to use an arbitrary number of correspondences, which we will take advantage of. We will refer to this solver as *5pt solver*.

For those interested, Kukelova et al. give an algorithm (with source code) that automatically creates solvers for algebraic problems [15].

2.2.3 Focal lengths computation

Bougnoux [3] gives a concise formula to compute focal length from the fundamental matrix, if principal point position is known. In the next two formulae the matrix $\mathtt{I}_2$ is defined as $\mathtt{I}_2 = \mathrm{diag}(1, 1, 0)$, and $\mathbf{e}_1, \mathbf{e}_2$ are the epipoles of first and second image correspondingly.

$$f_1^2 = -\frac{\mathbf{p}_2^{\mathrm{T}}\left[\mathbf{e}_2\right]_\times \mathtt{I}_2\mathtt{F}\mathbf{p}_1\mathbf{p}_1^{\mathrm{T}}\mathtt{F}^{\mathrm{T}}\mathbf{p}_2}{\mathbf{p}_2^{\mathrm{T}}\left[\mathbf{e}_2\right]_\times \mathtt{I}_2\mathtt{F}^{\mathrm{T}}\mathtt{I}_2\mathtt{F}\mathbf{p}_2} \tag{2.6}$$

$$f_2^2 = -\frac{\mathbf{p}_1^{\mathrm{T}}\left[\mathbf{e}_1\right]_\times \mathtt{I}_2\mathtt{F}^{\mathrm{T}}\mathbf{p}_2\mathbf{p}_2^{\mathrm{T}}\mathtt{F}\mathbf{p}_1}{\mathbf{p}_1^{\mathrm{T}}\left[\mathbf{e}_1\right]_\times \mathtt{I}_2\mathtt{F}\mathtt{I}_2\mathtt{F}^{\mathrm{T}}\mathbf{p}_1} \tag{2.7}$$

The second equation may be derived from the first by changing indices and transposing the $\mathtt{F}$ matrix.

Note that it is not possible to find a focal length unless the principal point is known. As these formulae show, TP: delet generically the focal lengths are different for different principal point positions.

2.2.4 Failure cases

We consider the failure cases that may occur while estimating focal length using 7pt algorithm and the Bougnoux formula.

Degeneracies

The first possible degeneracy is that the 7 epipolar constraints 2.4 are *linearly dependent*. This means that the system is underconstrained and there will be generically an infinite number of solutions. A configuration of 7 correspondences which lead to such a system is said to be in singular position.

[4] Enter 'Polynomial eigenvalue solutions to minimal problems in computer vision' in the search field.

Agarwal et al., in a thorough study [1], analyze using algebraic geometry TP: reorder the words to get a valid English sentence a degeneration when neither of 3 solutions that 7pt gives are actually valid Fundamental matrices . It is shown in the paper that meeting the rank constraint is not enough for a matrix to be a fundamental matrix, as it needs to be of rank precisely 2. There may be a degeneracy when the matrix *is of rank 1*[5]. However, the constraint "is not of rank 1" cannot be expressed algebraically in an easy way and doesn't reduce the number of degrees of freedom of $\mathtt{F}$, as it only affects an infinitely small fraction of matrices. The preferable way of dealing with this degeneracy is thus to select only valid subset of 3 solutions given by 7pt algorithm.

Another degeneracy may occur when the camera pair has *intersecting (or parallel) optical axes*. In that case it is impossible to determine the focal lengths [12]. We now formulate a concise condition which describes when this can happen.

Lemma 2.2.1. *Optical axes of cameras* $\mathtt{P}_1, \mathtt{P}_2$ *as lines in projective* $\mathbb{P}^3$ *space intersect*[6] *if and only if the principal points* $\mathbf{p}_1$, $\mathbf{p}_2$ *satisfy the epipolar constraint*

$$\mathbf{p}_2^T \mathtt{F} \mathbf{p}_1 = 0.$$

Moreover, if the principal points are zero points $\mathbf{p}_1 = \mathbf{p}_2 = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$, *the optical axes intersect if and only if the rightmost lowest element of the corresponding fundamental matrix* $\mathtt{F}$ *is zero*

$$\mathtt{F}_{3,3} = 0$$

.

Proof. The first part. If optical axes intersect, the point at their intersection is projected to images' principal points. Thus, principal points form a correspondence and must satisfy the epipolar constraint. Conversely, if the principal points satisfy the epipolar constraint, they must lie in an epipolar plane (see figure 2.2). The optical axes are then also lying in the epipolar plane, and each two projective lines in a plane intersect.

The second part. Consider the following:

$$\mathbf{p}_2^{\mathrm{T}} \mathtt{F} \mathbf{p}_1 = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathtt{F}_{1,1} & \mathtt{F}_{1,2} & \mathtt{F}_{1,3} \\ \mathtt{F}_{2,1} & \mathtt{F}_{2,2} & \mathtt{F}_{2,3} \\ \mathtt{F}_{3,1} & \mathtt{F}_{3,2} & \mathtt{F}_{3,3} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathtt{F}_{1,3} \\ \mathtt{F}_{2,3} \\ \mathtt{F}_{3,3} \end{bmatrix} = \mathtt{F}_{3,3}$$

□

A further degeneracy [8] occurs when the plane defined by the baseline and the optical axis of one camera *is perpendicular* to the plane defined by the baseline and optical axis of the other camera. In this case the Bougnoux formulae have zero denominator and thus fail to estimate the focal lengths.

Invalid solutions

There are also cases when the matrix and the focal lengths can be computed but don't correspond to any real camera configuration.

Some solutions that may need to be filtered out are *complex* $\mathtt{F}$ *matrices*. In the 7pt algorithm, as we are solving a 3rd degree polynomial, we may end up with 2 complex and one real solution TP: reorder words to make it a valid E sentence. This might be

[5] We suppose that algorithm does not return zero solution $\mathtt{F} = 0_{3,3}$. There indeed exists a formulation in which such solution naturally does not come up.

[6] In projective space we say that two parallel lines intersect at a point at infinity.

a problem if the real solution `F` is a matrix on rank 1, and thus we don't have any valid estimate.

Hartley in a work [9] shows that the focal lengths computed may be TP: all *complex*, which again leads to impossible camera configuration. This happens because the focal length enter the Bougnoux equation 2.6 only squared, so a negative estimation of f^2 leads to a purely imaginary focal length and can occur even when all the coefficients are real. Hartley argues that this is unacceptable and addresses the problem by developing a new algorithm to find a matrix which will have a valid focal length. We will address the degeneracy in this work later, and provide a clear and concise way to correct such a matrix.

Another problem is described in conjunction with the notion of *cheirality* in [11]. In the projective camera model, we allow the points from behind the camera to be projected to image plane. This introduces an ambiguity when decomposing the essential matrix into rotation and translation. Four different camera pairs can be TP: constructed imagined TP: , never comes before "that" in this way. This is a slavism., which will have the same essential matrix, corresponding to two possible orientations of each camera, one that has points in front of it and one that has points behind.

Normally, we are able to pick the right reconstruction by selecting the camera pair such that cameras have points in front of them in world space. However, with presence of large noise it can happen that no such camera pair will exist. Usually the configuration that has the points in front of it is chosen TP: delete , but this strategy can produce a wrong configuration.

2.3 Algebraic Geometry

This section serves as a brief introduction of the topic of algebraic geometry. We follow the notation of Cox et al. from [5][4]. We refer an interested reader to these books.

First we define the key algebraic object of the algebraic geometry.

Definition 2.3.1. Given a polynomial ring $\mathbb{C}[x_1, x_2, \ldots, x_n]$, *an ideal* I is such a subset of the ring that is closed under addition as well as under multiplication by a polynomial from $\mathbb{C}[x_1, x_2, \ldots, x_n]$. Specifically,

$$\forall f_1, f_2 \in I \ : \ f_1 + f_2 \in I$$
$$\forall f \in I, h \in \mathbb{C}[x_1, x_2, \ldots, x_n] \ : \ hf \in I.$$

We can generate a geometrical object from an ideal, as outlined in the next definition.

Definition 2.3.2. We define a function $\mathbf{V}(I) = V$, where $I \subset \mathbb{C}[x_1, x_2, \ldots, x_n]$, and $V \subset \mathbb{C}^n$.

$$\mathbf{V}(I) = \{a \in \mathbb{C}^n \mid f(a) = 0 \text{ for all } f \in I\}$$

Intuitively, V is a set of solutions to a (possibly infinite) system of polynomial equations I.

The geometrical sets that can be constructed in such a way are called varieties.

Definition 2.3.3. *A variety* V is such a set $V \subset \mathbb{C}^n$ that $\mathbf{V}(I) = V$ for some system of polynomial equations $I \subset \mathbb{C}[x_1, x_2, \ldots, x_n]$.

We can also generate an ideal back from a variety.

Definition 2.3.4. We define a function $\mathbf{I}(V) = I$, where $V \subset \mathbb{C}^n$, and $I \subset \mathbb{C}[x_1, x_2, \ldots, x_n]$.

$$\mathbf{I}(V) = \{f \in \mathbb{C}[x_1, x_2, \ldots, x_n] \mid f(a) = 0 \text{ for all } a \in V\}.$$

Intuitively, I is a maximal set of polynomial equations that determines the geometrical set V.

A known result [5] is that $\mathbf{V}(\mathbf{I}(V)) = V$. Under a certain assumption[7], the opposite is also true $\mathbf{I}(\mathbf{V}(I)) = I$.

An important notion of Gröbner bases is somewhat analogical to linear bases from linear algebra. Defining Gröbner bases requires more involved terms, so we won't do it here. Instead, we state an important result, which offers intuitive understanding of them.

Definition 2.3.5. If G is *a Gröbner basis* $G = \{g_1, g_2, \ldots, g_m\} \subset I$ of an ideal $I \subset \mathbb{C}[x_1, x_2, \ldots, x_n]$, then:

$$I = \langle g_1, g_2, \ldots, g_m \rangle = \{\sum_{i=1}^{m} h_i g_i \mid h_1, h_2, \ldots, h_m \in \mathbb{C}[x_1, x_2, \ldots, x_n]\}$$

An important result [5] states that every ideal defined on a finite-dimensional polynomial ring has a finite Gröbner basis. Continuing the analogy with linear bases, a Gröbner basis of a set of polynomials of degree 1 (i.e., linear equations) system is indeed its linear basis.

