Robust Focal Length Computation

Oleh Rybkin

rybkiole@fel.cvut.cz

May 26, 2017

Robust Focal Length Computation

Oleh Rybkin

May 26, 2017

BACHELOR PROJECT ASSIGNMENT

Student: Oleh R y b k i n

Study programme: Open Informatics

Specialisation: Computer and Information Science

Title of Bachelor Project: Robust Focal Length Computation

Guidelines:

1. Review the state of the art in camera focal length computation from Fundamental matrices in [1-8] and in references therein.
2. Experiment with the standard approach to focal length computation [1,2] and describe its limitations.
3. Suggest an improvement of the focal length calibration and implement it in Matlab.
4. Provide experimental evaluation of the new method and comparison to the state of the art focal length computation.

Bibliography/Sources:

[1] Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision, Cambridge University Press, 2nd ed.,2003.
[2] Bougnoux, S.: From Projective to Euclidean space under any practical situation, a criticism of self-calibration. In Proc. 6th International Conference on Computer Vision, Bombay, India, pages 790-796, January 1998.
[3] Kanatani, K., Matsunaga, C.: Closed-form expression for focal lengths from the fundamental matrix. Proc. Asian Conference on Computer Vision, vol. 1, pp. 128-133, 2000.
[4] Kanatani, K., Nakatsuji, A., Sugaya, Y.: Stabilizing the focal length computation for 3D reconstruction from two uncalibrated views. The International Journal of Computer Vision, 66(2), pp. 109-122, 2006.
[5] Sturm, P.: On focal length calibration from two views. Proc. IEEE International Conference on Computer Vision and Pattern Recognition, pp. 145-150, 2001.
[6] Sturm, P., Cheng, Z.L., Chen, P.C.Y., Poo, A.N.: Focal length calibration from two views: method and analysis of singular cases. Computer Vision and Image Understanding, 99(1), pp. 58-95, 2005.
[7] Stewénius, H. et al.: "A Minimal Solution for Relative Pose with Unknown Focal Length". In: Computer Vision and Pattern Recognition. 2005.
[8] Kukelova, Z.: PhD Thesis "Algebraic Methods in Computer Vision". Thesis Advisor: Tomas Pajdla. Center for Machine Perception. Dept. of Cybernetics. FEE, CTU in Prague, 2013.

Bachelor Project Supervisor: Ing. Tomáš Pajdla, Ph.D.

Valid until: the end of the summer semester of academic year 2017/2018

L.S.

prof. Dr. Ing. Jan Kybic   prof. Ing. Pavel Ripka, CSc.
  Head of Department   **Dean**

Prague, January 6, 2017

Acknowledgements

I would like to thank my advisor Tomáš Pajdla for introducing me into the world of research and his numerous patient explanations of how things work and how the work is done. I would also like to thank Zuzana Kúkelová for precious comments and suggestions.

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 instructions for observing the ethical principles in the preparation of university 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

Abstract

The problem of automatically computing focal lengths of a pair of cameras from corresponding pair of images has long been a daunting task for 3D reconstruction community. A number of methods were developed, but the commonly held view is that neither of them works good enough to be used in practical situations. We focus on the particular task of computing focal lengths from the point correspondences, which we deem to be the missing link for the problem solution.

We especially focus on existing algebraic solvers for computing the fundamental matrix and the Bougnoux formula for computing the focal lengths therefrom. We survey these methods, as well as iterative methods [10, 4] proposed as their extensions, and analyze their performance. Our results show that the number of imaginary estimates, as well as the error of the estimation, declines with growing number of correspondences used. Moreover, based on our analysis we suggest that the computation of the ratio of focal length $r = f_2/f_1$ is more robust than computation of f_1 or f_2 alone. We propose an improvement to the solver of [10] based on this suggestion.

We furthermore assess performance of the methods in degenerate situations, and show that for bigger levels of noise the effect of the degeneracies significantly decreases. Specifically, the degenerate case of intersecting optical axes is shown to almost vanish for realistic levels of noise.

We finally analyze the problem of computing focal length from the theoretical standpoint of algebraic geometry, and give two new formulae for computing camera focal length from a fundamental matrix. We show that using the right of them might help to avoid a known degeneracy. Specifically, the degeneracy where 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, and where Bougnoux ([3]) formula fails can in some cases be avoided. The degeneracy reduces to the case where all three formulae fail.

Keywords: computer vision, 3D reconstruction, minimal problems, focal length, Gröbner basis

Abstrakt

Problém automatického výpočtu ohniskových vzdáleností dvou kamer z odpovídajících obrázků je stále obtížnou úlohou pro komunitu 3D rekonstrukce. Pro vyřešení tohoto problému byla navržena řada metod. Má se ale za to, že žádná z nich nefunguje natolik dobře, aby mohla být použita v praktické situaci. V této tezi se proto zaměříme na úlohu výpočtu ohniskové vzdálenosti z korespondencí v obrázku, kterou považujeme za chybějící článek pro vyřešení problému.

Zejména se zaměříme na existující algebraické solvery pro výpočet fundamentální matice, a na Bougnouxův vzorec, který z ní vypočte ohniskovou vzdálenost. Tyto metody prozkoumáme a zanalyzujeme jejich výkonnost. Ukážeme, že počet imaginárních odhadů, jakož i chyba odhadu ohniskové vzdálenosti klesají s rostoucím počtem použitých korespondencí. Rovněž zanalyzujeme degenerace metod a jejich efektivitu v degenerovaných situacích, stejně jako výkon existujících iteračních solverů [10, 4], a navrhneme zlepšení solveru [10].

Dále provedeme analýzu problému výpočtu ohniskové vzdálenosti z hlediska algebraické geometrie. Ukážeme, že kromě Bougnouxova vzorce existují další dva vzorce pro výpočet ohniskové vzdálenosti kamery z fundamentální matice. Ukážeme, že použitím správného z těchto vzorců se můžeme v některých případech vyhnout známé degeneraci. Konkrétně takové, kde rovina definovaná baselineou a optickou osou jedné kamery je kolmá na rovinu definovanou baselineou a optickou osou druhé kamery, a kde Bougnouxův vzorec [3] selhává. Degenerace se redukuje na případ, kdy selhávají všechny tři vzorce.

Klíčová slova: počítačové vidění, 3D rekonstrukce, minimální problémy, ohnisková vzdálenost, Gröbnerovy báze

Contents

List of Figures

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.

1. Introduction

1.1. Motivation

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

Estimating epipolar geometry with unknown focal lengths is an important issue in practical problems. In the laboratory environment it is possible to calibrate the camera beforehand using established procedures [9]. 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 [25, 24, 21] to estimate camera external and internal parameters is to extract points of interests [2, 19] with tentative correspondences, and use RANSAC [23] 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 correspondences.

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 needed number of correspondences is 7. Hartley and Zisserman [9] 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 [11] also summarizes an algorithm by Longuet-Higgins [18] 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., 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 [10] argues that in many practical cases this approach cannot be used. Other methods [10, 27, 14] were proposed to solve this problem. Neither of these approaches, however, works decisively better.

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

1.2. Related work

Hartley and Silpa-Anan [10] develop an iterative algorithm which incorporates heuristic estimates (prior knowledge) of focal lengths. The authors use optimization on a certain cost function which includes the Sampson error, priors of focal lengths and priors of principal points. They show that by allowing the algorithm to move principal point better results can be obtained. The algorithm shows competitive, although not decisively better performance in comparison to 7pt approach. In this thesis, we propose an improvement to this algorithm and demonstrate better performance.

Chandraker [4] further explores the idea of using priors of focal lengths. He defines a simpler cost function, and uses epipolar constraints as hard constraints, instead of including Sampson error in the cost function. The work considers two same focal lengths, but the algorithm can be easily extended to incorporate two focal lengths.

Nakatsuji et al. [15] show that even when the two focal lengths are 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 the focal length is recomputed each time from the sample until a real focal length is found.

Kanazawa et al. [14] 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. [27], 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.

1.3. Thesis structure

Our contributions are presented in Chapters 3, 4 and 5.

In Chapter 2 we survey the basic concepts and establish notation for the thesis. In Chapter 3 we provide an analysis of the known methods for focal length computation and show their performance. In Chapter 4 we use algebraic geometry to further analyze the problem. In Chapter 5 we suggest improvements to the current methods using our analysis from Chapter 3. We also provide a survey of methods that use prior focal length information.

2. Basic notions

2.1. Camera geometry

The book [9] 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 [9].

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 $\sim$:

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

The 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 $\mathbb{P}^{n-1}$ 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 represented by a point from the space $\mathbb{P}^n$. 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 point $\mathtt{l} = \begin{bmatrix} a & b & c \end{bmatrix}^{\mathrm{T}}$.

The Fig. 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.
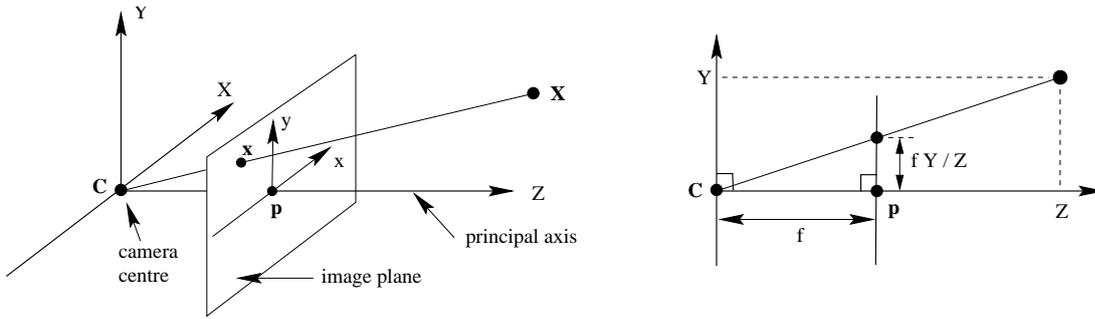
Figure 2.1. Projective camera model. Courtesy of [9]

Definition 2.1.6. *A iatrix[1]* K is a matrix of the form

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

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 Fig. 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 $\mathbf{x}_i = K\mathbf{X}_i$.

Extrinsic parameters of a camera i are described by a rotation matrix in the world space 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* P_i can be constructed from camera parameters and projects points $\mathbf{X}_i$ from world space $\mathbb{P}^3$ to the image plane $\mathbb{P}^2$:

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

.

2.1.2. Epipolar geometry

Epipolar geometry, one of the key topics of this thesis, is a geometry of two cameras seeing the same 3D object.

This section is described in depth in chapter 9 of [9]. 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 $R = R_2 R_1^T$. If $R, \mathbf{t}$ are known, extrinsics $R_1, R_2, \mathbf{C}_1, \mathbf{C}_2$ are defined up to a projective transformation.

Geometry of two cameras and a point in the world space seen by both cameras is usually described in terms of epipoles, which are the key concept of epipolar geometry.

Definition 2.1.8. *The baseline* is the line in 3D joining two camera centers.

---

[1]Throughout this work we assume unity aspect ratio and zero skew.
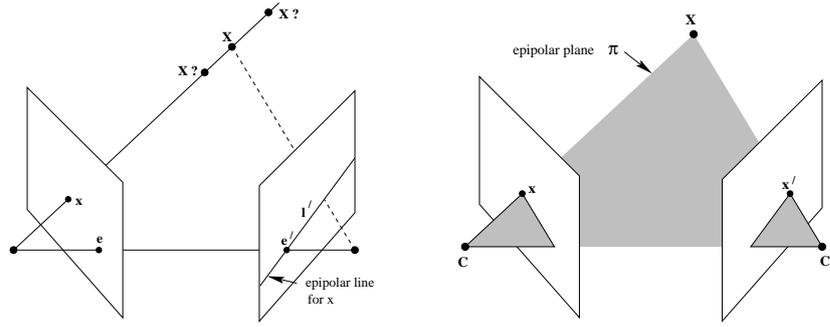
Figure 2.2. Camera pair. Courtesy of [9]

Definition 2.1.9. *The epipole* $\mathbf{e}_i$ *is the point in image plane that lies at the baseline. Equivalently, it is the point to which the center of another camera projects.*

Definition 2.1.10. *An epipolar line* $\mathbf{l}$ *is a line in image plane on which the epipole of that plane lies.*

Definition 2.1.11. *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 contains.*

The information about camera epipolar geometry can be described by a single matrix as defined next. If interested, see extended discussion and proofs in [9].

Definition 2.1.12. *An essential matrix* $\mathbf{E}$ *is a matrix of the form* $\mathbf{E} = \mathbf{R}\left[\mathbf{t}\right]_{\times}$. *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:

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

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):

$$2\mathbf{E}\mathbf{E}^{\mathrm{T}}\mathbf{E} - \mathrm{trace}\left(\mathbf{E}\mathbf{E}^{\mathrm{T}}\right)\mathbf{E} = 0. \tag{2.2}$$

The equations 2.1 and 2.2 together are called Demazure polynomials [20].

