Two-View Geometry Estimation by Random Sample and Consensus

A Dissertation Presented to the Faculty of the Electrical Engineering of the Czech Technical University in Prague in Partial Fulfillment of the Requirements for the Ph.D. Degree in Study Programme No. P2612 - Electrotechnics and Informatics, branch No. 3902V035 - Artificial Intelligence and Biocybernetics, by

Ondřej Chum

July 1, 2005

Thesis Advisor

Dr. Ing. Jiří Matas

Center for Machine Perception
Department of Cybernetics
Faculty of Electrical Engineering
Czech Technical University in Prague
Karlovo náměstí 13, 121 35 Prague 2, Czech Republic
fax: +420 224 357 385, phone: +420 224 357 465
http://cmp.felk.cvut.cz

Abstract

The problem of model parameters estimation from data with a presence of outlier measurements often arises in computer vision and methods of robust estimation have to be used. The RANSAC algorithm introduced by Fishler and Bolles in 1981 is the a widely used robust estimator in the field of computer vision. The algorithm is capable of providing good estimates from data contaminated by large (even significantly more than 50%) fraction of outliers. RANSAC is an optimization method that uses a data-driven random sampling of the parameter space to find the extremum of the cost function. Samples of data define points of the parameter space in which the cost function is evaluated and model parameters with the best score are output.

This thesis provides a detailed analysis of RANSAC, which is recast as time-constrained optimization – a solution that is optimal with certain confidence is sought in the shortest possible time.

Next, the concept of randomized cost function evaluation in RANSAC is introduced and its superiority over the deterministic evaluation is shown. A provably optimal strategy for the randomized cost function evaluation is derived.

A known discrepancy, caused by noise on inliers, between theoretical prediction of the time required to find the solution and practically observed running times is traced to a tacit assumptions of RANSAC. The proposed LO-RANSAC algorithm reaches almost perfect agreement with theoretical predictions without any negative impact on the time complexity.

A unified method of estimation of model and its degenerate configuration (epipolar geometry and homography of a dominant plane) at the same time without a priori knowledge of the presence of the degenerate configuration (dominant plane) is derived.

Next, it is shown that using oriented geometric constraints that arise from a realistic model of physical camera devices, saves non-negligible fraction of computational time. No negative side effect are related to the application of the oriented constraints.

An algorithm exploiting (possibly noisy) match quality to modify the sampling strategy is introduced. The quality of a match is an often freely available quantity in the matching problem. The approach increases the efficiency of the algorithm while keeping the same robustness as RANSAC in the worst-case situation (when the match quality is unrelated to whether a correspondence is a mismatch or not).

Most of the algorithms in the thesis are motivated by (and presented on) estimation of a multiview geometry. The algorithms are, however, general robust estimation techniques and can be easily used in other application areas too.

Acknowledgments

I would like to express my thanks to my colleagues at CMP who I had the pleasure to work with, especially Štěpán Obdržálek, and Michal Perd'och. My thanks belongs to Dr. Tomáš Werner, who also read the manuscript of the thesis, and Dr. Tomáš Pajdla for passing some of their knowledge of geometry on me. I am greatly indebted to my advisor Dr. Jiří Matas for guiding me throughout my research. His support, considerable help, and patience have been of outmost importance for finishing my PhD study. I thank to my family and to my friends for their support that made it possible for me to finish this thesis.

I gratefully acknowledge Toyota Motor Corporation and the grant of the Grant Agency of the Czech Technical University CTU 0306013 which supported my research.

Contents

Keywords: RANSAC, robust estimation, epipolar geometry, homography

1 Introduction

In the last few years, the field of computer vision has made significant advances in practical application of its theory. Many applications have emerged in numerous fields, including 3D reconstruction (Fig. 1.1) in architecture and telecommunications, augmented reality in the film industry, object recognition (Fig. 1.2), tracking in traffic surveillance, registration of medical images and others. Commonly the applications require (implicitly or explicitly) solving the so-called correspondence problem: finding correspondences in two or more images of a 3D scene taken from arbitrary viewpoints viewed with possibly different cameras and in different illumination conditions.

Obtaining correspondence in sets of images includes many steps: detection of regions (points) of interest [18, 1, 60, 38, 32, 35], tentative matching of detected regions [32, 60, 35] based on local (affine invariant) descriptors [17, 32, 49]. It is generally accepted that incorrect matches – *outliers* – cannot be avoided in the stage of the matching process where only local image descriptors are compared. The mismatches, due to phenomena like occlusions, depth discontinuities and repetitive patterns, are detected and removed by robust methods. In a final step of the matching process, the robust estimators search for sets of matches consistent with some global geometric constraint that arises from the fact that the same 3D scene was observed in all images. Robust estimation is the topic of this thesis.