TP: Written very well :-)

[7] When the ideal concerned is a so-called radical ideal [5].

3 Related work

Hartley and Silpa-Anan [9] develop an iterative algorithm which incorporates heuristic estimates (prior knowledge) of focal lengths. They also show that by allowing the algorithm to move principal point TP: , a better results can be obtained. The algorithm shows competitive, although not decisively better performance in comparison to 7pt approach.

Nakatsuji et al. [14] show that even when the two focal lengths are known to be the same, the 7pt algorithm (which assumes they are different) yields better accuracy than methods which assume the same focal length. When the 7pt algorithm with the Bougnoux formula fail to produce a real focal length, he authors use 'subsampling' procedure. Points are subsampled from the set of inliers and focal length is recomputed from the sample each time until a real focal length is found.

Kanazawa et al. [13] use three views for computing camera parameters (only two-view correspondences are needed for the method). Exploiting three different view pairs they are able to give more stable results than two-view methods.

DeepFocal, a recent algorithm by Scott Workman et al. [25], uses a Convolutional Neural Network to estimate the focal length directly from one image. The neural network outperforms other approaches based on one view. While an interesting and fresh idea, authors didn't compare DeepFocal to any existing two-view approaches, therefore it is difficult to assess the advantages of the work.

4 Experimental analysis of existing solutions

We assess the quality and stability of the 7pt algorithm with the Bougnoux formula. We consider this pipeline as essentially integral TP: baseline procedure for estimation of the focal lengthTP: s from image correspondences.

4.1 Experimental setup

We assume zero skew and unity aspect ratio in both cameras and consider a partially calibrated problem, i.e., the principal points are known. We assume that the images were preprocessed TP: not clear what "translated" means in a way that principal points are always in the center. This means that the calibration matrices after this preprocessing are of shape

$$\mathtt{K}_i = \begin{bmatrix} f_i & 0 & 0 \\ 0 & f_i & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The cameras are allowed to have different focal lengths.

4.1.1 Evaluation procedure

The setup for most experiments in the work is as follows:

Data: TP: The Number of correspondences used TP: is equal to n, level of noise σ TP: "level" is not precise enough". What is σ?

1 **begin**
2 Create a camera pair $\mathtt{P}_1$, $\mathtt{P}_2$ with focal lengths f_1,f_2;
3 Create a point cloud $\mathtt{X} \in \mathbb{R}^{3\times n}$ of n points in 3D ;
4 Project $\mathtt{X}$ to the cameras $\mathtt{P}_1$ and $\mathtt{P}_2$, producing the ground truth image point sets $\mathtt{x}_1^{gt}$, $\mathtt{x}_2^{gt} \in \mathbb{R}^{2\times n}$;
5 Apply additive noise drawn from $\mathcal{N}(0, \sigma^2)$ to both x, y coordinates of each point from $\mathtt{x}_1^{gt}$ and $\mathtt{x}_2^{gt}$. This produces observed image point sets $\mathtt{x}_1^{ob}$ and $\mathtt{x}_2^{ob}$ correspondingly ;
6 Select 7 pairs of corresponding points $\mathtt{x}_1$ and $\mathtt{x}_2$, and let the remaining pairs to be the test set $\mathtt{x}_1^{test}, \mathtt{x}_2^{test}$;
7 Get the $\mathtt{F}_i$ estimates of the fundamental matrix $\mathtt{F}$ from the correspondences $\mathtt{x}_1, \mathtt{x}_2$ (e.g. $i = 1 \dots 3$ when using the 7pt algorithm 1). Leave only matrices with real elements;
8 Choose the estimation that optimizes Sampson error 4.1 on the test set;
9 Estimate the focal lengths from the fundamental matrix $\mathtt{F}$ using the formulae 2.6 and 2.7;
10 **end**

Algorithm 3: Workflow

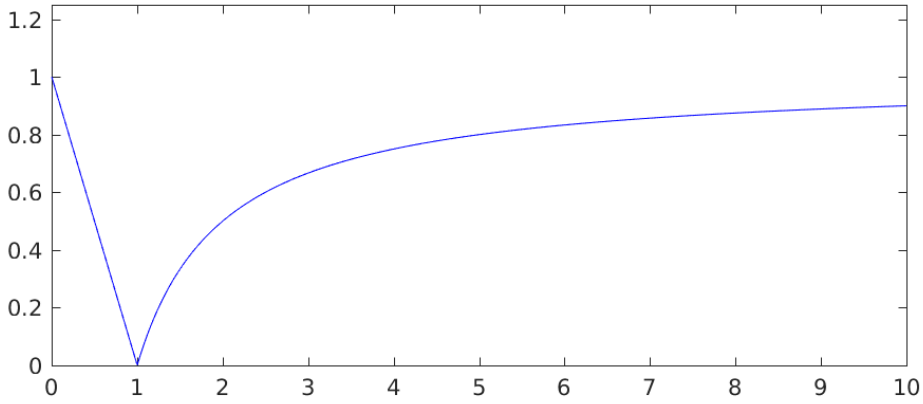


Figure 4.1 Error function for $f_{gt} = 1$

4.1.2 Estimation quality measure

We treat error in focal length estimation as multiplicative, so that the error in case where the ground truth $f_{gt} = 1000$ and computed value $f_{ob} = 2000$ is the same as error in case where the ground truth $f_{gt} = 100$ and computed value $f_{ob} = 200$. More precisely the error is given by

$$e(f_{gt}, f_{ob}) = 1 - \frac{\min(f_{gt}, f_{ob})}{\max(f_{gt}, f_{ob})}.$$

The error function for fixed ground truth value $f_{gt} = 1$ is shown at the figure 4.1.

For measuring the quality of a fundamental matrix we use the Sampson error:

$$S(\mathtt{F}) = \sum_i \frac{(\mathbf{x}_2^{\mathrm{T}}\mathtt{F}\mathbf{x}_1)^2}{(\mathtt{F}\mathbf{x}_1)_1^2 + (\mathtt{F}\mathbf{x}_1)_2^2 + (\mathbf{x}_2^{\mathrm{T}}\mathtt{F})_1^2 + (\mathbf{x}_2^{\mathrm{T}}\mathtt{F})_2^2} \tag{4.1}$$

4.2 Performance analysis for generic situations

In this section we suppose a generic scene, in which a degeneracy is unlikely to occur.

In experiments, we generate a random set of points and a random camera pairTP: s so that the set of correspondences in each image span at least 1000 × 1000 pixel square.

4.2.1 Overall analysis

We analyze the behavior of the pipeline against different numbers of given correspondences and different levels of noise. The quality of the estimation is assessed by computing the mean TP: of what? and two TP: which? quantiles of focal lengths computation error e. In these experiments we discard each all imaginary estimates.

The graph 4.2 shows TP: a rapid decline in error with bigger TP: growing number of correspondences. Using 20 correspondences we can achieve an error close to negligible.

The graph 4.3 shows the growth of the error with increasing noise. The number of TP: delete: used correspondences TP: used is 8 for both algorithms. We see that the error growth is approximately linear and quite mild, it can be seen that with additive noise of 3 pixels the error is still bearable. For realistic values of noise around 1 pixel, the error is reasonably small.

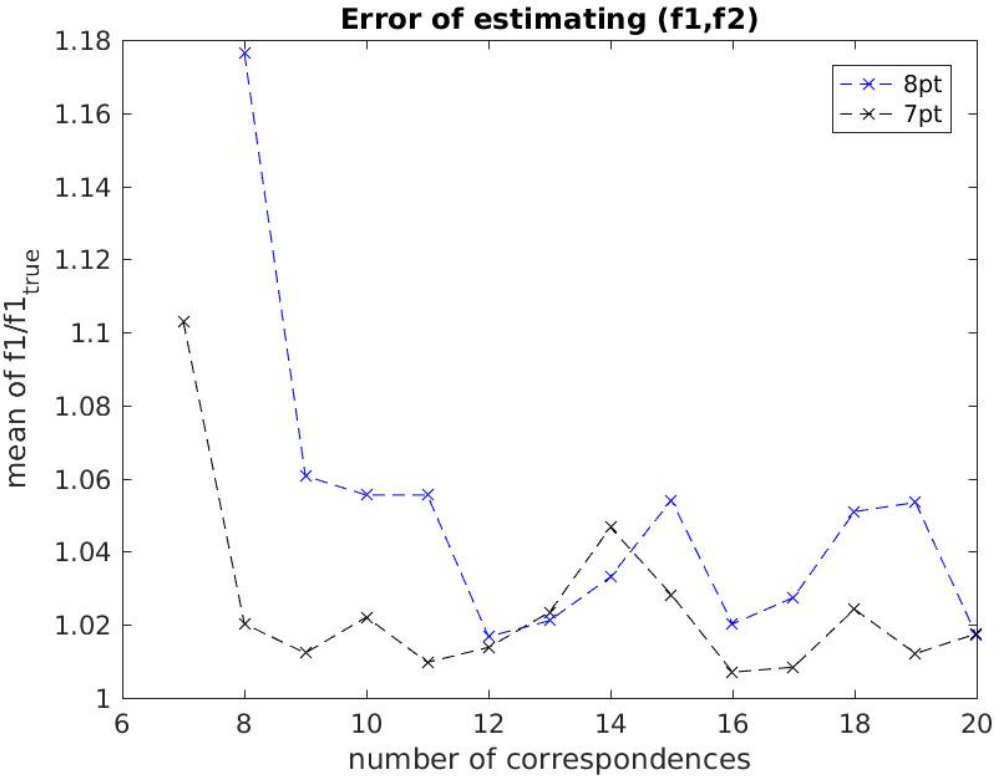


Figure 4.2 TP: A Comparison of TP: the 7pt and 8pt algorithms. Mean error TP: Not clear what that is. Explain better. in focal length and two its quantiles TP: which are shown. The level of noise σ is TP: equal to 1. Imaginary estimates were excluded TP: delete: from experiment. TP: If possible, run for more experiments and plot meadian, 0.25 and 0.75 quantiles.