Essential matrix does not capture intrinsic camera parameters. For this purpose, the fundamental matrix is used.

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

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

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 $\mathsf{P}_1, \mathsf{P}_2$, 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 $\mathbf{l}_2 = \mathbf{F}\mathbf{x}_1$

Definition 2.1.14. We say that a point in one camera $\mathbf{x}_1$ and a point in another camera $\mathbf{x}_2$ form a *correspondence* if there exists a 3D point $\mathbf{X}$ that projects to $\mathbf{x}_1$ and $\mathbf{x}_2$.

For each correspondence $\mathbf{x}_1, \mathbf{x}_2$, the fundamental matrix satisfies the so-called Epipolar constraint:

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

As with essential matrices, a fundamental matrix can only be of rank two. On the other hand, every matrix of rank two is a fundamental matrix [9]. A fundamental matrix satisfies the rank constraint

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

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 the world space, it is possible in some cases to determine the positions of points X as well as intrinsic and extrinsic parameters of both cameras. 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 thereof were developed to this end, the most notable being the RANSAC family of algorithms [7].

2.2.1. Fundamental matrix computation

First, we need to estimate the fundamental matrix $\mathbf{F}$ of a camera pair. In fact, all the information we need is contained in this matrix. Various algorithms exist that can compute the matrix $\mathbf{F}$ 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 present can be formulated as so-called algebraic solvers, which work by way of constructing a set of constraints on the matrix $\mathbf{F}$ given the correspondences, and then solving the constraints in an algebraically precise way, giving all possible solutions. If the constraints are represented by polynomial equations the techniques of algebraic geometry can be used to solve these constraints.

We first define the 7pt algorithm. The algorithm uses (at least) 7 different 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[3].

The algorithm uses SVD procedure which serves as an implicit optimization of the epipolar constraints. In case that precisely 7 epipolar constraints are given, SVD can

---

[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 entering the system, which is $3 \cdot 1^7$.

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 B. Equivalently, the matrices F_1, F_2 will be a basis of the pace of matrices that satisfy seven epipolar constraints.

Algorithm 1: 7pt

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 F

begin

    Populate matrix $B \in \mathbb{R}^{n \times 9}$ with columns $\mathbf{b}_i$: $\mathbf{b}_i \leftarrow \mathrm{vec}(\mathbf{x}_{2,i} \otimes \mathbf{x}_{1,i})$;

    Take the right singular vectors $\mathbf{f}_1, \mathbf{f}_2$ corresponding to the two smallest singular values and transform them back to matrices $F_1, F_2 \in \mathbb{R}^{3 \times 3}$;

    if $\det(F_2) = 0$ **then**

        return F_2;

    else

        Solve the 3rd degree polynomial in x: $\det(xF_1 + F_2) = 0$, and choose the real roots x_i.;

        return $x_i F_1 + F_2$;

    end

end

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. No smaller number of correspondences can be used, however. It can be proven that the problem of computing fundamental matrices with two unknown focal lengths is inherently ill-posed when less than 7 correspondences are given. The algorithm thus solves the problem using minimal possible number of correspondences.

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, that is, it is more loyal to the data. We will compare the performance of this algorithm to 7pt in the later parts of the work. Note that this algorithm gives precisely one real solution.

Algorithm 2: 8pt

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 F

begin

    Populate matrix $B \in \mathbb{R}^{n \times 9}$ 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 $F \in \mathbb{R}^{3 \times 3}$;

    return F;

end

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 all calibration information is known, we can estimate essential matrix E instead of F. This can be done by transforming points so that effects of intrinsic camera parameters

are removed:

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

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

$$\tilde{\mathbf{x}}_2^T E\tilde{\mathbf{x}}_1 = \mathbf{x}_2^T K_2^{-T} E K_1^{-1}\mathbf{x}_1 = \mathbf{x}_2^T F\mathbf{x}_1 = 0.$$

The algebraic solver for the 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 describe it fully.

Nister et al. [22] show in detail how to construct an algebraic solver for this problem. The paper makes extensive use of theory of algebraic geometry [6, 5]. It is very instructing to read through the construction of this solver.

We will, however, use another solver, described by Kukelova et al. in [17]. The code can be found online at "http://cmp.felk.cvut.cz/mini/"[4]. This solver, contrary to the one by Nister 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 [16].

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 I_2 is defined as $I_2 = \text{diag}(1, 1, 0)$, and $\mathbf{e}_1, \mathbf{e}_2$ are the epipoles of the first and second image correspondingly.

$$f_1^2 = -\frac{\mathbf{p}_2^T\left[\mathbf{e}_2\right]_\times I_2 F\mathbf{p}_1\mathbf{p}_1^T F^T\mathbf{p}_2}{\mathbf{p}_2^T\left[\mathbf{e}_2\right]_\times I_2 F^T I_2 F\mathbf{p}_2} \tag{2.6}$$

$$f_2^2 = -\frac{\mathbf{p}_1^T\left[\mathbf{e}_1\right]_\times I_2 F^T\mathbf{p}_2\mathbf{p}_2^T F\mathbf{p}_1}{\mathbf{p}_1^T\left[\mathbf{e}_1\right]_\times I_2 F I_2 F^T\mathbf{p}_1} \tag{2.7}$$

The second equation may be derived from the first by switching indices and transposing the F matrix.

Note that it is not possible to find a focal length unless both principal points are known. As these formulae show, 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.

---

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

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.

Agarwal et al., in a thorough study [1], analyze by algebraic geometry techniques a degeneration when neither of 3 solutions that 7pt algorithm 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 the valid subset of the 3 solutions given by the 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 [13]. 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 Fig. 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:

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

$\square$

A further degeneracy [9] 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.

---

[5]We suppose that algorithm does not return zero solution $\mathtt{F} = 0_{3,3}$. Our formulation of the 7pt algorithm indeed does not naturally produce this solution.

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

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* 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. This might be a problem if the real solution F is a matrix on rank 1, and thus we don't have any valid result.

Hartley shows in [10] that the focal lengths computed may be imaginary, which again leads to impossible camera configuration. This happens because the focal length enter the Bougnoux formulae (Eq. 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.

Another problem is described in conjunction with the notion of *cheirality* in [12]. 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 constructed that 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, 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 [6][5]. 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 [6] 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 an 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 [6] 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.

---

[7]When the ideal concerned is a so-called radical ideal [6].

3. Experimental analysis of existing solutions

In this chapter we provide an assessment of the the quality and stability of the 7pt algorithm (algorithm 1) with the Bougnoux formula (Eq. 2.6). We consider this pipeline as essentially integral baseline procedure for estimation of the focal lengths from image correspondences.

3.1. 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 in a way that principal points are always in the center. This means that the calibration matrices after this preprocessing are of shape

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

Different cameras are allowed to have different focal lengths.

3.1.1. Evaluation procedure

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

Algorithm 3: Workflow

Data: The number of correspondences used n, noise 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 (algorithm 1)). Leave only matrices with real elements;

8   Choose the estimation that optimizes Sampson error (equation 3.2) 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
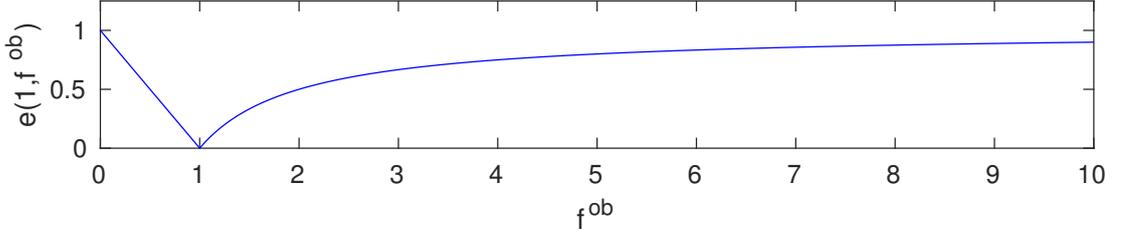
Figure 3.1. Error function 3.1 for the focal length, where f_{gt} is fixed to 1.

3.1.2. 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

$$e(f_{gt}, f_{ob}) = 1 - \frac{\min(f_{gt}, f_{ob})}{\max(f_{gt}, f_{ob})}. \tag{3.1}$$

The error function for fixed ground truth value $f_{gt} = 1$ is shown in the Fig. 3.1.

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

$$S(\mathbf{F}) = \sum_i \frac{(\mathbf{x}_2^\mathrm{T} \mathbf{F} \mathbf{x}_1)^2}{(\mathbf{F}\mathbf{x}_1)_1^2 + (\mathbf{F}\mathbf{x}_1)_2^2 + (\mathbf{x}_2^\mathrm{T}\mathbf{F})_1^2 + (\mathbf{x}_2^\mathrm{T}\mathbf{F})_2^2}. \tag{3.2}$$

3.2. Performance analysis for generic situations

In this section we analyze the focal length computation performance for a generic scene, in which a degeneracy is unlikely to occur.

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

3.2.1. 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 median of focal lengths computation error e and its 0.25 and 0.75 quantiles. In these experiments we discard each all imaginary estimates.

The Fig. 3.2 shows a rapid decline in error with growing number of correspondences. This is explained by the fact that both algorithms use SVD procedure to optimize error on epipolar constraints. With bigger number of correspondences the error caused by random noise is averaged and the estimation becomes more stable.

The Fig. 3.3 shows the growth of the error with increasing noise. The number of correspondences used is 8 for both algorithms. We see that the error growth is approximately linear and quite mild, e.g., with additive noise of 3 pixels the error is still bearable. For realistic values of noise around 1 pixel, the error is reasonably small.

Another important trend is shown in the Fig. 3.4 is the decrease of number of imaginary estimates with increasing number of correspondences used. We suggest that imaginary focal lengths are due to high level of noise, and appear less often when the noise is averaged over a big number of correspondences.
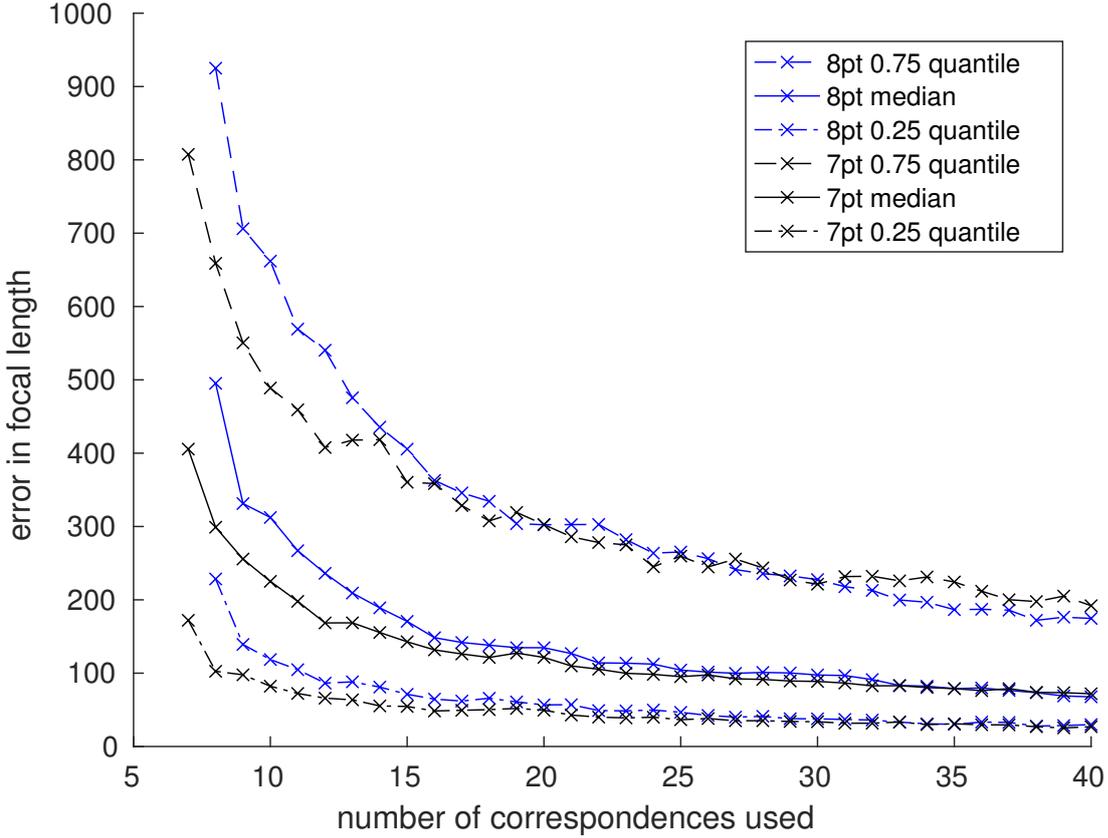
Figure 3.2. A Comparison of the 7pt and 8pt algorithms against the number of correspondences used. Median error $f - f^{true}$ in focal length and two its 0.25, 0.75 quantiles are shown. The ground truth f^{true} is equal to 1500. The level of noise σ is equal to 1. Imaginary estimates are excluded.