The problem of parameter estimation from data contaminated by outliers is known as robust estimation in the field of robust statistics [24, 46]. The RANSAC[1] algorithm introduced by Fishler and Bolles in 1981 [15] is the most widely used robust estimator in the field of computer vision. RANSAC and similar techniques [57, 39, 3] have been applied in the context of short baseline stereo [54, 58], wide baseline stereo matching [45, 60, 47], motion segmentation [54], mosaicing [37], detection of geometric primitives [11], robust eigenimage matching [31] and elsewhere.

The algorithm is capable of providing good estimates from data contaminated by large (even significantly more than 50%) fraction of outliers. RANSAC is an optimization method that uses a data-driven random sampling of the parameter space to find the extremum of the cost function. Samples of data define points of the parameter space in which the cost function is evaluated. Finally, model parameters with the best score are output.

1.1. Goals of the Thesis

The goal of the thesis is to provide a comprehensive description of the RANSAC algorithm together with detail analysis of the behavior and drawbacks of the method; design novel techniques

[1] **RAN**dom **SA**mple **C**onsensus

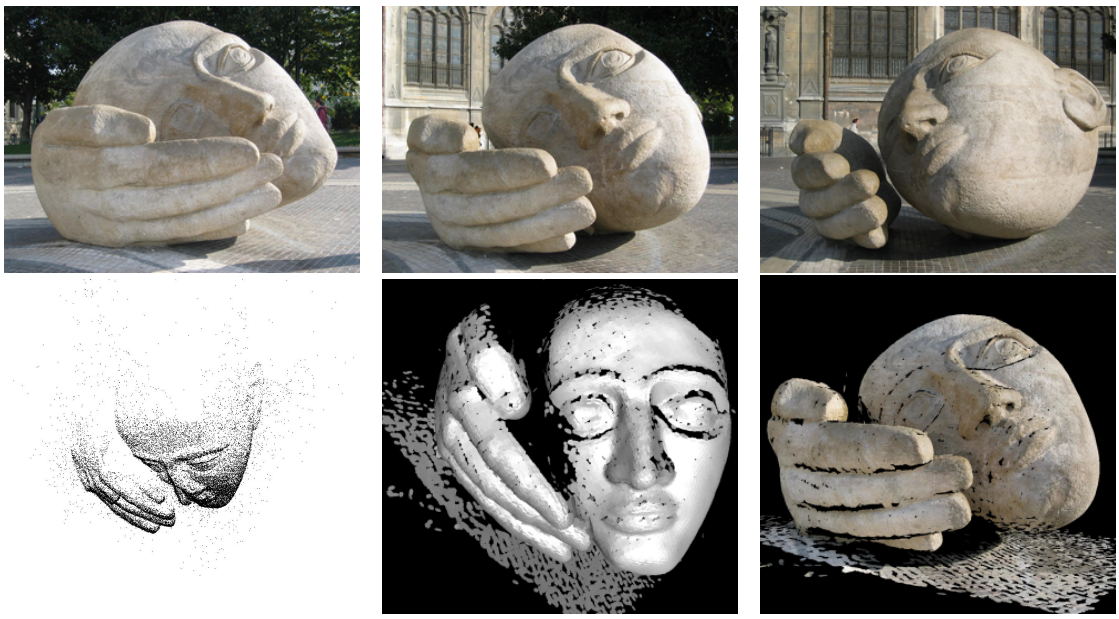


Figure 1.1.: Automatic 3D reconstruction from images. Three images from the image sequence (top row), reconstructed points (bottom left), reconstructed surface (bottom middle), and textured 3D model (bottom right). The images were taken from [12].

enhancing the algorithm and eliminating or reducing the drawbacks. In particular, the following weaknesses of RANSAC are addressed:

- sensitivity to the noise on inliers;
- sensitivity to the presence of degenerate configurations;
- issues of the speed of the algorithm, addressing both the hypothesis generation and model verification.

1.2. Contributions of the Thesis

This thesis contributes to the state of the art in robust estimation of geometric models. First, it provides a detailed analysis of RANSAC, which is formulated as time-constrained optimization. A solution that is optimal with certain confidence is sought in the shortest possible time.

As a second contribution, the concept of randomized cost function evaluation in RANSAC is introduced and its superiority over the deterministic evaluation is shown. A provably optimal strategy for the randomized cost function evaluation is derived.