Another important trend showed on the graph 4.4 is the decrease of number of imaginary estimates with increasing number of used correspondences. We suggest that imaginary focal lengths are due to high level of noise, and appear less often when the noise is averaged over big number of correspondences.

We conclude that when the estimate of the focal lengths is real we can generally expect it to be quite precise and usable for most practical cases. More correspondences insure good performance, and moderate amounts of noise can be handled well. The 7pt algorithm performs better for most applications.

4.2.2 Ratio of focal lengths is robust

We Tdelete: empirically show TP: empirically that the ratio $r = f_2/f_1$ TP: computation is more robust than TP: computation of f_1 or f_2 TP: alone.

To show it we conduct an experiment with a 300 different noise samples applied to the same camera geometry. The figure 4.5 shows scattered estimates of (f_1, f_2) vector. The method used for this experiment is 7pt with 7 correspondences. The noise level σ is 1. A few outliers were filtered out.

The graph TP: add a reference to the figure: Fig. **??** shows scattered estimated points in the space (f_1, f_2). TP: not a valid E sentence: If this is space is not to have additional structure, we expect these estimations to lie on circles around the ground truth, perhaps

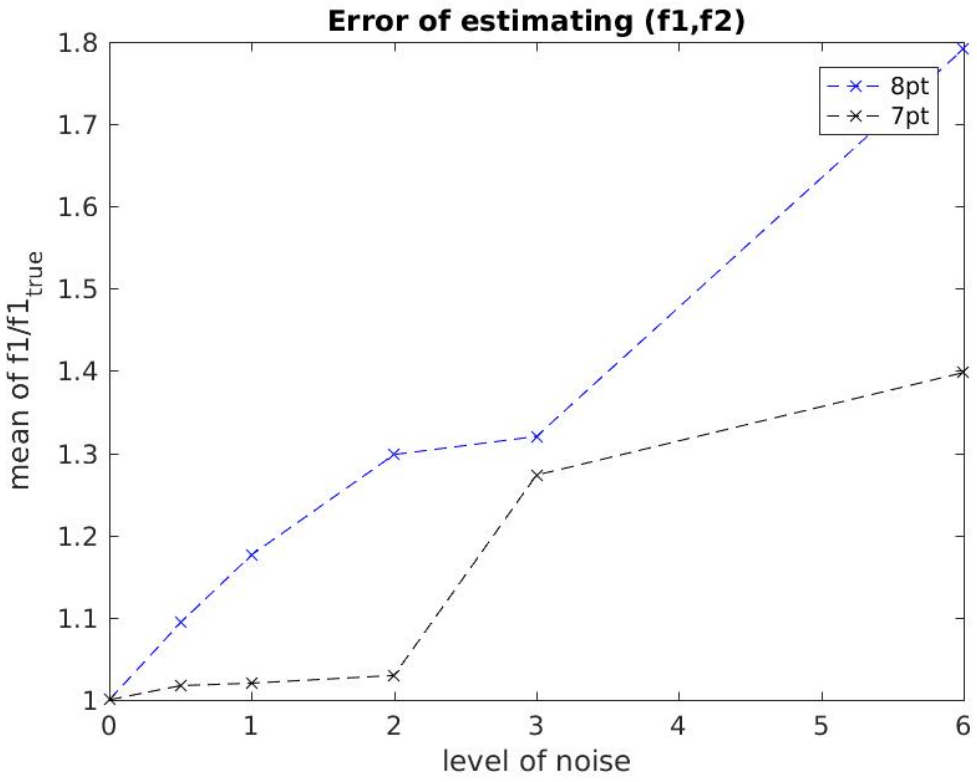


Figure 4.3 Comparison of 7pt and 8pt algorithms against noise level σ. 40 correspondences were used. TP: If possible, run for more experiments and plot meadian, 0.25 and 0.75 quantiles.

similar to a 2 dimensional Gaussian distribution. However, it can be seen that while the deviation of focal lengths is relatively big, the points clearly tend to the line $r = 3/4$ (given in blue), which is indeed the ground truth for the ratio. Imaginary estimates are also shown, and hint to a possibility that there is some structure TP: reformulate: to them also.

We conclude that TP: the computation of the ratio of the focal lengths r TP: by Bougnoux formula is more robust than the TP: the computation of focal lengths themselves TP: by the Bougnoux formula. TP: delete: which fact hopefully can be TP: We will use this fact later to construct a more robust solver.

4.2.3 Imaginary focal lengths

We empirically show that the ratio of the focal lengths is robust even when both focal lengths are imaginary.

TP: ? ... Why must they be both imaginary? Did not we see that we can always get good estimates by square rooting the absolute values of the quares of the focal length estimates?

Our results also show that the probability of getting an imaginary focal length is not connected to distance between optical axes.

Figure 4.6 shows the experiment. The graph is TP: a $\rightarrow$ the histogram of different errors in ratio (measured as $f_2/f_1 - (f_2/f_1)_{gt}$. The method used for this experiment is 7pt with 7 correspondences. The noise level σ is 1. A few outliers lie far off the graph.

Clearly, the likelihood that a focal lengths ratio is estimated better is bigger when

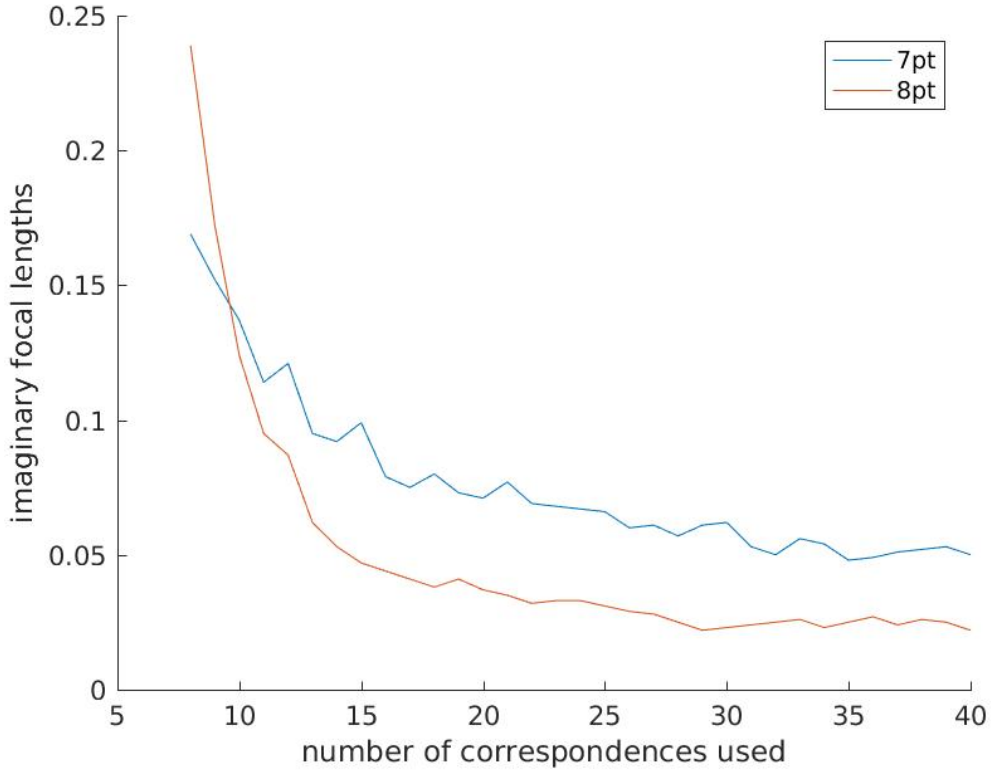


Figure 4.4 Fraction of imaginary focal length estimates produced by 7pt and 8pt algorithms. Noise level σ is 1.

the focal lengths are real. However, the graph shows that the ratios computed from imaginary focal lengths also fall reasonably close to the ground truth, in the sense that TP: the mean of their distribution TP: reformulate: seemingly converges to TP: the ground truth. We suggest that there is a hidden variable that is susceptible to noise and makes both focal lengths imaginary when TP: the level of noise TP: are too → is high. It doesn't, however, affect TP: the ratio of focal lengths, or at least affects it to a lesser degree. TP: It actually looks that it does not affect the absolute values of the squares of the focal lengths. The variable seems to be selecting the sign of the square of the focal lengths such that the higher is the noise the more probable is that the sign will be negative.

Figure 4.7 shows another experiment. One of our guesses was that the probability of having imaginary estimate somehow correlates with closeness to degeneracy. In this experiment we assess the closeness TP: by the length of the shortest transversal between the optical axes TP: delete:transversal length. If the length is zero, TP: the optical axes intersect. TP: delete: , and conversely, the bigger is transversal length, the further they are from intersecting.

We show the estimated ratios $r = f_2/f_1$. TP: We see that even as transversal length increases, there is still approximately the same amount of imaginary estimates. → We see that the number of imaginary estimates does not depend on the length of the shortest transversal. We also see how both ratios computed from real and imaginary estimates tend to the ground truth TP: delete: line. TP: We plot the ratios estimated from imaginary values of the focal lengths with negative sign and in red to distinguish