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, however, interestingly enough, the 8pt algorithm gives less imaginary estimates.

3.2.2. Ratio of focal lengths is robust

We will continue to analyze the 7pt algorithm, as it apparently performs better than the 8pt algorithm on average. The 7pt algorithm also has the theoretical advantage in that it never gives a matrix of rank 3 (remember that a fundamental matrix is always of rank 2).

We show empirically that the ratio $r = f_2/f_1$ computation is more robust than computation of f_1 or f_2 alone. To show it we conduct an experiment with 300 different noise samples applied to the same camera geometry. We exclude extreme outliers.

The Fig. 3.5 shows scattered estimated points in the space (f_1, f_2). If the distribution of estimates in this space doesn't have additional structure, we expect these estimations to lie around the ground truth, perhaps similar to a two-dimensional Gaussian distribution. However, it can be seen that while the deviation of focal lengths from the ground truth point is relatively big, the points clearly tend to the line $f_2/f_1 = 3/4$
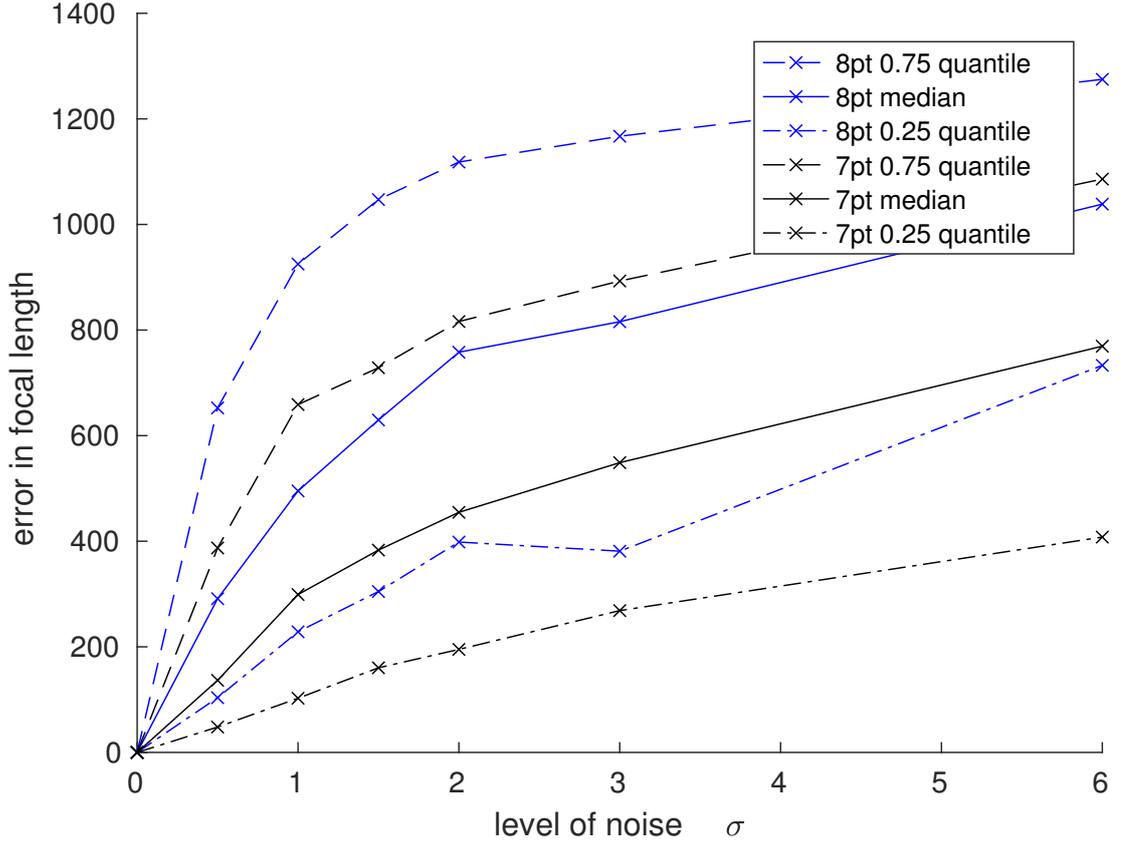
Figure 3.3. A comparison of the 7pt and 8pt algorithms against noise level σ. The ground truth f^{true} is equal to 1500. 40 correspondences are used.

(given in blue), which is indeed the ground truth for the ratio. Imaginary estimates are also shown in absolute values, which, interestingly, also tend to have the right ratio. We will expand on this in the next section.

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

3.2.3. Imaginary focal lengths

We empirically show that the ratio of the focal lengths is robust even when both focal lengths are imaginary. The probability of getting an imaginary focal length also isn't connected to distance between optical axes.

Fig. 3.6 shows the experiment. The figure is the histogram of different errors in ratio (measured as $f_2/f_1 - (f_2/f_1)_{gt}$). A small number of outliers lies far off the graph. Clearly, the likelihood that a focal lengths ratio is estimated better is bigger when the focal lengths are real. However, the Fig. 3.6 shows that the ratios computed from imaginary focal lengths also fall reasonably close to the ground truth, in the sense that the mean of their distribution seem to converge to the ground truth ratio.

The distribution of the imaginary focal lengths itself is shown in the Fig. 3.8. We see that for bigger values of error, the number of imaginary estimates is roughly equal to the number of real ones, in other words, the likelihood of getting an imaginary focal length estimate is around 50% when the error in focal length estimate is bigger than
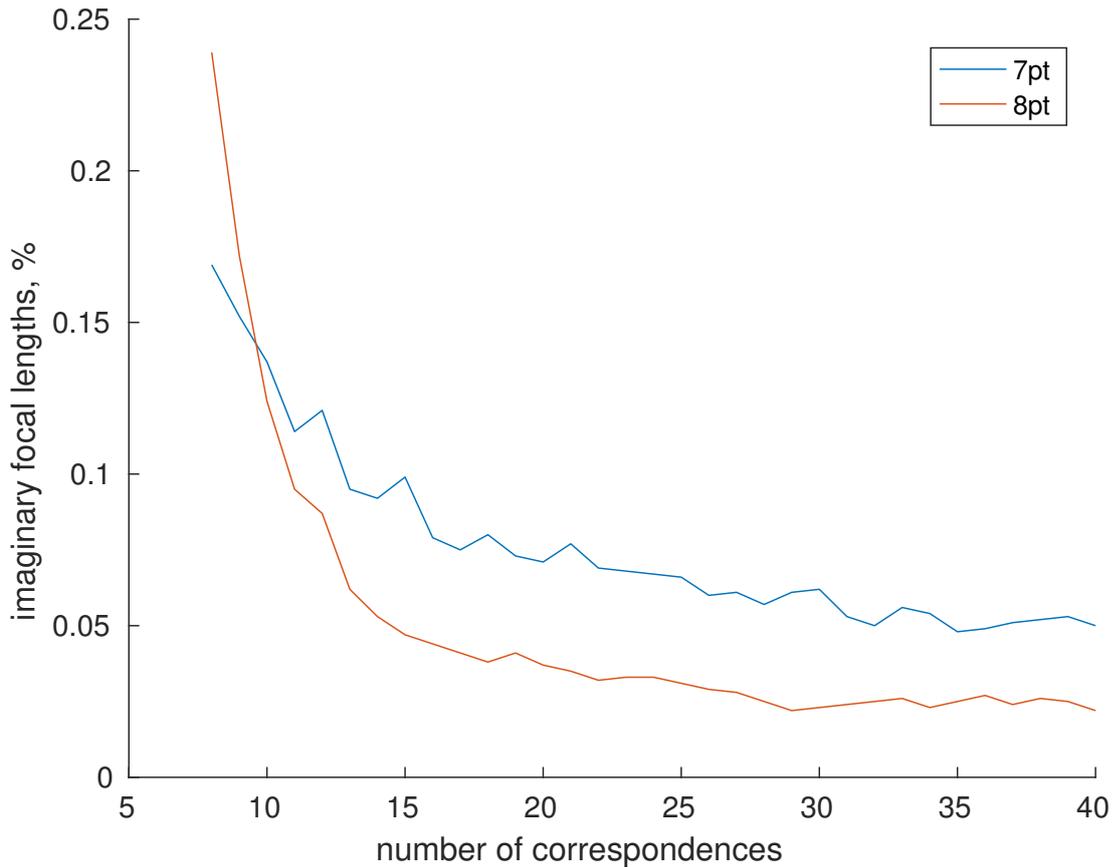
Figure 3.4. The fraction of imaginary focal length estimates by the 7pt and 8pt algorithms. The noise level σ in image measurements is equal to 1.

twofold.

We suggest that there is a hidden variable that is susceptible to noise and makes both focal lengths imaginary when the level of noise is high. It doesn't, however, 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 in a way that the higher the noise is the more probable it is that the sign will be negative.

One of our guesses was that the probability of having imaginary estimate can increase when the situation becomes close the degenerate. In the next experiment we assess the degree of degeneracy by the length of shortest transversal between the optical axes. When the length is zero, the optical axes intersect.