Third, a known discrepancy (caused by noise on inliers) between theoretical prediction of the time required to find the solution and practically observed running times is traced to a tacit

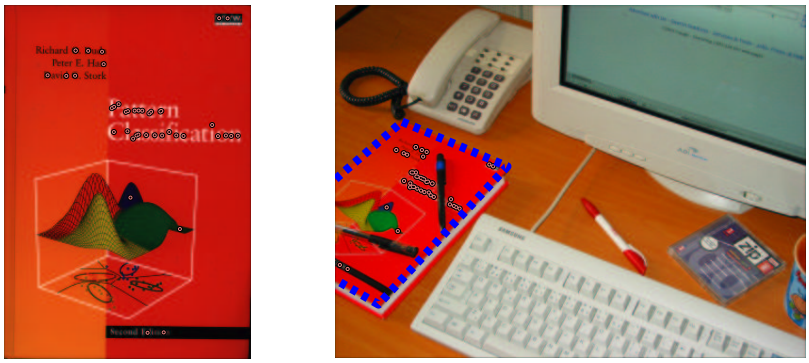


Figure 1.2.: Object recognition: Where is the book? A scanned cover of a book (left, 712×1024) and an image acquired by an uncalibrated camera (right, 1600×1200).

assumptions of RANSAC. The algorithm is modified to reach almost perfect agreement with theoretical predictions without any negative impact on the time complexity.

Fourth, an unified approach estimating model and its degenerate configuration (epipolar geometry and homography of a dominant plane) at the same time without a priori knowledge of the presence of the degenerate configuration (dominant plane) is derived.

As another contribution, it is shown that using oriented geometric constraints that arise from a realistic model of physical camera devices, saves non-negligible fraction of computational time. No negative side effect are related to the application of the oriented constraints.

Finally, an algorithm exploiting (possibly noisy) match quality to modify the sampling strategy is introduced. The quality of a match is an often freely available quantity in the matching problem. The approach increases the efficiency of the algorithm while keeping the same robustness as RANSAC in the worst-case situation (when the match quality is unrelated to whether a correspondence is a mismatch or not).

Most of the algorithms in the thesis are motivated by (and presented on) estimation of a multiview geometry. The algorithms are, however, general robust estimation techniques and can be easily used in other application areas as well (with the exception of the oriented geometric constraints).

1.3. Structure of the Thesis

RANSAC A detailed description of the original RANSAC algorithm is presented in chapter 3. Conditions under which RANSAC works worse than expected or even fails to find the solution are explained. The literature on RANSAC and robust estimation techniques based on or similar to RANSAC is reviewed in chapter 4. Each method is analyzed and compared with the standard RANSAC algorithm.

LO-RANSAC An extension of the RANSAC procedure is proposed in chapter 5. By applying local optimization (LO) to models with a score (quality) better than all previous ones, an al-

gorithm with the following desirable properties is obtained: (i) a near perfect agreement with theoretical (i.e. optimal) performance and (ii) lower sensitivity to noise and poor conditioning. The LO is shown to be executed so rarely that it has minimal impact on the execution time.

The concept of LO-RANSAC is further generalized in section 5.3, where the desired model is estimated via an approximating model with a simpler model (in terms of degrees of freedom). Such an approach significantly decreases the computational complexity of the algorithm. The potential of the generalized algorithm is demonstrated on two new matching algorithms that are straightforward application of LO-RANSAC: (i) an algorithm for simultaneous estimation of epipolar geometry and radial distortion and (ii) an algorithm estimating epipolar geometry from three region-to-region correspondences.

DEGENSAC In chapter 6, a RANSAC-based algorithm for robust estimation of epipolar geometry from point correspondences in the possible presence of a dominant scene plane is presented. The algorithm handles scenes with (i) all points in a single plane, (ii) majority of points in a single plane and the rest off the plane, (iii) no dominant plane. It is not required to know a priori which of the cases (i) – (iii) occurs.

The algorithm exploits a theorem that if five or more of seven correspondences are related by a homography then there is an epipolar geometry consistent with the seven-tuple as well as with all correspondences related by the homography. This means that a seven point sample consisting of two outliers and five inliers lying in a dominant plane produces an epipolar geometry which is wrong and yet consistent with a high number of correspondences. The theorem explains why RANSAC often fails to estimate epipolar geometry in the presence of a dominant plane.

Rather surprisingly, the theorem also implies that RANSAC-based homography estimation is faster when drawing non-minimal samples of seven correspondences than minimal samples of four correspondences.

PROSAC In chapter 7, the Progressive Sampling Consensus (PROSAC) is introduced. The PROSAC algorithm exploits the linear ordering defined on the set of correspondences by a similarity function used in establishing tentative correspondences. Unlike RANSAC, which treats all correspondences equally and draws random samples from the full set, PROSAC samples are drawn from progressively large sets of top-ranked correspondences. Under the mild assumption that the similarity measure predicts correctness of a match better than random guessing, we show that PROSAC achieves large computational savings.