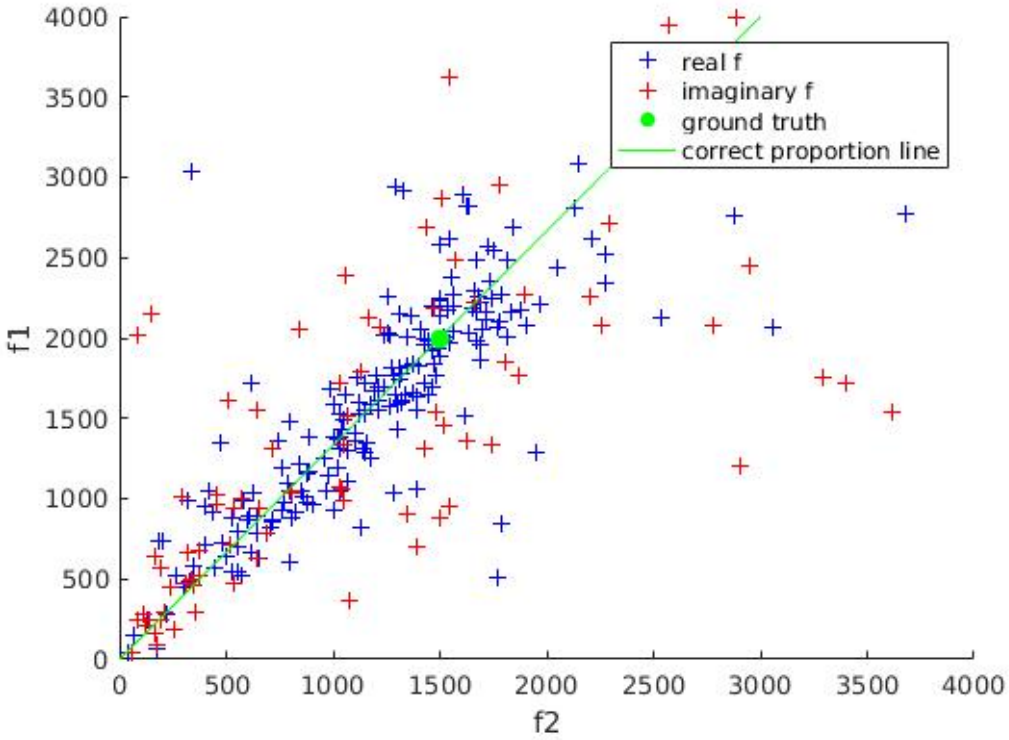


Figure 4.5 Scatter plot of 7pt focal lengths (f_1, f_2) estimates for one scene. Line through ground truth ratio $r = f_2/f_1$ is given for reference. Ground truth point also shown in green. Imaginary estimates shown in absolute value. Noise level σ is 1. 7 correspondences were used.

it from the ratios estimated from the real vallues, which are shown in blue.

4.3 Performance analysis for close to degenerate situations

TP: delete: While in TP: In theory TP: , there does not exist a method to recover focal length when the optical axes intersectTP: . TP: Here we show that with even small amount of noise TP: in what? TP: , the configuration no longer is singular. The focal lengthTP: s TP: then → thus can be computed in almost all practical situations. However, we expect TP: delete: a pose estimation TP: delete: method to deteriorate as the configuration becomes close to singularTP: delete: , due to numerical instability. Here TP: where? we show the extent of this deterioration.

4.3.1 Intersecting optical axes

TP: moved here: In this scenario, two TP: non-parallel optical axes are initially in a plane and we lift one camera away from the plane to distance d. The optical axes become skew lines. TP: delete: as we move one of it away, and the distance between them is equal to d. We compare the quality of estimations for different distances d and levels of noise σ.

In our setup the distance between camera centers is 1 m and the mean distance from TP: delete: a cameraTP: s to TP: delete:a TP: scene pointTP: s is 5 m. The algorithm

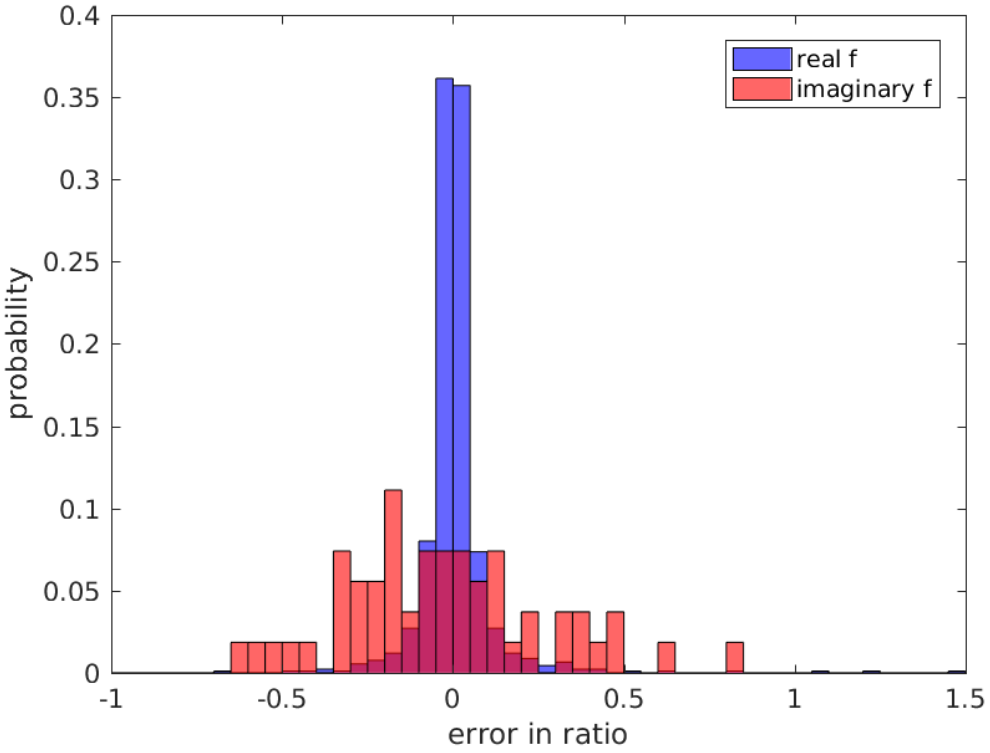


Figure 4.6 Distribution of errors in ratioTP: , i.e. $r - r_{true}$TP: , produced by 7pt TP: algorithm. Separate distributions are shown for the case with both real focal lengths and for the case with at least one imaginary TP: focal legtn. TP: Expain what is in blue and what is in orange first Red color is TP: shown where an orange is supersimposed TP: with → over blue. 40 correspondences were used, and level of noise σ TP: is → was 1. TP: A small number of outliers lieTP: s far off the graph.

used is 7pt and Bougnoux formula.

We empirically assess TP: the robustness of 7pt algorithm and Bougnoux formula in degenerate cases.

Figure 4.8 shows ... TP: Explain what is shown in the figure. What is on the axes and what does tht eplot mean

... that for small values of noise, $\sigma \ll 0.1$ px, TP: the distance between TP: the optical axes significantly affects 7pt TP: method → algorithm performance. The performance deteriorates greatly with decreasing distance TP: between optical axes. For more realistic noise levels, however, figure TP: Fig. **??** shows that the distance haTP: s little impact on TP: the quality of estimation. TP: The performance is almost the same for 1 pixel noise TP: delete: the performance for intersecting axes and axes that TP: have → are 10 cm TP: apart TP: delete: transversal is almost the same. A 50 cm transversal configuration, which we assume is not influenced by the degeneration is marginally more suitable for reconstruction. Even for configuration with axes distance zero, the performance is still acceptable (because of the noise presence TP: im image measurements), with 75% of estimations being off by a factor of at most 2.

We suggest this happens because TP: delete: the presence of noise drives the system further away from TP: a degenerate situation. TP: delet: The noise effect on camera geometry TP: seemingly → aprarently has a bias towards moving optical axes further

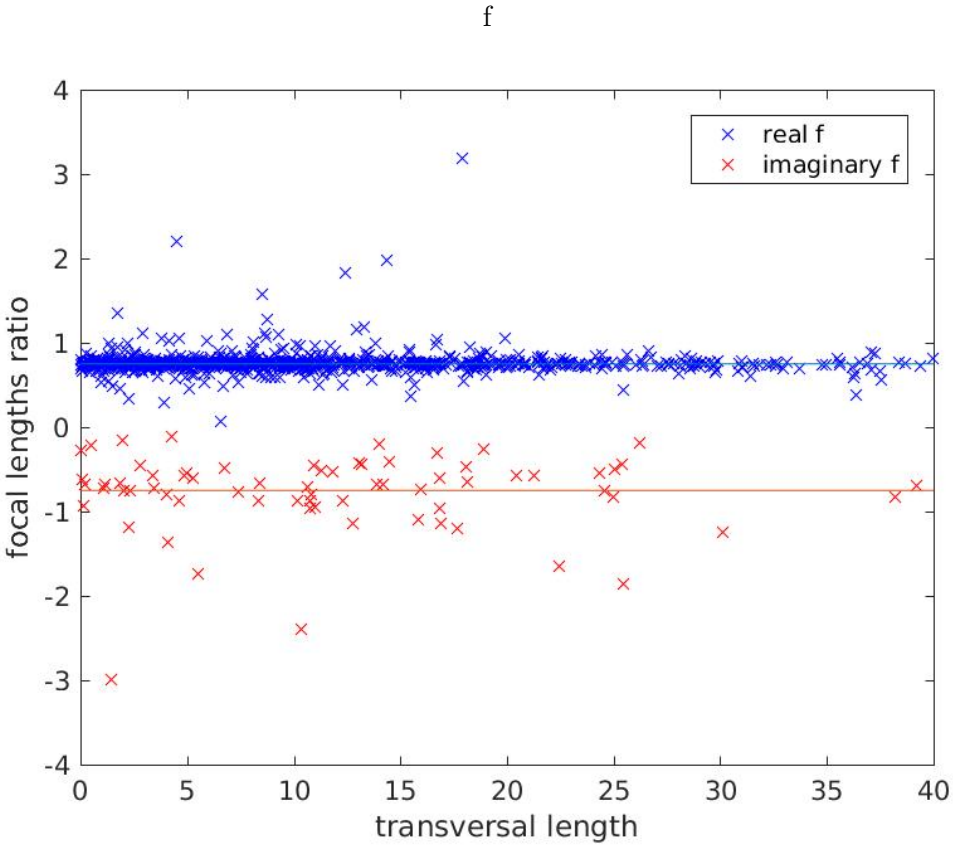


Figure 4.7 Scatter plot of focal lengths ratio $r = f_2/f_1$ estimates produced by 7pt. Ratios corresponding to imaginary focal lengths are plotted as negative to distingish them (they are positive). On X axis is the distance between optical axes. Small number of outliers lie far off the graph. 40 correspondences were used.

from each other, which reduces the effect of numerical instability.

4.3.2 Parallel optical axes

ATP: n TP: particular → interesting case TP: delete: of previous section configuration, is when the optical axes intersect at a point at infinity, i.e. TP: when they are parallel.

In this scenario, two optical axes are initially parallel and we rotate one axis away from the common plane of the axes by TP: delete: an angle α. As we rotate one axis away, it is no longer parallel to the other one. We compare the quality of estimations for different TP: distances d → angles α? and levels of noise σ.