The Fig. 3.7 shows the estimated ratios $r = f_2/f_1$. We see that the fraction of imaginary estimates does not depend on the length of shortest transversal. We also see how both ratios computed from real and imaginary estimates tend to the ground truth. We plot the ratios estimated from imaginary values of the focal lengths with negative sign and in red to distinguish it from the ratios estimated from the real values, which are shown in blue.
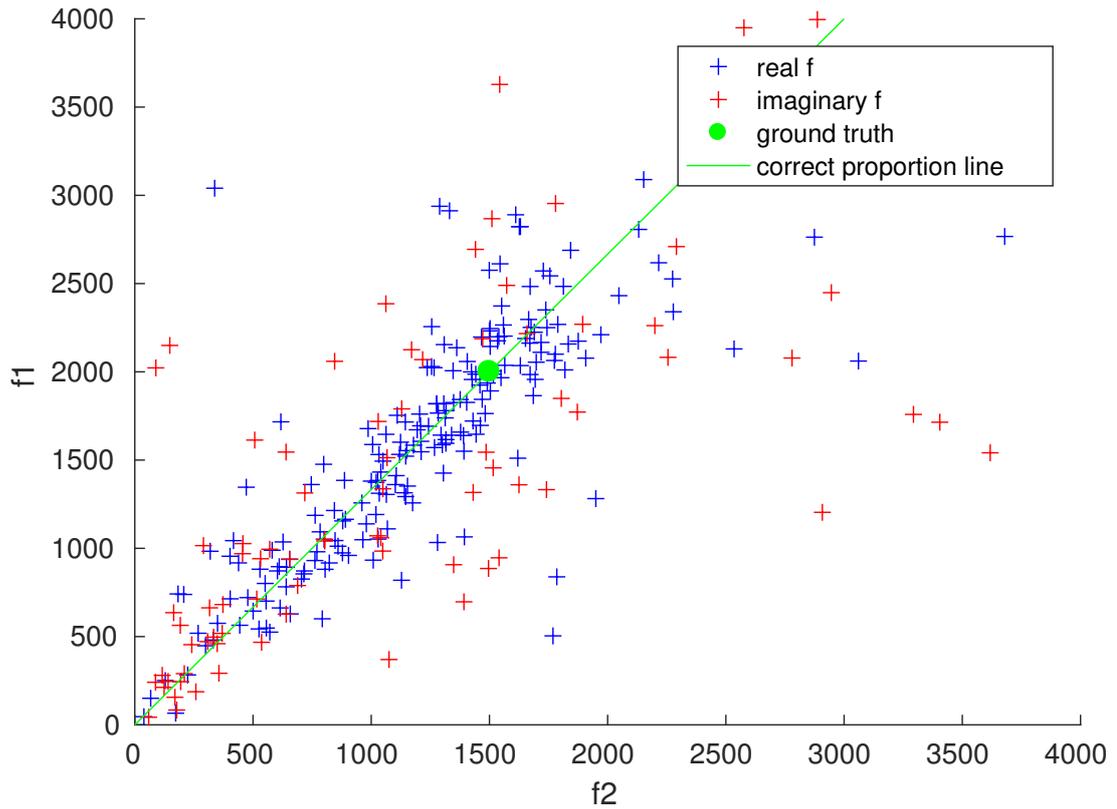
Figure 3.5. The scatter plot of the focal lengths (f_1, f_2) estimates by the 7pt algorithm for one scene. Line through ground truth ratio $r = f_2/f_1$ is given for reference. Ground truth point also shown in green. Imaginary estimates are shown in absolute value. The noise level σ in image measurements in image measurements is equal to 1. Seven correspondences are used.
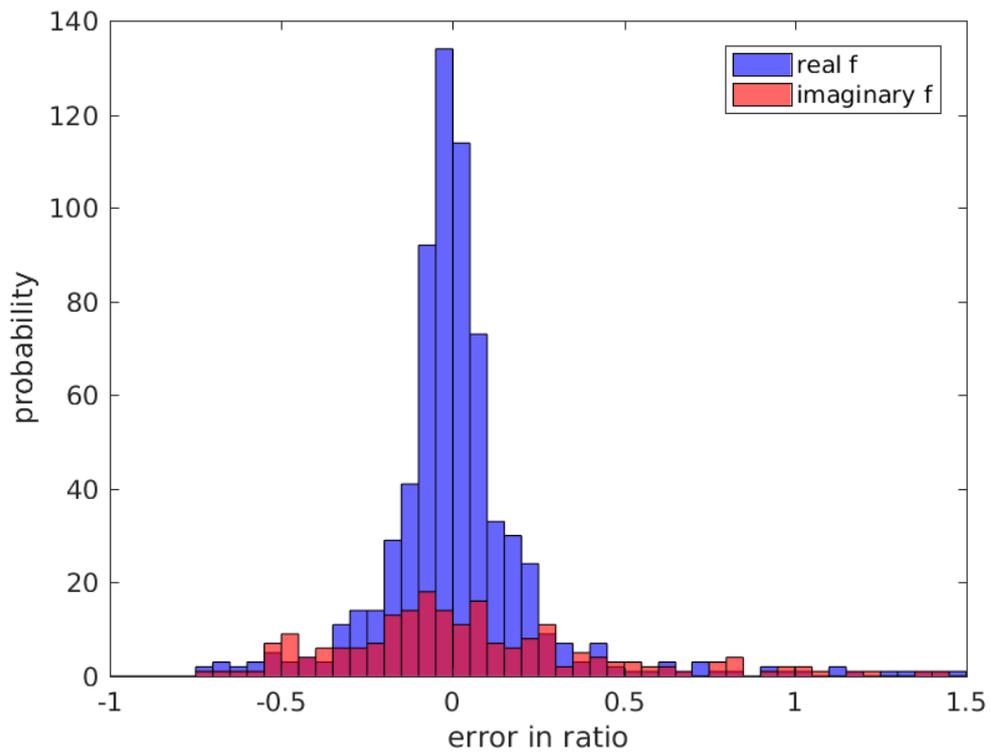
Figure 3.6. The histogram of errors in the ratio of focal length estimations by 7pt algorithm, i.e. $f_2/f_1 - f_2^{true}/f_1^{true}$. The cases with at least one imaginary focal length are shown in a separate histogram (orange), the cases where both focal length were real are shown in blue. Red color is shown where an orange is supersimposed over blue. 40 correspondences are used, and level of noise σ is equal to 1. A small number of outliers lies far off the graph.
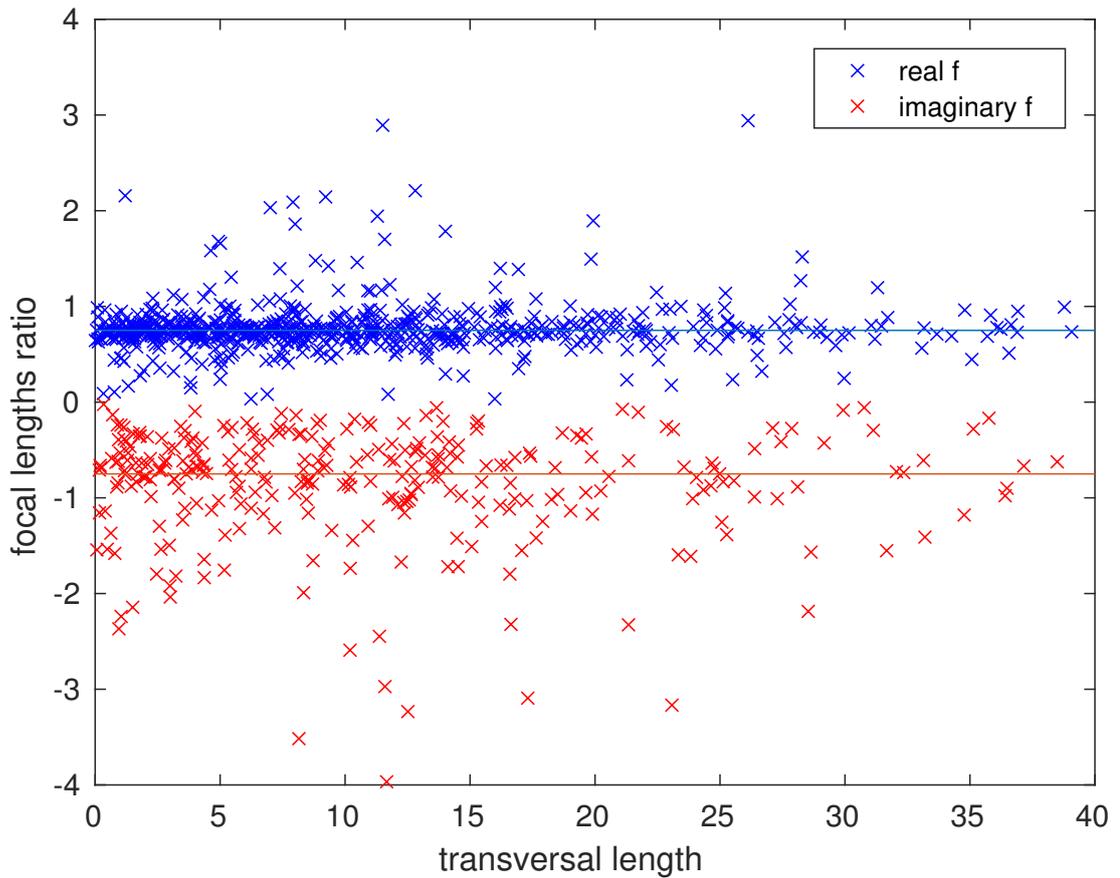
Figure 3.7. Scatter plot of focal lengths ratio $r = f_2/f_1$ estimates produced by 7pt algorithm. Ratios corresponding to imaginary focal lengths are plotted as negative to distinguish them (they are positive). On X axis is the distance between optical axes. A small number of outliers lies far off the graph.. 40 correspondences are used.
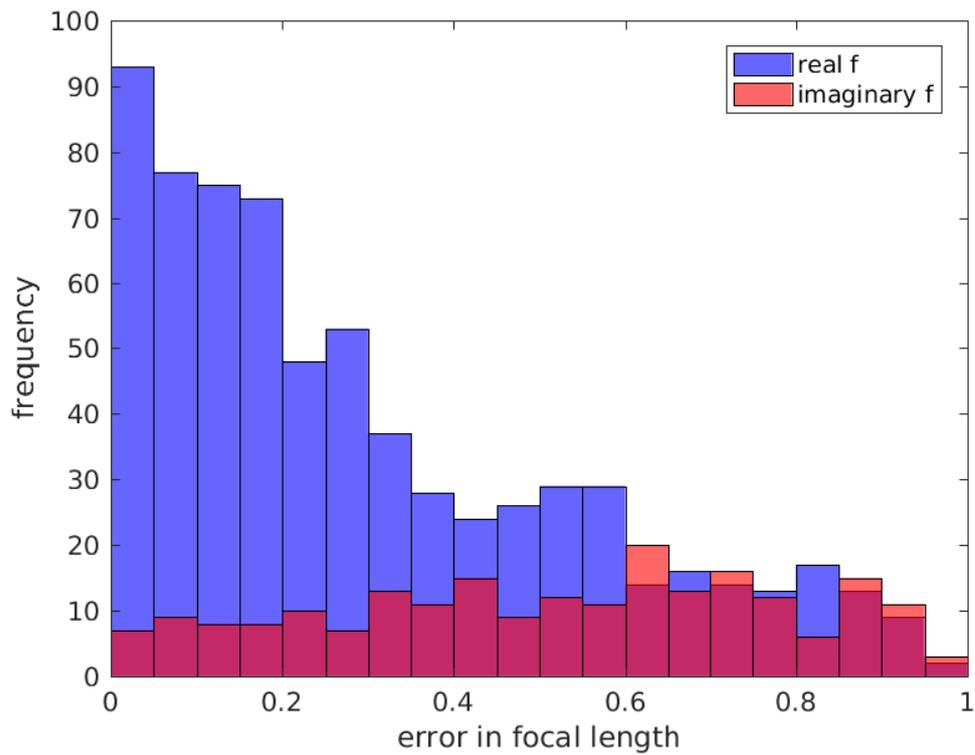
Figure 3.8. The distribution of multiplicative errors (equation 3.1) in the focal lengths, produced by 7pt algorithm. Separate distributions are shown for the case with both real focal lengths (blue) and for the case with at least one imaginary focal length (orange). Red color is shown where the orange is supersimposed over blue. 40 correspondences are used, and level of noise σ is equal to 1.

3.3. Performance analysis for close to degenerate situations

In theory, there does not exist a method to recover focal length when the optical axes intersect. With a small amount of noise in the image measurements, however, the configuration no longer is singular. The focal lengths thus can be computed in almost all practical situations. Nonetheless, we expect pose estimation to deteriorate as the configuration becomes close to singular due to numerical instability. In this section we show the extent of this deterioration. For the fundamental matrix estimation, we will use the 7pt algorithm, as it apparently performs better than the 8pt algorithm.

3.3.1. Intersecting optical axes

In this scenario, two 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. 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 the cameras to the scene points is 5 m.

The Fig. 3.9 shows the distribution of multiplicative errors in the focal length estimates by the 7pt algorithm. The figure shows 4 different levels of noise, and each line represent a distance d. We see that for small values of noise, $\sigma \ll 0.1$ px, the distance between the optical axes significantly affects 7pt algorithm performance. The performance deteriorates greatly with decreasing distance between optical axes. For more realistic noise levels, however, the Fig. 3.9 shows that the distance has little impact on the quality of estimation. With 1 pixel noise, the performance is almost the same and axes that are 10 cm or 1 cm apart. 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 in image measurements), with 75% of estimations being off by a factor of at most 2.

We suggest this happens because the noise in image measurements drives the system further away from a degenerate situation.

3.3.2. Parallel optical axes

An interesting configuration is when the optical axes intersect at a point at infinity, i.e., 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 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 angles α and levels of noise σ.

Fig. 3.10 shows the results. We see that the quality of estimation is much worse for truly parallel axes and small noise in image correspondences does not save the situation. However, the degeneracy almost disappears already when the angle α reaches 0.1°.

3.3.3. Conclusions

We have shown that the 7pt algorithm performs better for most applications. However, interestingly, 8pt algorithm gives less imaginary estimates.

Using more correspondences allows 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.
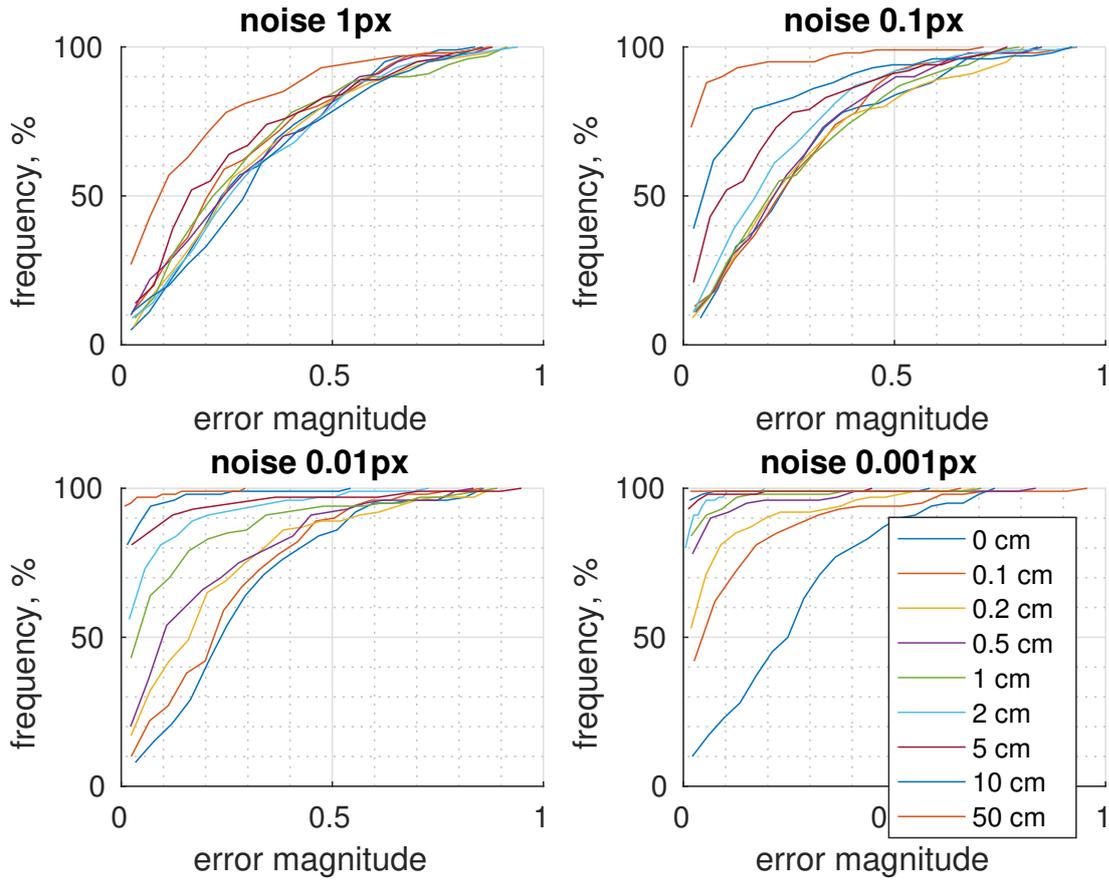
Figure 3.9. The cumulative distribution of the estimated focal length multiplicative error (equation 3.1) for nearly intersecting optical axes. Different lines show different distances between optical axes in centimeters. 7pt algorithm is used

We confirmed that the imaginary focal lengths are indeed signs of a highly corrupt solution, and the fundamental matrices that decompose into imaginary focal lengths should be discarded.

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.
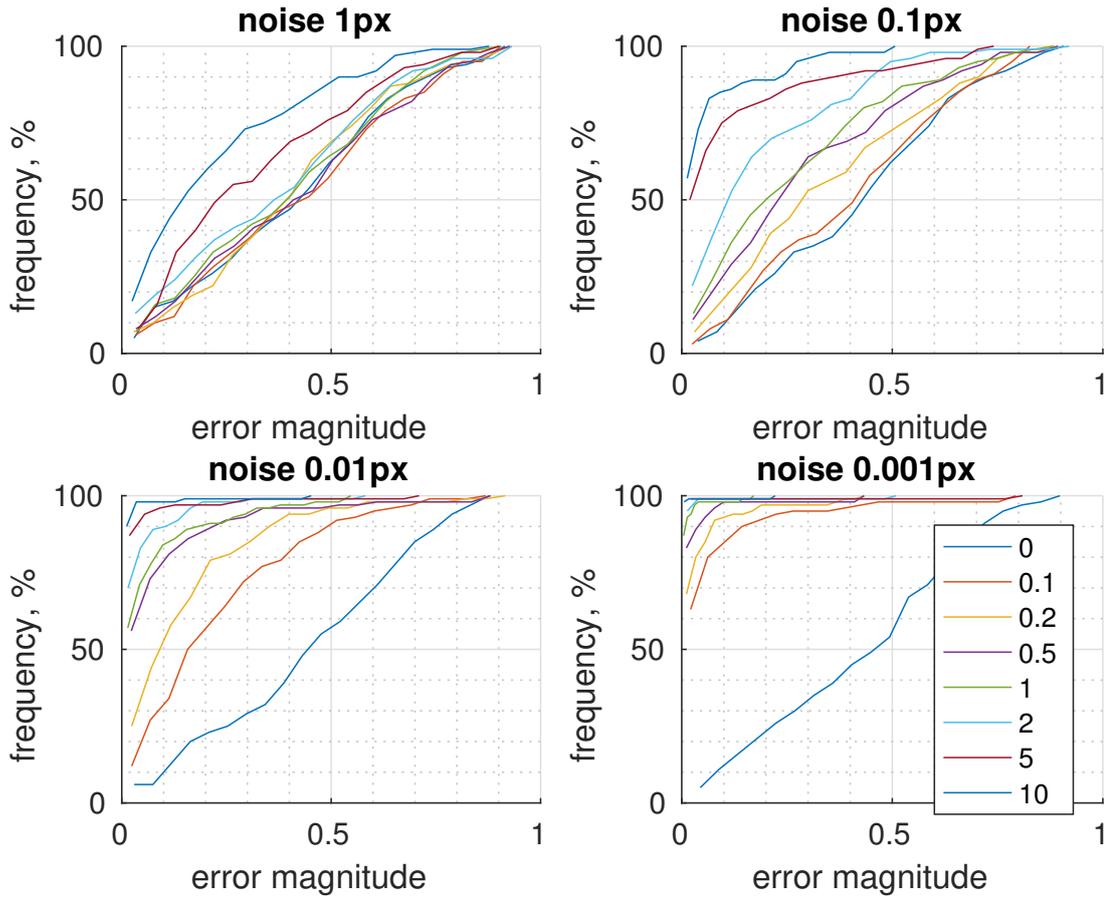
Figure 3.10. The cumulative distribution of the estimated focal length multiplicative error (equation 3.1) for nearly parallel optical axes. Different lines show different angles between optical axes in degrees. 7pt algorithm is used

4. Algebraic analysis of focal lengths computation

In this section we analyze the methods for computing focal lengths from point correspondences by the algebraic geometry techniques. We show that indeed three and only three fundamental matrices (some possibly complex or of rank 1) can be derived using algebraic geometry. We analyze the degeneracies of the computation. We also show that three different formulae exist for computing focal lengths from a Fundamental matrix. One of these is the Bougnoux formula 2.6, and two others weren't known before.

4.1. Analysis

We analyze the set of valid fundamental matrices of camera pairs calibrated up to focal lengths using techniques of algebraic geometry. To this end, we use Macaulay2 [8], 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 language, we will use in this section Macaulay2 code snippets directly. 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 field of rational numbers $\mathbb{Q}^{11}$ of 11 unknowns, 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 it can remove large spurious components corresponding to cases when $f_1 = 0$ or $f_2 = 0$, which can't happen in any real camera systems.