TP: delete: The f TP: Figure 4.9 shows the results. We see that in this caseTP: delete: , TP: the quality of estimation TP: is much worse for trulyfor truly parallel axes and small noise in image corerspondences does not save the situation. TP: delete: At the elevation angle of 0.1° the degeneracy is no longer apparent but for smaller angles performance visibly deteriorates. TP: Interestingly, degeneracy almost disappear when the angle α reaches 0.1°.

TP: We should also analyse the case when one optical axis is perpendicular to the plane spanned by the baseline and the other axis and show that one of the three formulas always works. We should put this perhaps after the algebraic analysis.

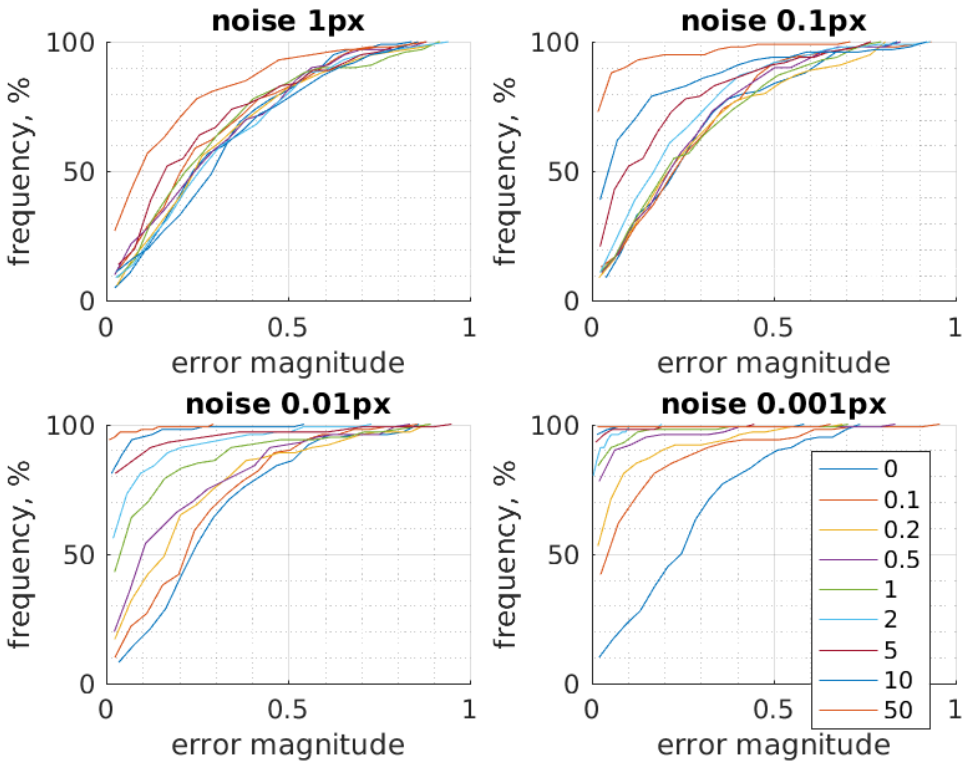


Figure 4.8 Cumulative distribution of estimated focal length multiplicative error for nearly intersecting optical axes. Different lines show different distances between optical axes in centimeters.

4.3.3 Conclusions

We have seen that using more correspondences allowTP: s us to make much better results and get imaginary estimates less often. The ratio $r = f_2/f_1$ seems to be more robust than the focal lengths themselves. The ratio is also usable even when the focal lengths are imaginary.

Of two types of camera configuration degeneracy, intersecting optical axes do not pose a significant risk for camera reconstruction, especially for real-life noise levels. A configuration with parallel or nearly parallel optical axes, however, is a harder case where more than a half of reconstructions may be off by a factor of two and more.

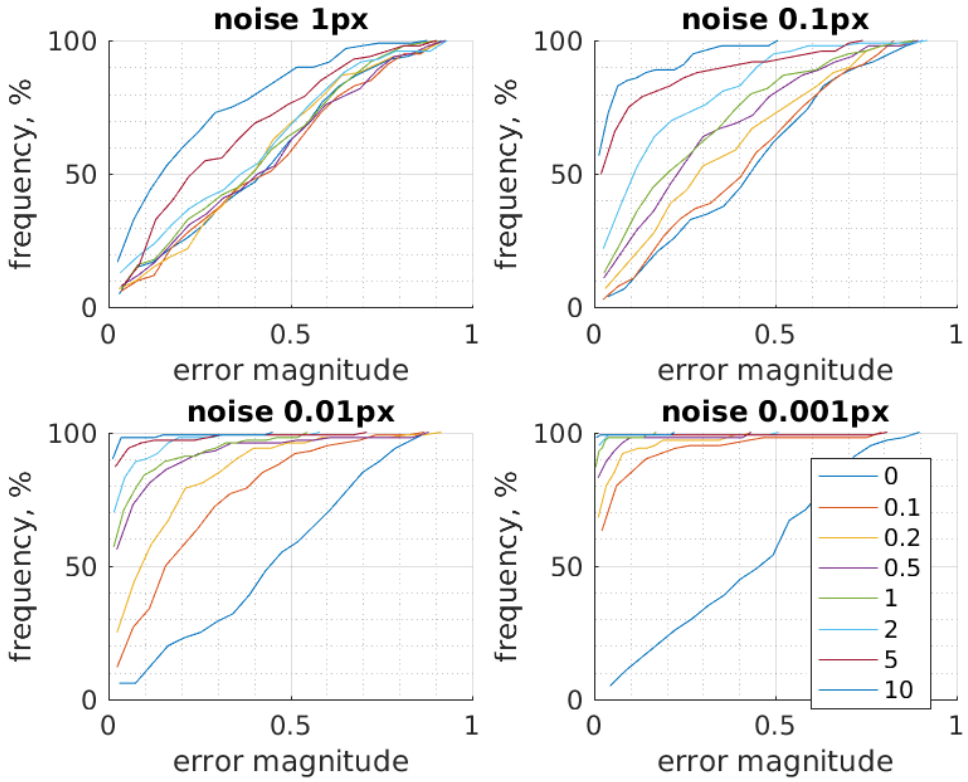


Figure 4.9 Cumulative distribution of estimated focal length multiplicative error for nearly parallel optical axes. Different lines show different angles between optical axes in degrees.

5 TP: Set of fundamental matrices → Algebraic analysis of focal lengths computation

In this section we analyze TP: the 7pt algorithm → focal lenghts computations from the algebraic geometry point of view. We show that indeed three and only three fundamental matrices (some possibly complex or of rank 1) can be derived using algebraic geometry. We analyse degeneracies and show that the only major degeneracy occurs when $\mathtt{F}_{3,3} = 0$. In this case the focal lengths are not determined, i.e. they may be anything from the interval $< -\infty, \infty >$. We also show that three different formulae exist for computing focal lengths from a Fundamental matrix (one being the Bougnoux formula). TP: explain how it solves the probelm of the degeneracy described by HZ ... this is an important contribution of our work

5.1 Analysis

We analyse the set of valid fundamental matrices TP: of up to focal lengths calibrated camera pairs over complex numbers using techniques of algebraic geometry. To this end, we use Macaulay2 [7], a programming language and also a software pack for algebraic geometry and abstract commutative algebra. Macaulay2 tends to have an intuitive interface for many algebraic operations and, in general, code can be read as easily as mathematical equations. For this reason, instead of mathematical TP: operations → language, we will in use this section Macaulay2 code snippets TP: directly. TP: This also has the advantages of our results being easy reproducible. → This helps to make our results reproducible, too.

In algebraic terms, we choose to describe the set of valid Fs as a variety in a complex numbers' field $\mathbb{C}^{11}$ of 11 variables, 9 for elements of the fundamental matrix F and 2 for focal lengths f_1, f_2. We assume that the principal points $\mathbf{p}_1$, $\mathbf{p}_2$ are zero.

The ideal Gs is the ideal generated by rank 2.5 and trace constraints 2.2 and saturated by the ideal $\langle f_1\, f_2 \rangle$. The saturation is desirable, because the case when $f_1 = 0$ or $f_2 = 0$ doesn't correspond to any real camera system.

TP: This actually does not have to remove all cases when $f_1 = f_2 = 0$ but usually removes all big spurious components. For instance, consider $\langle x\, z, y\, z \rangle$ saturated by $\langle x, y \rangle$, which gets us $\langle x\, z, y\, z, z$. We see that $x = y = z = 0$ is still possible but the z-axis has been removed.

Below is a snippet of Macaulay2 code explaining TP: the construction of Gs.

```
R = QQ[f1,f2,f11,f12,f13,f21,f22,f23,f31,f32,f33, MonomialOrder=>Lex]
F = matrix{{f11,f12,f13},{f21,f22,f23},{f31,f32,f33}}
K1 = matrix{{f1, 0, 0}, {0, f1, 0}, {0, 0, 1}}
K2 = matrix{{f2, 0, 0}, {0, f2, 0}, {0, 0, 1}}
E = transpose(K2)*F*K1 -- Essential matrix
G = ideal(det(E)) + minors(1, 2*E*transpose(E)*E - trace(E*transpose(E))*E);
dim G, codim G, degree G
Gs = saturate(G,ideal(f1*f2)); -- det(K1),det(K2) are non-zero
dim Gs, codim Gs, degree Gs
```

The ideal Gs contains conditions under which a matrix is generically a fundamental matrix. The variety $\mathbf{V}(Gs)$ therefore is a variety that contains all valid fundamental matrices. It also contains more matrices, specifically those of rank 0 and 1. Generically, however, we expect a matrix from the variety to be of rank 2. Moreover, the variety does contain some complex matrices. In practical situations however, when considering a solver constructed this way, we can just sort out the spurious solutions afterwards as there will be only three solutions in total.

The 7pt algorithm can be regarded as intersecting the variety $\mathbf{V}(Gs)$ it with 7 hyperplanes, and it gives us three different solutions. Therefore we would expect the ideal Gs to have dimension 7 TP: NO. We would expect this to be 8 because the solutions are one-dimensional subspaces since all is homogeneous and we are in a projective space. and degree 3, but in reality, as constructed, it has dimension of 8 and degree of 58. This is explained by the fact that the variety $\mathbf{V}(Gs)$ has an additional component of dimension 8 and degree 58. Executing 7pt algorithm can be viewed as adding epipolar constraints to the ideal, and later we will see that this component doesn't survive the procedure.