Below is a snippet of Macaulay2 code explaining 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
Et = transpose E
G = ideal(det(E)) + minors(1, 2*E*Et*E - trace(E*Et)*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 -- dimension, codimension and degree
```

The ideal Gs contains constraints 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, when considered in the algebraically closed field $\mathbb{C}$, 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. The solutions are actually one-dimensional subspaces, and therefore we would expect the ideal Gs to have dimension 8, which it indeed has.

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 the above, we see that the ideal M has only one generator - the det(F) polynomial. This means that the algebraic constraints 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 rank constraint 2.1 is needed to solve the system, and the trace constraint is extraneous. The trace constraint, however, allow us to compute the focal lengths.

4.2. Computing focal lengths

In the next snippet we 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. It turns out that the Gröbner basis of the eliminated ideal s_i contains the determinant of F and additional three polynomials, from which a formula for computing the focal length that wasn't eliminated can be deduced. 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,

$$
f_2^2 \;=\; -\frac{c23_{0,1}}{c23_{0,0}} \;=\; -\frac{\mathrm{F}_{3,3}(\mathrm{F}_{1,1}\mathrm{F}_{2,3}\mathrm{F}_{3,1} + \mathrm{F}_{1,2}\mathrm{F}_{2,3}\mathrm{F}_{3,2} - \mathrm{F}_{1,3}\mathrm{F}_{2,1}\mathrm{F}_{3,1} - \mathrm{F}_{1,3}\mathrm{F}_{2,2}\mathrm{F}_{3,2})}{(\mathrm{F}_{1,1}^2\mathrm{F}_{1,3}\mathrm{F}_{2,3} - \mathrm{F}_{1,1}\mathrm{F}_{1,3}^2\mathrm{F}_{2,1} + \mathrm{F}_{1,1}\mathrm{F}_{2,1}\mathrm{F}_{2,3}^2 + \mathrm{F}_{1,2}^2\mathrm{F}_{1,3}\mathrm{F}_{2,3}}.\tag{4.1}
$$
$$
{\quad\; - \mathrm{F}_{1,2}\mathrm{F}_{1,3}^2\mathrm{F}_{2,2} + \mathrm{F}_{1,2}\mathrm{F}_{2,2}\mathrm{F}_{2,3}^2 - \mathrm{F}_{1,3}\mathrm{F}_{2,1}^2\mathrm{F}_{2,3} - \mathrm{F}_{1,3}\mathrm{F}_{2,2}^2\mathrm{F}_{2,3})}
$$

4. Algebraic analysis of focal lengths computation

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 other formulae are:

$$f_2^2 = -\frac{c22_{0,1}}{c22_{0,0}} = -\frac{F_{3,3}(F_{1,1}F_{3,1}F_{3,3} + F_{1,2}F_{3,2}F_{3,3} - F_{1,3}F_{3,1}^2 - F_{1,3}F_{3,2}^2)}{(F_{1,1}^2F_{1,3}F_{3,3} - F_{1,1}F_{1,3}^2F_{3,1} + F_{1,1}F_{2,1}F_{2,3}F_{3,3} + F_{1,2}^2F_{1,3}F_{3,3}} \\ \overline{\quad - F_{1,2}F_{1,3}^2F_{3,2} + F_{1,2}F_{2,2}F_{2,3}F_{3,3} - F_{1,3}F_{2,1}F_{2,3}F_{3,1} - F_{1,3}F_{2,2}F_{2,3}F_{3,2})}}.$$

$$(4.2)$$

$$f_2^2 = -\frac{c21_{0,1}}{c21_{0,0}} = -\frac{F_{3,3}(F_{2,1}F_{3,1}F_{3,3} + F_{2,2}F_{3,2}F_{3,3} - F_{2,3}F_{3,1}^2 - F_{2,3}F_{3,2}^2)}{(F_{1,1}F_{1,3}F_{2,1}F_{3,3} - F_{1,1}F_{1,3}F_{2,3}F_{3,1} + F_{2,1}^2F_{2,3}F_{3,3} - F_{2,1}F_{2,3}^2F_{3,1}} \\ \overline{\quad + F_{1,2}F_{1,3}F_{2,2}F_{3,3} - F_{1,2}F_{1,3}F_{2,3}F_{3,2} + F_{2,2}^2F_{2,3}F_{3,3} - F_{2,2}F_{2,3}^2f_{3,2})}}. \quad (4.3)$$

The formulae for the other focal length 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 of stability and quality of estimations. However, as the denominators of the formulae are different, each of them has its own degeneracies, where another one may succeed instead. We proceed to give examples of such situations.

It can be shown by substitution that the Bougnoux formula vanishes whenever either first or second column of F is the zero vector. Of the formulae we found, however, the third formula (Eq. 4.3) does not necessarily vanish when the first column of F is zero, and the second formula (Eq. 4.2) does not necessarily vanish when the second column is zero. When the third column is zero, all the formulae either vanish or become $p(F) f_2^2 = 0$. The focal lengths cannot be reconstructed in this case per lemma 2.2.1, as $F_{3,3} = 0$.

Example 4.2.1. The pair of cameras P_1, P_2 both have their optical axes almost perpendicular to the baseline (tilted by an angle of $\pi/10$) and experience the degeneration where 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 [9]. The calibration matrices are identity matrices. More precisely, the fundamental matrix of the pair is:

$$R^{tilt} = \begin{bmatrix} \cos(\pi/10) & 0 & \sin(\pi/10) \\ 0 & 1 & 0 \\ -\sin(\pi/10) & 0 & \cos(\pi/10) \end{bmatrix}$$

$$F = R * [t]_\times = R^{tilt} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} R^{tilt} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \approx \begin{bmatrix} 0 & 0.3 & 0.3 \\ 0 & 0.9 & 0 \\ 0 & 0.1 & 0.9 \end{bmatrix}$$

By substitution, the first and second (Eqs. 4.2, 4.1) vanish on this example,, i.e., assume the form $0\ f_2^2 = 0$, but the third formula (Eq. 4.3) successfully gives a correct focal length estimate.

The example 4.2.1 shows that there among all configurations known to be degenerate for the Bougnoux formula, there exist some that can be successfully solved by using the correct formula of the existing three. We see that specifically the degeneration 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 can be, at least in

some cases, solved. We also see that in the example 4.2.1 the configuration corresponds to a fundamental matrix which has the first column equal to zero vector.

Note that there still remains a 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). The next example shows when this degeneracy might occur.

Example 4.2.2. The pair of cameras P_1, P_2 are similar as in the example 4.2.1, but without tilt, i.e., they both have their optical axes perpendicular to the baseline. They also experience the degeneration where 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. The calibration matrices are identity matrices. More precisely, the fundamental matrix of the pair is:

$$\mathtt{F} = \mathtt{R} * [t]_\times = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

By substitution, all three formulae (Eqs. 4.3 4.2, 4.1) vanish on this example, i.e., assume the form $0\ f_2^2 = 0$

Based on the examples 4.2.1, 4.2.2 we *conjecture* that the remaining degeneracy is the case when the both camera have their optical axes perpendicular to the baseline, and also experience the degeneration where 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 all three formulae should fail.

4.3. The ratio formula

We present a formula to compute directly the ratio $r = f_2/f_1$. Note that unlike focal length formulae, there can be only one algebraically independent formula for computing the ratio.

$$r = \frac{f_2}{f_1} = \frac{\begin{aligned} &\mathtt{F}_{1,1}^2\mathtt{F}_{3,1}^2\mathtt{F}_{3,3} + 2\mathtt{F}_{1,1}\mathtt{F}_{1,2}\mathtt{F}_{3,1}\mathtt{F}_{3,2}\mathtt{F}_{3,3} - \mathtt{F}_{1,1}\mathtt{F}_{1,3}\mathtt{F}_{3,1}^3 - \mathtt{F}_{1,1}\mathtt{F}_{1,3}\mathtt{F}_{3,1}\mathtt{F}_{3,2}^2 \\ &+ \mathtt{F}_{1,2}^2\mathtt{F}_{3,2}^2\mathtt{F}_{3,3} - \mathtt{F}_{1,2}\mathtt{F}_{1,3}\mathtt{F}_{3,1}^2\mathtt{F}_{3,2} - \mathtt{F}_{1,2}\mathtt{F}_{1,3}\mathtt{F}_{3,2}^3 + \mathtt{F}_{2,1}^2\mathtt{F}_{3,1}^2\mathtt{F}_{3,3} + 2\mathtt{F}_{2,1}\mathtt{F}_{2,2}\mathtt{F}_{3,1}\mathtt{F}_{3,2}\mathtt{F}_{3,3} \\ &- \mathtt{F}_{2,1}\mathtt{F}_{2,3}\mathtt{F}_{3,1}^3 - \mathtt{F}_{2,1}\mathtt{F}_{2,3}\mathtt{F}_{3,1}\mathtt{F}_{3,2}^2 + \mathtt{F}_{2,2}^2\mathtt{F}_{3,2}^2\mathtt{F}_{3,3} - \mathtt{F}_{2,2}\mathtt{F}_{2,3}\mathtt{F}_{3,1}^2\mathtt{F}_{3,2} - \mathtt{F}_{2,2}\mathtt{F}_{2,3}\mathtt{F}_{3,2}^3 \end{aligned}}{\begin{aligned} &\mathtt{F}_{1,1}^2\mathtt{F}_{1,3}^2\mathtt{F}_{3,3} - \mathtt{F}_{1,1}\mathtt{F}_{1,3}^3\mathtt{F}_{3,1} + 2\mathtt{F}_{1,1}\mathtt{F}_{1,3}\mathtt{F}_{2,1}\mathtt{F}_{2,3}\mathtt{F}_{3,3} - \mathtt{F}_{1,1}\mathtt{F}_{1,3}\mathtt{F}_{2,3}^2\mathtt{F}_{3,1} \\ &+ \mathtt{F}_{1,2}^2\mathtt{F}_{1,3}^2\mathtt{F}_{3,3} - \mathtt{F}_{1,2}\mathtt{F}_{1,3}^3\mathtt{F}_{3,2} + 2\mathtt{F}_{1,2}\mathtt{F}_{1,3}\mathtt{F}_{2,2}\mathtt{F}_{2,3}\mathtt{F}_{3,3} - \mathtt{F}_{1,2}\mathtt{F}_{1,3}\mathtt{F}_{2,3}^2\mathtt{F}_{3,2} \\ &- \mathtt{F}_{1,3}^2\mathtt{F}_{2,1}\mathtt{F}_{2,3}\mathtt{F}_{3,1} - \mathtt{F}_{1,3}^2\mathtt{F}_{2,2}\mathtt{F}_{2,3}\mathtt{F}_{3,2} + \mathtt{F}_{2,1}^2\mathtt{F}_{2,3}^2\mathtt{F}_{3,3} - \mathtt{F}_{2,1}\mathtt{F}_{2,3}^3\mathtt{F}_{3,1} + \mathtt{F}_{2,2}^2\mathtt{F}_{2,3}^2\mathtt{F}_{3,3} - \mathtt{F}_{2,2}\mathtt{F}_{2,3}^3\mathtt{F}_{3,2} \end{aligned}} \tag{4.4}$$

Computing the ratio directly from F can offer greater speed and more stability than computing it from the focal lengths. Indeed, we observed that for close to degenerate situations, the formula, although marginally, is a better estimator.

4.4. Computing Fundamental matrices

We show in detail how the 7pt solver works from the viewpoint of algebraic geometry.

We first simulate a random set of correspondences. We skip this code for brevity and assume that the matrix B is a matrix from the 7pt algorithm (algorithm 1), i.e., the matrix, the right null space of which is the linear space of all (vectorized) matrices satisfying 7 certain epipolar constraints 2.4.

4. Algebraic analysis of focal lengths computation

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 simulated 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 together the ideals Gs and IB we combine the constraints and obtain an ideal that corresponds to a variety of solutions of the 7pt solver for our correspondences.

Variety $\mathbf{V}(GBs)$ can consist of either 5 or 13 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.

We provide an example of what the minimal primes (ideals, corresponding to component varieties) may look like.

Example 4.4.1. The *minimal primes* of an ideal containing seven epipolar constraints and the rank constraint. The actual minimal primes are 6 ideals, which correspond to 5 component varieties over $\mathbb{Q}$. One of the prime ideals does not correspond to any variety over $\mathbb{Q}$, but would decompose into 8 point varieties over $\mathbb{C}$. We exclude this ideal from the example for the sake of brevity. The remaining 5 ideals are as follows:

1. $\langle F_{3,3}, F_{3,2}, F_{3,1}, F_{2,3}, F_{2,2}, F_{2,1}, F_{1,3}, F_{1,2}, F_{1,1} \rangle$
2. $\langle 16000F_{3,2}+31F_{3,3}, 3200F_{3,1}+3F_{3,3}, 12000F_{2,3}+11F_{3,3}, 8000000F_{2,2}-9F_{3,3}, 1200000F_{2,1}+ F_{3,3}, 4000F_{1,3} - F_{3,3}, 6000000F_{1,2} - F_{3,3}, 4800000F_{1,1} - 7F_{3,3}, f_2 + 1500, f_1 - 2000 \rangle$
3. $\langle 16000F_{3,2}+31F_{3,3}, 3200F_{3,1}+3F_{3,3}, 12000F_{2,3}+11F_{3,3}, 8000000F_{2,2}-9F_{3,3}, 1200000F_{2,1}+ F_{3,3}, 4000F_{1,3} - F_{3,3}, 6000000F_{1,2} - F_{3,3}, 4800000F_{1,1} - 7F_{3,3}, f_2 + 1500, f_1 + 2000 \rangle$
4. $\langle 16000F_{3,2}+31F_{3,3}, 3200F_{3,1}+3F_{3,3}, 12000F_{2,3}+11F_{3,3}, 8000000F_{2,2}-9F_{3,3}, 1200000F_{2,1}+ F_{3,3}, 4000F_{1,3} - F_{3,3}, 6000000F_{1,2} - F_{3,3}, 4800000F_{1,1} - 7F_{3,3}, f_2 - 1500, f_1 + 2000 \rangle$
5. $\langle 16000F_{3,2}+31F_{3,3}, 3200F_{3,1}+3F_{3,3}, 12000F_{2,3}+11F_{3,3}, 8000000F_{2,2}-9F_{3,3}, 1200000F_{2,1}+ F_{3,3}, 4000F_{1,3} - F_{3,3}, 6000000F_{1,2} - F_{3,3}, 4800000F_{1,1} - 7F_{3,3}, f_2 - 1500, f_1 - 2000 \rangle$

The first ideal corresponds to the component of dimension 2. It represents the degenerate situation when the matrix F is the zero matrix. 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.

The 12 (or 4) other components are of dimension 0 are point varieties (all unknowns are determined there). They can be divided into 3 groups (or 1 group), which correspond(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 (Eqs. 4.3, 4.2, 4.1) in the second degree, so it is impossible to determine it's sign[1].

---

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

We see that each group corresponds 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. This explains why sometimes we can get only 4 components with nonzero f_1, f_2 instead of 12.

4.5. 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 [9] can be partially 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 camera relative pose with two unknown focal lengths.

5. Improvements

Using the analysis carried out in the previous chapters, we suggest improvements to the existing methods of focal length computing.

5.1. f-Ratio

We present and analyze the performance of a new algorithm, called f-Ratio, for robust focal lengths computation. The algorithm uses the premise that the ratio of focal lengths is more robust to achieve superior accuracy. The algorithm serves as a further proof that explicitly using the ratio of the focal lengths may by beneficial for the performance.

5.1.1. 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 the automatic generator [16]. The solver uses the Demazure polynomials, i.e. the rank and the constraints 2.1, 2.2 in terms of elements of the matrix F, f_1, and r.

A solver assuming fixed focal length $f_1 = f_2$ could be used instead, for example that of Torii et el. [26], by first rescaling

$$\mathbf{x}_2^r = \begin{bmatrix} r & 0 & 0 \\ 0 & r & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{x}_2.$$

The output of such solver would be f_1 and F. We can compute f_2 as $f_1 r$.

Both solvers give same results up to machine precision. Both solvers also yield 15 (possibly non-real) solutions.

Algorithm 4: f-Ratio

Data: A list of 7 right image points $\mathbf{x}_1$ and a list of the corresponding left image points $\mathbf{x}_2$, lists of remaining tentative correspondences $\mathbf{x}_1^{test}, \mathbf{x}_2^{test}$, 6pt solver *solve6pt*

Result: Fundamental matrix F

begin

    Estimate $\mathbf{F}_0$ from all correspondences;