The next snippet shows how to compute algebraic conditions in terms of elements of Fundamental matrix only, by eliminating f_1, f_2.

```
M = eliminate(Gs,{f1,f2});
dim M, codim M, degree M
-- the commands "mingens gb" give a minimal set of generators
-- for the groebner basis of an ideal
m = mingens gb M
```

After executing TP: the above, we see that the ideal M has only one generator - the det(F) polynomial. This means that the TP: algebraic constraints TP: on the set of fundamental matrices of up to focal lengths calibrated camera pairs are the same as for the fully uncalibrated camera case [?]. It can be deduced that a seven-tuple of corresponding points obtained by completely uncalibrated cameras can also be explained by cameras with two unknown focal lengths when the focal lengths are allowed to attain non-real values.

It also means that only the singularity constraintTP: ?? is needed to solve the system, and the Demazure constraints are extraneous.

5.2 Computing focal lengths

In the next snippet we also show how the Bougnoux formula [3] can be derived with algebraic geometry, given as a polynomial in the entries of F. This is done by eliminating one of the focal lengths from the ideal Gs. TP: It turns out that the Gröbner basis of the eliminated ideal s_i contains TP: the determinant of F and additional TP: three polynomials, and from each a formula for computing the focal length that wasn't eliminated can be deducted. This means that there exist three algebraically independent constraints on each focal length.

```
s2 = mingens gb eliminate(Gs,f1)
s1 = mingens gb eliminate(Gs,f2)
-- Formulae for f1
(m11,c11) = coefficients(s1_1_0,Variables=>{f1}) -- extract coefficients
(m12,c12) = coefficients(s1_2_0,Variables=>{f1}) -- extract coefficients
```

```
(m13,c13) = coefficients(s1_3_0,Variables=>{f1}) -- extract coefficients
-- Formulae for f2
(m21,c21) = coefficients(s2_1_0,Variables=>{f2}) -- extract coefficients
(m22,c22) = coefficients(s2_2_0,Variables=>{f2}) -- extract coefficients
(m23,c23) = coefficients(s2_3_0,Variables=>{f2}) -- extract coefficients
```

We see that there exist three formulae for each focal length, for example,

$$f_2^2 = -\frac{c23_{0,1}}{c23_{0,0}} = \\ -\frac{f_{3,3}(f_{1,1}f_{2,3}f_{3,1} + f_{1,2}f_{2,3}f_{3,2} - f_{1,3}f_{2,1}f_{3,1} - f_{1,3}f_{2,2}f_{3,2})}{(f_{1,1}^2 f_{1,3}f_{2,3} - f_{1,1}f_{1,3}^2 f_{2,1} + f_{1,1}f_{2,1}f_{2,3}^2 + f_{1,2}^2 f_{1,3}f_{2,3} - f_{1,2}f_{1,3}^2 f_{2,2} + f_{1,2}f_{2,2}f_{2,3}^2 - f_{1,3}f_{2,1}^2 f_{2,3} - f_{1,3}f_{2,2}^2 f_{2,3})}. \tag{5.1}$$

It can be checked that this formula is equivalent to the Bougnoux formula by expressing the latter directly in terms of Fundamental matrix elements. The two TP: other formulae are:

$$f_2^2 = -\frac{c21_{0,1}}{c21_{0,0}} = \\ = -\frac{f_{3,3}(f_{2,1}f_{3,1}f_{3,3} + f_{2,2}f_{3,2}f_{3,3} - f_{2,3}f_{3,1}^2 - f_{2,3}f_{3,2}^2)}{(f_{1,1}f_{1,3}f_{2,1}f_{3,3} - f_{1,1}f_{1,3}f_{2,3}f_{3,1} + f_{1,2}f_{1,3}f_{2,2}f_{3,3} - f_{1,2}f_{1,3}f_{2,3}f_{3,2} + f_{2,1}^2 f_{2,3}f_{3,3} - f_{2,1}f_{2,3}^2 f_{3,1} + f_{2,2}^2 f_{2,3}f_{3,3} - f_{2,2}f_{2,3}^2 f_{3,2})}. \tag{5.2}$$

$$f_2^2 = -\frac{c22_{0,1}}{c22_{0,0}} = \\ = -\frac{f_{3,3}(f_{1,1}f_{3,1}f_{3,3} + f_{1,2}f_{3,2}f_{3,3} - f_{1,3}f_{3,1}^2 - f_{1,3}f_{3,2}^2)}{(f_{1,1}^2 f_{1,3}f_{3,3} - f_{1,1}f_{1,3}^2 f_{3,1} + f_{1,1}f_{2,1}f_{2,3}f_{3,3} + f_{1,2}^2 f_{1,3}f_{3,3} - f_{1,2}f_{1,3}^2 f_{3,2} + f_{1,2}f_{2,2}f_{2,3}f_{3,3} - f_{1,3}f_{2,1}f_{2,3}f_{3,1} - f_{1,3}f_{2,2}f_{2,3}f_{3,2})}. \tag{5.3}$$

The formulae for focal length TP: f_1 may be obtained by transposing the fundamental matrix. The undertaken analysis of the formulae has shown that generically they differ little in terms TP: of stability and quality of estimations.

It can be shown by substitution that the Bougnoux formula vanishes when either first or second column of `F` is TP: the zero vector. Of the formulae we found, however, the first formula 5.1 does not vanish when the first column of `F` is zero, and the second one **??** does not vanish when the second column is zero. When the third column is zero, all three formulae vanish, as $\mathtt{F}_{3,3} = 0$.

TP: NOT true! One of them becomes $p(F)\ f_2^2 = 0$, implying $f_2^2 = 0$ for $p(F) \neq 0$.

We show that the degeneration when first or second column is zero occurs when the plane defined by the baseline and the optical axis of one camera is perpendicular to the plane defined by the baseline and optical axis of the other camera.

TP: TODO ... write down the example from the photo I sent you. Refere to HZ-2003 book and explain why this is an improvement of the state of the art.

Note that there TP: still remains an additional degeneracy when each of the three denominators vanishes. When this happens, the focal lengths cannot be reconstructed from the matrix `F`, as there are no constraints on them (they can have any value).

TP: Done until here

5.3 The ratio formula

We also present a formula to compute directly the ratio $r = f_2/f_1$.

TP: Perhaps the following could help to format log formulas. splitdfrac can be nested to split the numerators and denominators to more than two pieces. See tex.stackexchange.com/questions/18091 the-equation-inside-a-fraction.

$$r = \frac{f_2}{f_1} = \frac{\begin{array}{r} f_{1,1}^2 f_{3,1}^2 f_{3,3} + 2f_{1,1}f_{1,2}f_{3,1}f_{3,2}f_{3,3} - f_{1,1}f_{1,3}f_{3,1}^3 - f_{1,1}f_{1,3}f_{3,1}f_{3,2}^2 + f_{1,2}^2 f_{3,2}^2 f_{3,3} - f_{1,2}f_{1,3}f_{3,1}^2 f_{3,2} - f_{1,2}f_{1,3}f_{3,2}^3 + f_{2,1}^2 f_{3,1}^2 f_{3,3} \\ + 2f_{2,1}f_{2,2}f_{3,1}f_{3,2}f_{3,3} - f_{2,1}f_{2,3}f_{3,1}^3 - f_{2,1}f_{2,3}f_{3,1}f_{3,2}^2 + f_{2,2}^2 f_{3,2}^2 f_{3,3} - f_{2,2}f_{2,3}f_{3,1}^2 f_{3,2} - f_{2,2}f_{2,3}f_{3,2}^3 \end{array}}{\begin{array}{r} f_{1,1}^2 f_{1,3}^2 f_{3,3} - f_{1,1}f_{1,3}^3 f_{3,1} + 2f_{1,1}f_{1,3}f_{2,1}f_{2,3}f_{3,3} - f_{1,1}f_{1,3}f_{2,3}^2 f_{3,1} + f_{1,2}^2 f_{1,3}^2 f_{3,3} - f_{1,2}f_{1,3}^3 f_{3,2} + 2f_{1,2}f_{1,3}f_{2,2}f_{2,3}f_{3,3} \\ - f_{1,2}f_{1,3}f_{2,3}^2 f_{3,2} - f_{1,3}^2 f_{2,1}f_{2,3}f_{3,1} - f_{1,3}^2 f_{2,2}f_{2,3}f_{3,2} + f_{2,1}^2 f_{2,3}^2 f_{3,3} - f_{2,1}f_{2,3}^3 f_{3,1} + f_{2,2}^2 f_{2,3}^2 f_{3,3} - f_{2,2}f_{2,3}^3 f_{3,2} \end{array}} \tag{5.4}$$

Computing the ratio directly from `F` can offer TP: more → greater speed and TP: more stability than computing it from the focal lengths. Indeed, we observed that for close to degenerate situations, the formula, TP: if → although marginally TP: Please add a graph showing this., is a better estimator.

5.4 Computing Fundamental matriTP: ces

We show in detail how the 7pt solver workTP: s from the viewpoint of algebraic geometry.

We first simulate a random set of correspondences. Skipping this code for brevity, assumeTP: ing that the matrix `B` is a matrix from 7pt algorithm (1).TP: It is not clear what is matrix `B`! Please explain it.

In the next snippet we show detailed analysis of the computation results. The variety $\mathbf{V}(GBs)$ is the variety of all possible focal lengths that explain the generated correspondences.

```
m = transpose matrix {{f11,f12,f13,f21,f22,f23,f31,f32,f33}}
rank B
eB = B*m
IB = minors(1,eB)
GBs = Gs + IB
gGBS=mingens gb GBs
dim GBs, codim GBs, degree GBs
pGBS = minimalPrimes GBs
```

Ideal IB contains the epipolar constraints. By adding the ideals Gs and IB we combine the constraints and obtain an ideal that corresponds to a variaty of solutions for our correspondences.

Variety $\mathbf{V}(GBs)$ can consist of either 6 or 14 components, one of dimension 2, and the rest of dimension 0. We will show the significance of this fact and what are these components corresponding to.

TP: Can we write down the equations?

The ideal corresponding to the component of dimension 2 contains a polynomial $f_{3,3}$, i.e., it represents the degenerate situation when $f_{3,3}$ and the cameras' optical axes intersect (or are parallel) by lemma 2.2.1. The ideal also does not contain any polynomial in f_1 or f_2, so neither of focal lengths can be determined as there are no constraints on them. Because of this, the component has dimension 2 - its two degrees of freedom are f_1 and f_2.

Components of dimension 0 are point varieties (all TP: variables → unknowns are determined there). The first TP: what? contains the polynomials f_1 and f_2, which means TP: that it corresponds to a degenerate case when $f_1 = 0, f_2 = 0$. Technically,

such a fundamental matrix satisfies all constraints and is an example of TP: a matrix having rank less than two (see the discussion in section 2.2.4). This case could be avoided by using TP: variables → unknowns $w_1 = \frac{1}{f_1}, w_2 = \frac{1}{f_2}$. TP: I tried it and there is a case f(F,w1,w2)=0. Which still needs to be analyzed.

TP: The 12 (or 4) TP: When 12 and when 4? other components can be divided into 3 groups (or 1 group), which correspondTP: (s) to 3 different Fundamental matrices (a single matrix). Each group includes 4 point varieties in form $f_1 = \pm C_1, f_2 \pm C_2$ where C_1, C_2 are constants. The reason for this is that the focal length actually only enters the Bougnoux formula, as well as formulae we derived **??**, in TP: the second degree, so it is impossible to determine it's sign[1].

We see that each group correspondTP: s to one of the three fundamental matrices as returned by 7pt algorithm. Sometimes, however, two of them would be complex. In this case we will get only one group, as our code considers varieties over real numbers.

5.5 Conclusions

We have shown how Bougnoux formula and 7pt algorithm work in terms of algebraic geometry. We have derived 2 new formulae for computing focal lengths from fundamental matrix, and a formula for computing ratio $r = f_2/f_1$. No more algebraically independent formulae can be derived for this problem. We have shown that a known degeneracy TP: cite HZ-2003 here can be avoided by using the right focal length formula of the three. The degeneracy reduces to the case when all three formulae fail.

We confirmed that three and only three fundamental matrices can explain 7 correspondences. We have shown that the trace constraint is redundant while solving 7pt problem TP: What is 7pt problem? It was never mentioned before..

[1] Of course for any practical application, the situation is unambiguous, as the positive sign should always be chosen.

6 Improvements

6.1 f-Ratio

We present a new algorithm, called f-Ratio, for robust focal lengths computTP: ation. The algorithm uses information that the ratio of focal lengths is more robust to achieve superior accuracy. We analyze the performance of the algorithm.

6.1.1 Algorithm

The idea of the algorithm is to use a new solver that would compute F from 6 correspondences given the ratio $r = f_2/f_1$. We create such a solver using TP: the automatic generator [15]. The solver uses Demazure [19] polynomials in terms of elements of the matrix F, f_1, and r.

Otherwise TP: It is not clear why and when is the "otherwise" situation. Please explain. a solver assuming fixed focal length $f_1 = f_2$ could be used, for example that of Torii et el. [24], by first rescaling

$$v_r = \begin{bmatrix} r & 0 & 0 \\ 0 & r & 0 \\ 0 & 0 & 1 \end{bmatrix} v.$$

The output of such solver would be f_1.

In both cases, the solver yieldTP: s 15 (possibly non-real) solutions.

Data: A list of 7 right image points u and a list of the corresponding left image points v, lists of remaining tentative correspondences u_{test}, v_{test}, 6pt solver $solve6pt$
Result: Fundamental matrix F

```
begin
    Estimate F_0 from all correspondences;
    Estimate r = f_2/f_1 from F_0 ;
    foreach 6-tuple u_6 of points drawn from u do
        v_6 ← corresponding points to u_6 from v ;
        Fs_i ← solve6pt(u_6, v_6, r);
        F_i ← the matrix ∈ Fs_i which best explains u_test, v_test.
    end
    return F_i which best explains u_test, v_test.
end
```

Algorithm 4: f-Ratio

6.1.2 Performance

We assess the performance of the algorithm in a similar manner TP: to → as in the previous analysis. In this experiment, we compare our method f-Ratio to baseline 7pt TP: method → algorithm.

We also plot the best reconstruction which was found by the procedure 4. In most cases, the matrix chosen by the consensus set is not the best one available, which is reflected by this graph. We measure the "fitness" of the matrix in terms of focal lengths only in this case. TP: The previous sentence does not seem make much sense.

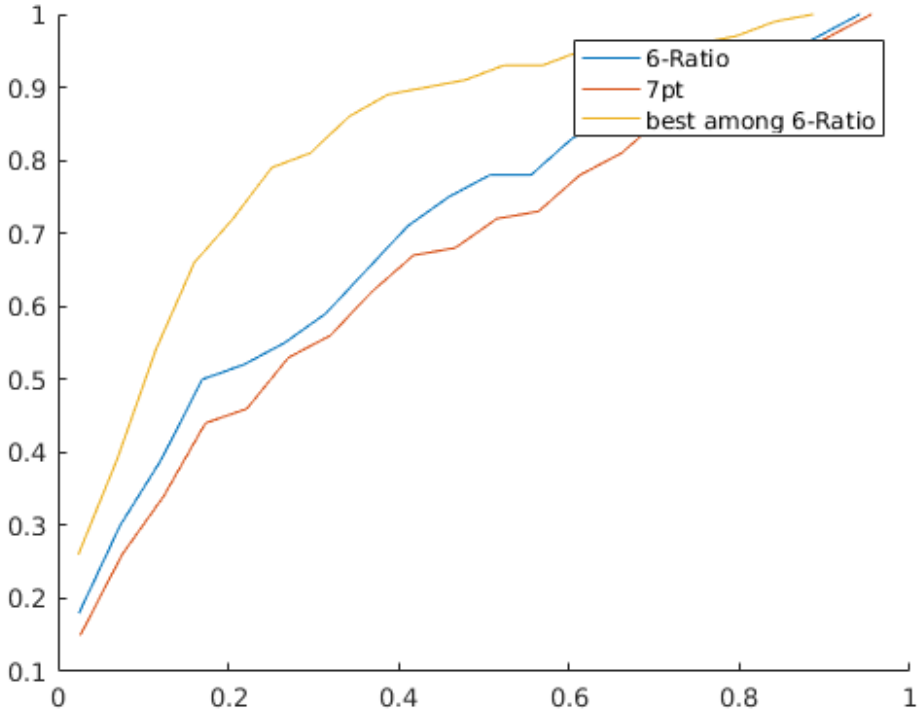


Figure 6.1 Cumulative distribution of estimated focal length multiplicative error showing performance of f-Ratio

In the next experiment, we add an additional amount of noise ($\sigma = 10$ pixels) to one of the points. This demonstrates that our method can cope with outliers even better than the standart RANSAC algorithm.

The results show that it is possible to reconstruct scenes TP: better with f-Ratio than with 7pt algorithm. However, since we need to test 7×15 matrices[1], the amount of computation needed is much bigger. The algorithm TP: add the reference **??** demonstrateTP: s that TP: valuing → using estimated ratio of focal lengthTP: s over TP: the focal lengths themselves may improve TP: the accuracy TP: of the algorithm.

6.2 Prior focal length

Hartley [9] describes an iterative algorithm for computing focal lengths from point correspondences, which incorporates prior information about focal lengths and principal points. The algorithm uses a Levenberg-Marquardt optimization TP: delete: to achieve this. In this section we describe an improvement to TP: the → his algorithm. Using

[1] At most. The 6pt solver usually gives from 9 to 14 real solution, but the maximal possible number is 15.

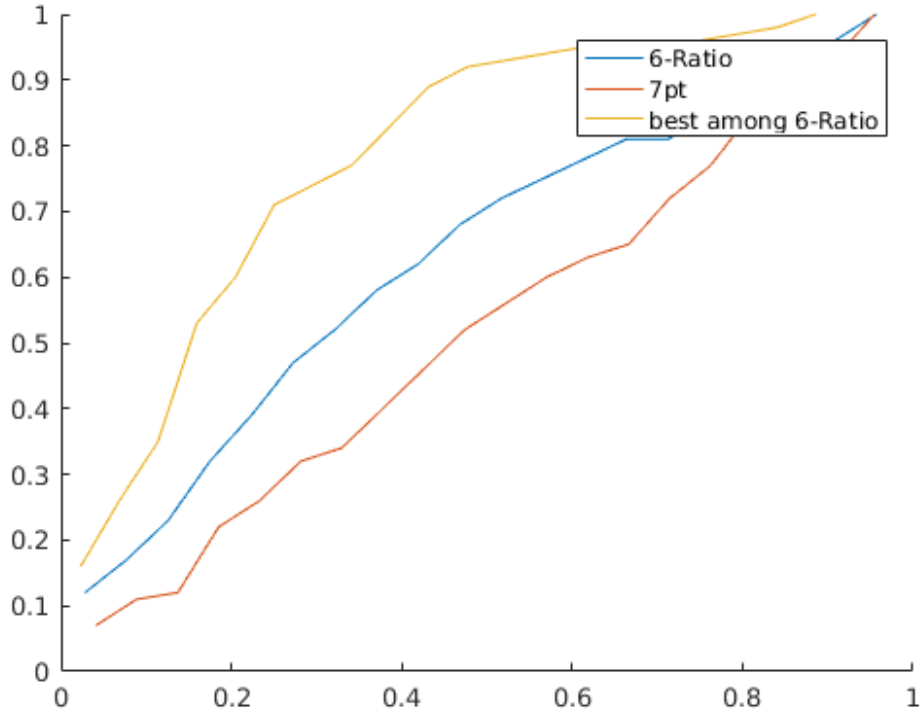


Figure 6.2 Cumulative distribution of estimated focal length multiplicative error showing performance of f-Ratio with one outlier correspondence. TP: Wha is "best among 6-Ratio"? PLease explain it here as well as in the text.

computed focal length ratio $r = f_2/f_1$, we are able to get better focal lengths estimatTP: es.