    Estimate $r = f_2/f_1$ from $\mathbf{F}_0$;

    foreach *6-tuple* $\mathbf{x}_1^6$ *of points drawn from* $\mathbf{x}_1$ **do**

        $\mathbf{x}_2^6 \leftarrow$ corresponding points to $\mathbf{x}_1^6$ from $\mathbf{x}_2$;

        $\mathbf{F}s_i \leftarrow solve6pt(\mathbf{x}_1^6, \mathbf{x}_2^6, r)$;

        $\mathbf{F}_i \leftarrow$ the matrix from $\mathbf{F}s_i$ which best explains $\mathbf{x}_1^{test}, \mathbf{x}_2^{test}$,

    end

    return F_i *which best explains* $\mathbf{x}_1^{test}, \mathbf{x}_2^{test}$

end

The solver needs to compute and test 7×15 fundamental matrices[1], which still may be bearable time for a RANSAC-based method.

One may construct a version of the algorithm which works with all the available points, however, in such an algorithm one would need to test a combinatorially growing number of matrices.

5.1.2. Performance

We assess the performance of the algorithm in a similar manner as in the previous analysis. Fig. 5.1 a comparison our f-Ratio method and the baseline 7pt algorithm.
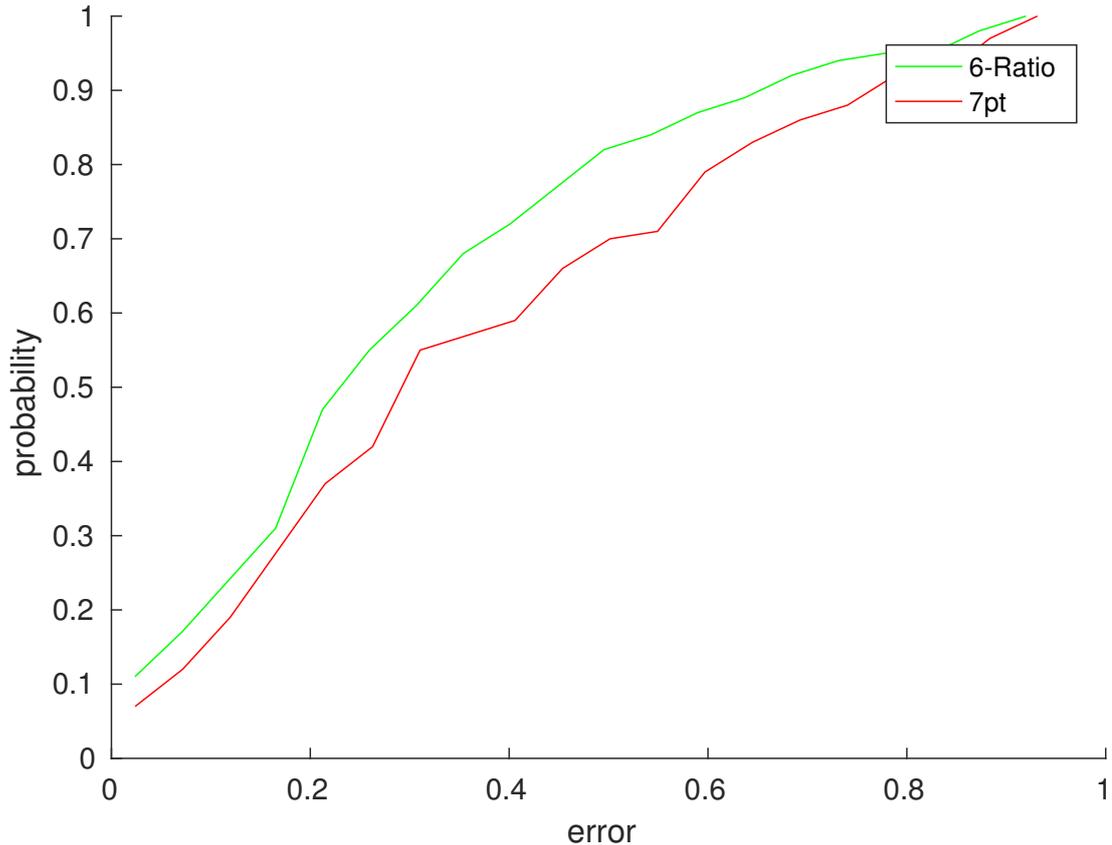
Figure 5.1. A comparison of 6-Ratio against the baseline. The cumulative distribution of the multiplicative error (equation 3.1) in the focal length estimates is shown.

In the next experiment, Fig. 5.2 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 standard RANSAC algorithm.

The results show that it is possible to reconstruct scenes better with f-Ratio than with 7pt algorithm. The algorithm 4 successfully demonstrates that explicitly using estimated ratio of focal lengths may improve the accuracy of the estimation.

---

[1]At most. The 6pt solver usually gives from 9 to 14 real solution, but the maximal possible number is 15.
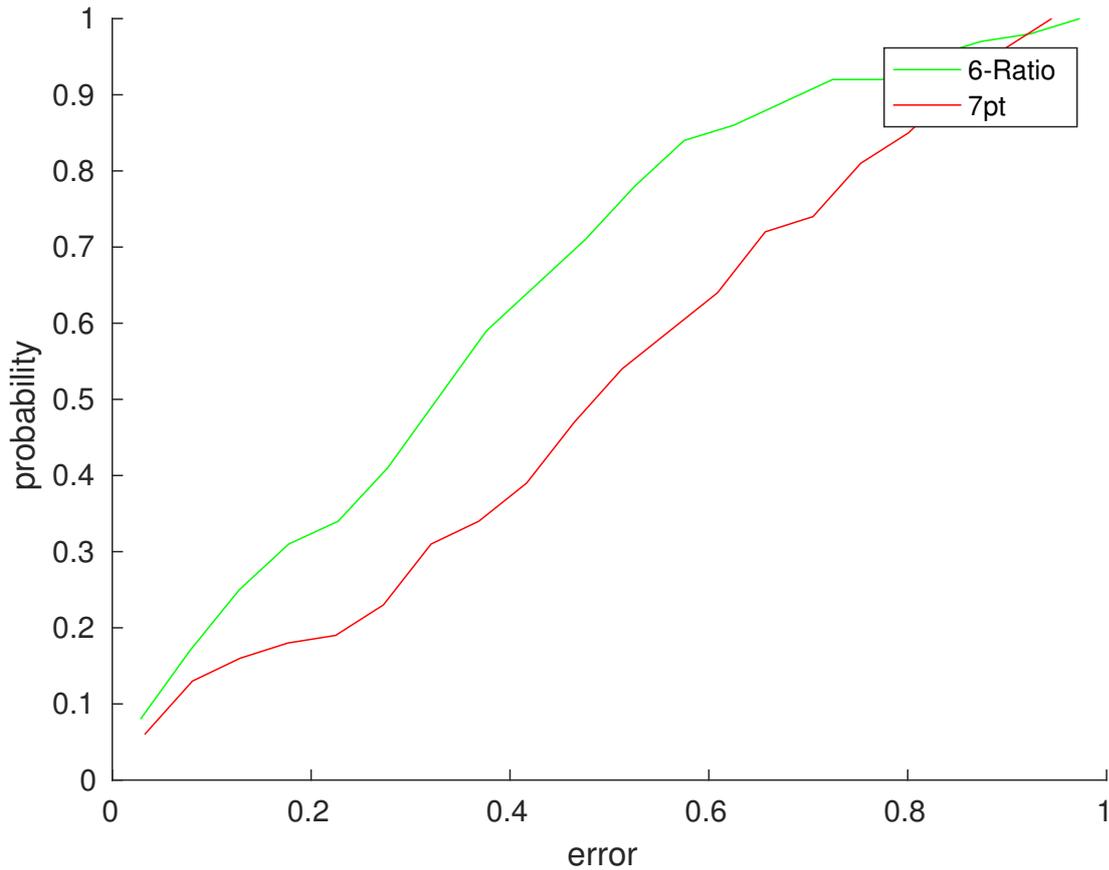
Figure 5.2. A comparison of 6-Ratio against the baseline in presence of outliers. The cumulative distribution of the multiplicative error (equation 3.1) in the focal length estimates is shown.

5.2. Prior focal length

Hartley [10] 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. In this section we describe an improvement to his algorithm. Using computed focal length ratio $r = f_2/f_1$, we are able to get better focal lengths estimates.

5.2.1. Original Hartley's algorithm

Original Hartley's algorithm optimizes a certain cost function given the weights w_i, point correspondences $\mathbf{x}_1$, $\mathbf{x}_2$, prior focal length $\bar{f}_1$, $\bar{f}_2$, 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.

Note that the focal lengths are not in the list of the parameters to 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 a term, the method should perform accurately even if the cameras used were indeed the same camera. The fourth and the fifth terms serve two purposes. Firstly, they prevent the

---

Algorithm 5: The cost function of Hartley [10]

Data: Fundamental matrix F, principal points $\mathbf{p}_1$, $\mathbf{p}_2$.
Result: Vector of costs (errors) **C**
begin

    From F, $\mathbf{p}_1$, $\mathbf{p}_2$ compute f_1, f_2;
    $C_{\mathbf{F}} \leftarrow$ Sampson error of the matrix F on the points $\mathbf{x}_1$, $\mathbf{x}_2$;
    $C_f \leftarrow$
    $w_1^2(f_1^2 - \bar{f}_1^2)^2 + w_2^2(f_2^2 - \bar{f}_2^2)^2 + w_d^2(f_1^2 - f_2^2)^2 + w_{z1}^2(f_{min}^2 - f_1^2)^2 + w_{z2}^2(f_{min}^2 - f_2^2)^2$;
    $C_{\mathbf{p}} \leftarrow w_p^2\|\mathbf{p}_1 - \bar{\mathbf{p}}_1\|^2 + w_p^2\|\mathbf{p}_2 - \bar{\mathbf{p}}_2\|$;
    return *Costs $C_{\mathbf{F}}$, $C_{\mathbf{p}}$, C_f;*

end

---

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 prevent focal lengths from converging to zero.

Hartley suggests initializing the optimization with a technique he calls calibrated reconstruction, which is described as algorithm 6. The algorithm takes point correspondences and prior information about camera calibration, and returns a matrix which is consistent with priors. Hartley also claims that the returned matrix has small Sampson error on the point correspondences.

Algorithm 6: Calibrated reconstruction [10]

Data: Point correspondences $\mathbf{x}_1$, $\mathbf{x}_2$, prior focal lengths $\bar{f}_1$, $\bar{f}_2$, prior principal
       points $\bar{\mathbf{p}}_1$, $\bar{\mathbf{p}}_2$.
Result: Fundamental matrix $\mathbf{F}'$
begin

    Compute a fundamental matrix F from point correspondences $\mathbf{x}_1$, $\mathbf{x}_2$ using the
      7pt algorithm;
    Create the prior calibration matrices $\bar{\mathsf{K}}_1$, $\bar{\mathsf{K}}_2$ from the $\bar{f}_1$, $\bar{f}_2$, $\bar{\mathbf{p}}_1$, $\bar{\mathbf{p}}_2$;
    Create an estimate of the essential matrix $\bar{\mathsf{E}} = \bar{\mathsf{K}}_2^T \mathsf{F} \bar{\mathsf{K}}_1$;
    Compute $(\mathsf{U}, \mathsf{S}, \mathsf{V}) = \mathrm{svd}(\bar{\mathsf{E}})$;
    Take the two biggest singular values s_1, s_2 Create an essential matrix
    $\mathsf{E} = \mathsf{U} \ \mathrm{diag}(\begin{bmatrix} s_1 & s_2 & 0 \end{bmatrix}) \ \mathsf{V}$;
    Create a fundamental matrix which is consistent with the priors $\mathsf{F}' = \bar{\mathsf{K}}_2^{-\mathrm{T}} \mathsf{E} \bar{\mathsf{K}}_1^{-1}$;
    return F';

end

5.2.2. Using ratio

We introduce a modification based on use of 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. We include the estimated ratio in 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 5.3 shows the experiment with 100 runs of each solver.
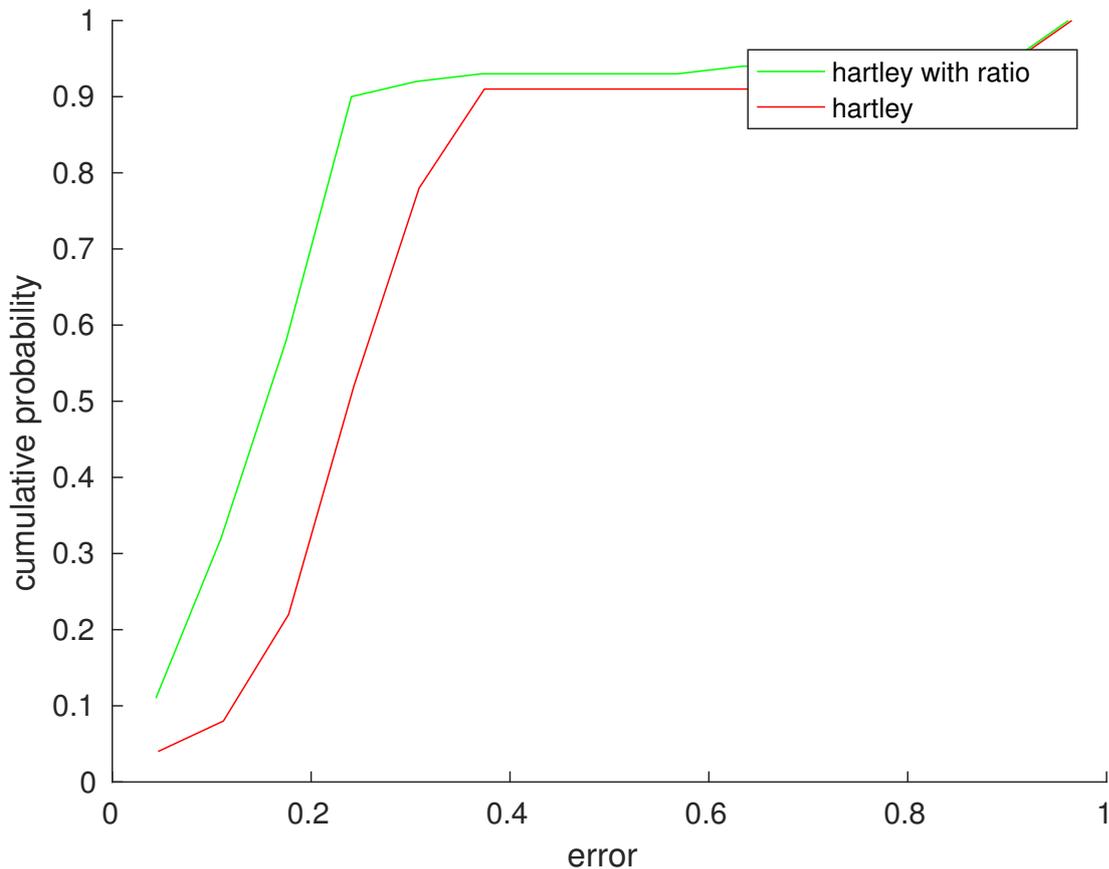
Figure 5.3. A comparsion of the Hartley algorithm [10] against our modified version which explicitly uses the ratio of the focal lengths $r = f_2/f_1$. The cumulative distribution of the multiplicative error (equation 3.1) in the focal length estimates. The ground truth focal lengths were $(3, 4)$. The 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 than weights on focal lengths priors.

5.2.3. Comparison of focal length computing methods

We summarize the performance of several methods for estimating the focal lengths from the point correspondences. Besides previously mentioned, we use a method similar to the one from [4]. The method derived in the work is described as the algorithm 7

Algorithm 7: The algorithm of Chandraker [4]

Data: Weights w_i, point correspondences $\mathtt{x}_1$, $\mathtt{x}_2$, prior focal length $\bar{f}_1$, $\bar{f}_2$, prior
    principal points $\bar{\mathbf{p}}_1$, $\bar{\mathbf{p}}_2$.

Result: Vector of costs (errors) $\mathbf{C}$

begin

　Estimate the fundamental matrix $\mathtt{F}$ using the 7pt algorithm;
　Define $p_1(f_1, f_2, \mathbf{p}_1, \mathbf{p}_2), p_2(f_1, f_2, \mathbf{p}_1, \mathbf{p}_2), p_3(f_1, f_2, \mathbf{p}_1, \mathbf{p}_2)$ as Kruppa
　 equations [9], rewritten as polynomials (see the paper [4]);
　Minimize the following cost function:

$$f_1^*, f_2^*, \mathbf{p}_1^*, \mathbf{p}_2^* =$$

$$= \min_{f_1, f_2, \mathbf{p}_1, \mathbf{p}_2, \lambda_1, \lambda_2, \lambda_3} \quad w_1^2(f_1 - \bar{f}_1)^2 + w_2^2(f_2 - \bar{f}_2)^2 + w_p^2\|\mathbf{p}_1 - \bar{\mathbf{p}}_1\|^2 + w_p^2\|\mathbf{p}_2 - \bar{\mathbf{p}}_2\|$$

$$\text{such that} \quad \lambda_1 p_1(f_1, f_2, \mathbf{p}_1, \mathbf{p}_2) + \lambda_2 p_2(f_1, f_2, \mathbf{p}_1, \mathbf{p}_2) + \lambda_3 p_3(f_1, f_2, \mathbf{p}_1, \mathbf{p}_2)$$

　return *focal lengths f_1, f_2, principal points $\mathbf{p}_1, \mathbf{p}_2$;*

end

Note that the algorithm 7, contrary to the algorithm 5 of Hartley, does not change the fundamental matrix during the optimization. Instead, it computes the fundamental matrix beforehand, and the optimization finds a suitable combination of principal points and focal lengths that would be consistent with the fundamental matrix..

We use an optimization procedure obtained in private conversation to do the optimization in the last step of the algorithm 7.

The figures 5.4, 5.5 show the distribution of focal length as computed by different methods against different levels of noise in prior focal lengths. The distribution for each case is depicted using MATLAB function boxplot which shows 25% to 75% quantile values as boxes with a horizontal line for median. The crosses show data beyond 1.5 times the interquartile range. We use high level of noise ($\sigma = 4$ pixels) to show that the regularization by priors allows us to handle bigger amount of noise. For smaller amounts noise the difference between the Bougnoux formula and prior methods is not as pronounced.

We can observe in the figures 5.4, 5.5 that different methods behave inherently in a different way. The bougnoux formula has the biggest standard deviation of all, generally is noisy and gives a lot of outliers. The formula is also much less biased, and, of course, is independent of the given prior. The method of Hartley, because of the chosen cost function, tends to drive the focal lengths close to each other. Thus, it overestimates the smaller focal length in the Fig. 5.4 and underestimates the bigger focal length in the Fig. 5.4. Our modified version of Hartleys' algorithm, as well as the method of Chandraker does not suffer from this. The algorithm of Chandraker apparently shows smaller standard deviation and is less influenced by the prior.
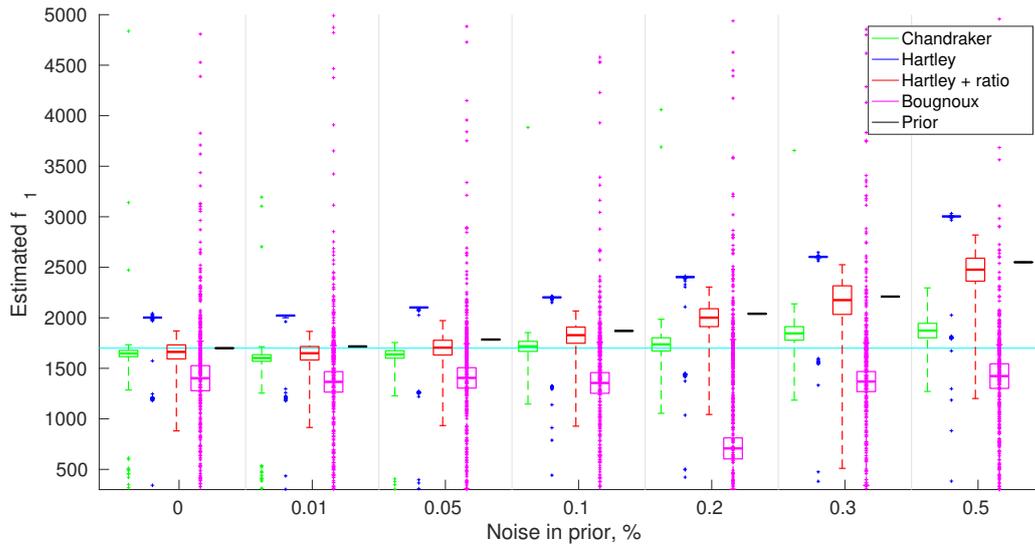
Figure 5.4. A comparison of the methods for computing focal lengths from point correspondences. The ground truth focal length f_1 was 1700 (shown in cyan), and the relative noise (axis x) was added to it to produce priors, i. e., the priors were: 1700, 1717, 1785, 1870, 2040, 2210, 2550. The ground truth principal points were both $(10,20)$, and the prior principal points were $(0,0)$. The number of correspondences used was 7, and the level of noise σ was equal to 1. The weights on principal point priors were 10 times smaller than weights on focal lengths priors. The term enforcing the ratio of focal length in our modified version has the same weight as the terms for the focal lengths themselves.
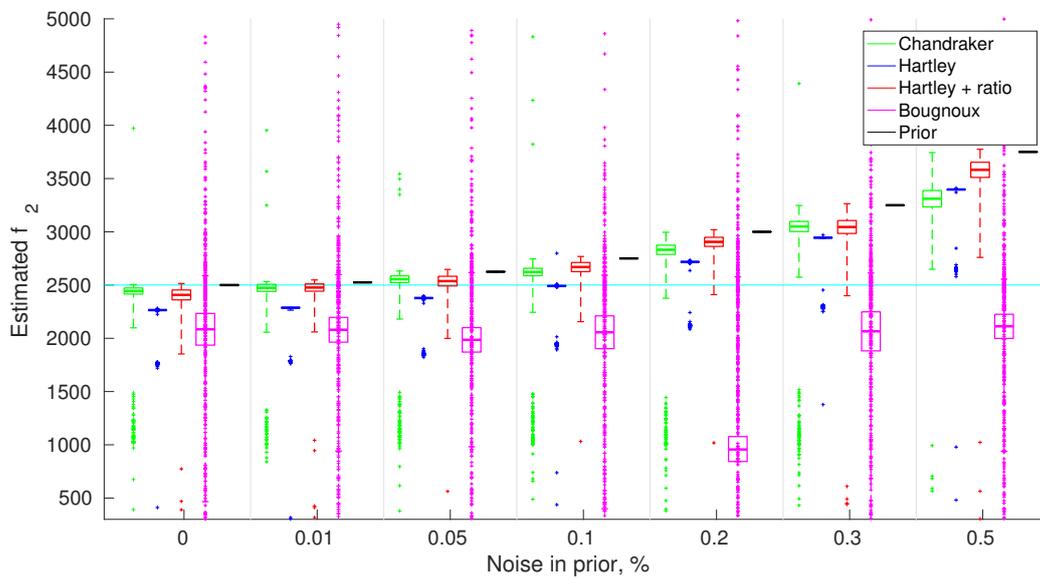
Figure 5.5. A comparison of the methods for computing focal lengths from point correspondences. The ground truth focal length f_2 was 2500 (shown in cyan), and the relative noise (axis x) was added to it to produce priors, i. e., the priors were: 2500, 2525, 2625, 2750, 3000, 3250, 3750.

5.3. Conclusions

We conclude that the fact that focal length ratio $r = f_2/f_1$ is robust can be used to construct more efficient algorithms. We show two examples of modifications to existing algorithms where we explicitly use a ratio estimation to improve the accuracy of the focal lengths computation. We show that using prior knowledge about focal length considerably bigger amounts of noise could be treated.

6. Conclusions

In this work, we have focused on the methods of computing the focal length from point correspondences.

In Chapter 2 we surveyed the existing methods for this task, as well as the basic concepts needed for understanding them. We provided specific degeneracies of the methods and other cases where they might fail.

In Chapter 3 we analyzed the performance of the different methods for computing the focal length. We have revealed several trends which these algorithms exhibit. Firstly, the error of the focal length estimation declines rapidly with growing number of correspondences. Secondly, with growing number of correspondences, the number of imaginary focal length estimates also declines. A conclusion can be made that the standard methods work fairly well for scenes where a sufficient number of inlier correspondences may be found.

We compared the performance of minimal solvers that use or don't use the rank constraint (Eq. 2.1). Our results found that using rank constraint isneficial for the performance, however, such a solver might produce a larger fraction of imaginary estimates.

We found that the computation of the ratio of focal length $r = f_2/f_1$ by the Bougnoux formula [3] is more robust than the computation of f_1 or f_2 alone. Interestingly, this robustness is preserved even when both focal length estimates are imaginary.

We furthermore assessed the performance of the methods in degenerate situations. The results showed that for bigger levels of noise in image measurements the effect of the degeneracies significantly decreases. Specifically, the effects of the intersecting optical axes degeneracy was shown to be mild under 1 pixel noise in image measurements already.

In Chapter 4 we analyzed the problem of computing focal length using the techniques of algebraic geometry. We have shown how the Bougnoux formula may be derived with these techniques and give two new formulae for computing camera focal length from a fundamental matrix, as well as one formula for the ratio of the focal length. We have shown that using the right formula helps avoiding a known degeneracy. Specifically, the degeneracy where 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, and where Bougnoux ([3]) formula fails can in some cases be avoided. The degeneracy reduces to the case where all three formulae fail.

In Chapter 5 we suggested improvements to the existing methods using our suggestion that computation of the ratio of focal length $r = f_2/f_1$ is more robust than computation of f_1 or f_2 alone. We have shown that this fact indeed may improve performance of a solver. We suggested a modification to the solver of [10] and demonstrated an improvement in focal length estimation quality over the original method. Finally, we presented a survey of optimization methods for focal lengths computing and have shown that the optimization methods using prior focal length information may exhibit robust performance even under big amounts of noise in image measurements.

A. Contents of the enclosed CD