6.2.1 Original TP: Hartley's algorithm

Original HartleyTP: 's algorithm optimizes a certain cost function given the weights w_i, point correspondences $\mathtt{x}_1$, $\mathtt{x}_2$, prior focal lenth $\bar{f}_1$, $\bar{f}_2$, minimal focal length f_{min} and prior principal points $\bar{\mathbf{p}}_1$, $\bar{\mathbf{p}}_2$. We give this cost function as algorithm 5.

Data: Fundamental matrix $\mathtt{F}$, principal points $\mathbf{p}_1$, $\mathbf{p}_2$.
Result: Vector of costs (errors) $\mathbf{C}$
begin
  From $\mathtt{F}$, $\mathbf{p}_1$, $\mathbf{p}_2$ compute f_1, f_2;
  $C_{\mathtt{F}} \leftarrow$ Sampson error of the matrix $\mathtt{F}$ on the points $\mathtt{x}_1$, $\mathtt{x}_2$;
  $C_f \leftarrow$
  $w_1^2(f_1^2 - \bar{f}_1^2)^2 + w_2^2(f_2^2 - \bar{f}_2^2)^2 + w_d^2(f_1^2 - f_2^2)^2 + w_{z1}^2(f_{min}^2 - f_1^2)^2 + w_{z2}^2(f_{min}^2 - f_2^2)^2$;
  $C_{\mathbf{p}} \leftarrow w_p^2\|\mathbf{p}_1 - \bar{\mathbf{p}}_1\|^2 + w_p^2\|\mathbf{p}_2 - \bar{\mathbf{p}}_2\|$;
  return *Costs* $C_{\mathtt{F}}$, $C_{\mathbf{p}}$, C_f;
end

Algorithm 5: Cost function of [9]

Note that the focal lengths aren't parameters of this cost function, as they are determined by the fundamental matrix and principal points.

The cost function of focal lengths C_f incorporates a few interesting ideas besides using prior knowledge in first two terms. Its third term drives the focal lengths to the same value, which probably reflects the fact that most real cameras have similar focal lengths from a relatively small (compared to infinity) range. Also using such term, TP: the method should perform accurately even if the cameras used were indeed the same camera. TP: the fourth and TP: the fifth terms serve two purposes. Firstly, they prevent the squared focal lengths f_1^2, f_2^2 from becoming negative, which addresses the problem of imaginary focal lengths estimates. Secondly, these terms also TP: serve as a guard → prevent focal lengths TP: from converging to zero.

Hartley suggests to initialize the optimization with a technique he calls "calibrated reconstruction". In our experiments, we use the matrix computed with 7pt algorithm instead.

6.2.2 Using ratio

We introduce a modification based on use of the estimated ratio of focal lengths $r = f_2/f_1$. Given the correspondences, we use 7pt algorithm 1 and the Bougnoux formula to estimate the ratio. Then we include it in the cost function on f from the function 5:

$$C_f = w_1^2(f_1^2 - \bar{f}_1^2)^2 + w_2^2(f_2^2 - \bar{f}_2^2)^2 + w_d^2((rf_1)^2 - f_2^2)^2 + w_{z1}^2(f_{min}^2 - f_1^2)^2 + w_{z2}^2(f_{min}^2 - f_2^2)^2$$

We show that such enhanced solver produces better focal length estimates. The graph 6.3 showTP: s the experiment with 100 runs of each solver.

6.3 Conclusions

We conclude that the fact TP: that focal length ratio $r = f_2/f_1$ is robust can be used to construct more efficient algorithms. We show two exampleTP: s of modifications to existing algorithms where we explicitly use a ratio estimation to improve TP: the accuracy TP: of the focal lengths computation .

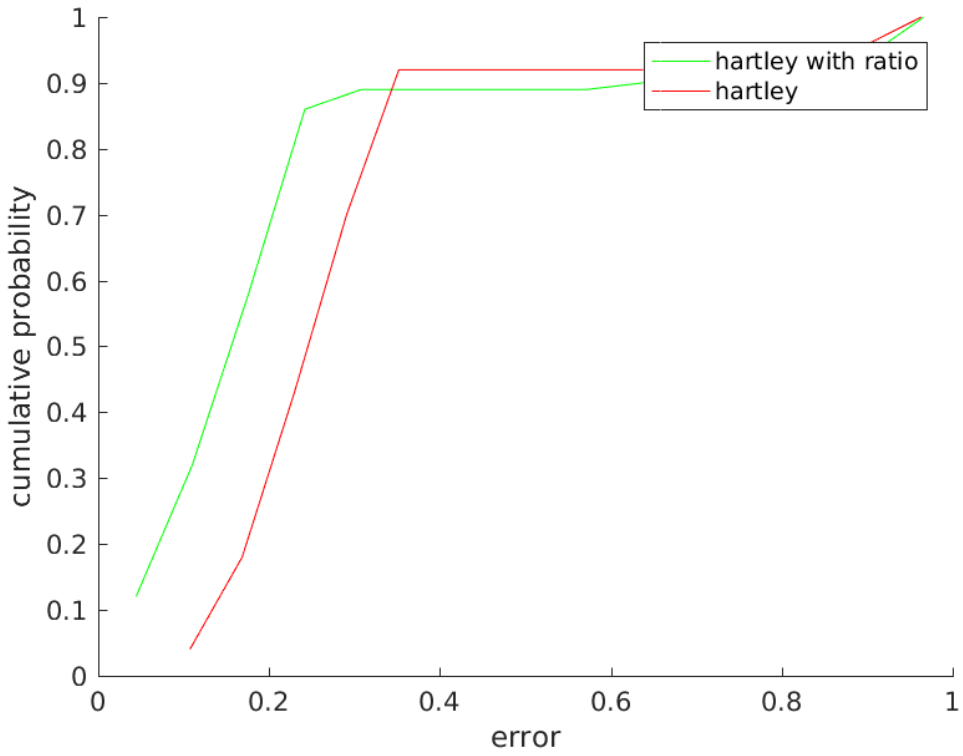


Figure 6.3 Cumulative distribution of estimated focal length multiplicative error of Hartley solver [9] with or without using ratio $r = f_2/f_1$ in the cost function. The ground truth focal lengths were $(3, 4)$. The prior focal lengths were drawn from an uniform distribution between $(3, 4)$ and $(4.5, 5.5)$. The ground truth principal points were zero, and the prior principal points were $(0.1, 0.1)$. The number of correspondences used was 40, and the level of noise σ was equal to 1. The weights on principal point priors were 10 times smaller then weights on focal lengths priors.

Bibliography

[1] Sameer Agarwal, Hon-leung Lee, Bernd Sturmfels, and Rekha R. Thomas. On the existence of epipolar matrices. *CoRR*, abs/1510.01401, 2015.

[2] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. Speeded-up robust features (surf). *Comput. Vis. Image Underst.*, 110(3):346–359, June 2008.

[3] Sylvain Bougnoux. From projective to euclidean space under any practical situation, a criticism of self-calibration. In *ICCV*, pages 790–798, 1998.

[4] David A. Cox, John Little, and Donal O'Shea. *Using Algebraic Geometry*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005.

[5] David A. Cox, John Little, and Donal O'Shea. *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra, 3/e (Undergraduate Texts in Mathematics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2007.

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

[7] Daniel R. Grayson and Michael E. Stillman. Macaulay2, a software system for research in algebraic geometry. Available at http://www.math.uiuc.edu/Macaulay2/.

[8] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004.

[9] Richard Hartley and Chanop Silpa-anan. Reconstruction from two views using approximate calibration. In *Proc. 5th Asian Conf. Comput. Vision*, pages 338–343, 2002.

[10] Richard I. Hartley. In defense of the eight-point algorithm. *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(6):580–593, June 1997.

[11] Richard I. Hartley. Chirality. *International Journal of Computer Vision*, 26(1):41–61, 1998.

[12] Kenichi Kanatani and Chikara Matsunaga. Closed-form expression for focal lengths from the fundamental matrix. In *in Proc. 4th Asian Conf. Computer Vision*, pages 128–133, 2000.

[13] Kenichi Kanatani, Atsutada Nakatsuji, and Yasuyuki Sugaya. Stabilizing the focal length computation for 3-d reconstruction from two uncalibrated views. *International Journal of Computer Vision*, 66(2):109–122, 2006.

[14] Yasushi Kanazawa, Yasuyuki Sugaya, and Kenichi Kanatani. Decomposing three fundamental matrices for initializing 3-d reconstruction from three views. *IPSJ Transactions on Computer Vision and Applications*, 6:120–131, 2014.

[15] Zuzana Kukelova, Martin Bujnak, and Tomas Pajdla. Automatic generator of minimal problem solvers. In *Computer Vision - ECCV 2008, 10th European Conference on Computer Vision, Proceedings, Part III*, pages 302–315, 2008.

[16] Zuzana Kukelova, Martin Bujnak, and Tomas Pajdla. Polynomial eigenvalue solutions to minimal problems in computer vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34:1381–1393, 2011.

[17] H. C. Longuet-Higgins. Readings in computer vision: Issues, problems, principles, and paradigms. chapter A Computer Algorithm for Reconstructing a Scene from Two Projections, pages 61–62. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1987.

[18] David G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision*, 60(2):91–110, November 2004.

[19] Demazure M. Sur deux problemes de reconstruction. Technical Report 882, INRIA, 1988.

[20] Pierre Moulon, Pascal Monasse, Renaud Marlet, and Others. Openmvg. an open multiple view geometry library. https://github.com/openMVG/openMVG.

[21] David Nistér. An efficient solution to the five-point relative pose problem. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(6):756–777, June 2004.

[22] Johannes Lutz Schönberger and Jan-Michael Frahm. Structure-from-motion revisited. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[23] Noah Snavely, Steven M. Seitz, and Richard Szeliski. Photo tourism: Exploring photo collections in 3d. In *ACM SIGGRAPH 2006 Papers*, SIGGRAPH '06, pages 835–846, New York, NY, USA, 2006. ACM.

[24] Akihiko Torii, Zuzana Kukelova, Martin Bujnak, and Tomás Pajdla. In *ACCV Workshops (2)*, 2010.

[25] S. Workman, C. Greenwell, M. Zhai, R. Baltenberger, and N. Jacobs. Deepfocal: A method for direct focal length estimation. In *2015 IEEE International Conference on Image Processing (ICIP)*, pages 1369–1373, Sept 2015.