```
/
└── RobustFocal ................ folder with the all files of the thesis
    ├── results ................ folder containing graphs that are included in the
    │   │                        thesis
    │   └── ⋮
    ├── lib ..................... folder with auxiliary files required for the exper-
    │   │                        iments
    │   ├── generated_solvers .... solvers generated for the work using automatic
    │   │   │                      generator [16]
    │   │   └── ⋮
    │   ├── utils ............... utility scripts
    │   │   └── ⋮
    │   ├── geometry ............ geometrical utility scripts
    │   │   └── ⋮
    │   └── scene_generator ...... scripts that generate a sample scene with points
    │       │                      and cameras
    │       └── ⋮
    ├── data .................... folder with synthetic data generated for the ex-
    │   │                        periments
    │   └── ⋮
    └── src ..................... folder with the source files of the thesis
        ├── analysis_experiments . folder containing experiments for Chapter 4.
        │   └── ⋮
        ├── algeom_experiments ... folder containing experiments for Chapter 4.
        │   └── ⋮
        ├── thenew_experiments ... folder containing experiments for Chapter 5.
        │   └── ⋮
        ├── geometry ............. folder containing geometrical functions written
        │   │                      for the project.
        │   └── ⋮
        ├── utils ................ folder containing utility functions written for the
        │   │                      project.
        │   └── ⋮
        ├── paths.m .............. script which add all required paths to the MAT-
        │                          LAB environment
        └── toy.m ................ script which runs a toy example of the focal
                                   length computation
```

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] Manmohan Krishna Chandraker. Two-view focal length estimation for all camera motions using priors. Technical report, 2004.

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

[6] 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.

[7] 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.

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

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

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

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

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

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

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

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

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

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

[23] Rahul Raguram, Ondrej Chum, Marc Pollefeys, Jiri Matas, and Jan-Michael Frahm. USAC: A universal framework for random sample consensus. *IEEE Trans. Pattern Anal. Mach. Intell.*, 35(8):2022–2038, 2013.

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

[25] 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.

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

[27] 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.