R-RANSAC ($T_{d,d}$, SPRT) In chapter 8, a randomized cost function evaluation strategy for RANSAC is presented. The method finds, like RANSAC, a solution that is optimal with user-specified probability. Computational savings are achieved by evaluating a statistical test on only a fraction of data points. Two tests are discussed in the chapter. A non-optimal, but simple and intuitive, $T_{d,d}$ test is introduced first. A provably optimal model verification strategy is designed for the situation when the contamination of data by outliers is known, i.e. the algorithm is the fastest possible (on average) of all randomized RANSAC algorithms guaranteeing confidence in the solution. The derivation of the optimality property is based on Wald's theory of sequential

decision making. Finally, the R-RANSAC with Sequential Probability Ratio Test (SPRT), which does not require the a priori knowledge of the fraction of outliers and has results close to the optimal strategy, is introduced.

Oriented constraints The efficiency of epipolar geometry estimation by RANSAC is improved by exploiting the oriented epipolar constraint in chapter 9. Performance evaluation shows that the enhancement brings up to a two-fold speed-up. The orientation test is simple to implement, is universally applicable and takes negligible fraction of time compared with epipolar geometry computation.

1.4. How to Read the Thesis

The RANSAC algorithm is described and analyzed in chapter 3 in detail. Further chapters build on this chapter and often refer to it. Starting from chapter 5, each chapter addresses one robust estimation problem. The problems are demonstated and the proposed algorithms are tested on two-view geometry estimation. Basic knowledge on the level of [22] of the two-view geometry is expected. These chapters can be read on their own.

1.5. Authorship

I hereby certify that the results presented in this thesis were achieved during my own research in cooperation with my thesis advisor Jiří Matas, published in [5, 34, 6, 36, 7, 8, 9, 3], with Štěpán Obdržálek, published in [36, 7], with Tomáš Werner, published in [8, 9], and Josef Kittler, published in [3].

2 Notation and Concepts

The mathematical notation used in the thesis is summarized in the following table.

$a, b, \lambda, \ldots$	scalars
$\mathbf{x}, \mathbf{e}, \mathbf{l}, \mathbf{C}, \ldots$	vectors
$\mathtt{F}, \mathtt{H}, \ldots$	matrices
$\mathbb{R}$, $\mathbb{R}^+$, $\mathbb{R}_0^+$	real, positive real, and non-negative real numbers respectively
$\mathbb{R}^n$, $\mathbb{P}^n$	n-dimensional spaces
$\mathcal{U}, \mathcal{S}, \ldots$	sets
$\lvert\mathcal{S}\rvert$	cardinality of the set
$\lVert\mathbf{a}\rVert$	vector norm
$\lVert\mathtt{A}\rVert$	Frobenius norm
$\mathtt{A}^+$	matrix pseudoinverse
$\mathbf{a} \sim \mathbf{b}$, $\mathtt{A} \sim \mathtt{B}$	vectors (matrices) are equal up to a non-zero scale factor: $\exists\lambda \in \mathbb{R} \setminus \{0\} : \lambda\mathbf{a} = \mathbf{b}$
$\mathbf{a} \not\sim \mathbf{b}$, $\mathtt{A} \not\sim \mathtt{B}$	for all $\lambda \in \mathbb{R} : \lambda\mathbf{a} \neq \mathbf{b}$
$\mathbf{a} \overset{+}{\sim} \mathbf{b}$, $\mathtt{A} \overset{+}{\sim} \mathtt{B}$	vectors (matrices) are equal up to a positive scale factor: $\exists\lambda \in \mathbb{R}^+ : \lambda\mathbf{a} = \mathbf{b}$
$[\mathbf{e}]_\times : \mathbf{e} \in \mathbb{R}^3$	a matrix satisfying $[\mathbf{e}]_\times\mathbf{x} = \mathbf{e} \times \mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^3$

Table 2.1.: Used fonts and symbols.

Let $\mathbb{X}$ be a universum of observations (data points) and $\mathcal{U}$ be a finite set of data points $\mathbf{x}_i \in \mathbb{X}$, $i = 1 \ldots N$. Let Θ_M be a space of parameters of some model M.

In the description of RANSAC, a notion of a data point being consistent with model with parameters θ is required. We assume that an error function

$$\rho_M(\theta, \mathbf{x}) : \Theta_M \times \mathbb{X} \to \mathbb{R}_0^+$$

is given together with a threshold $\Delta_M \in \mathbb{R}_0^+$. A data point $\mathbf{x}$ is consistent with (supports) a model with parameters θ iff

$$\rho_M(\theta, \mathbf{x}) \leq \Delta_M.$$

We drop the subscript M in the following definitions since only one model is thought at a time. Let the support of the model with parameters θ be denoted by function

$$\mathcal{S}(\theta) : \Theta \to \exp(\mathcal{U}),$$

so that

$$S(\theta, \mathcal{U}, \Delta) = \{\mathbf{x} \in \mathcal{U} |\ \rho(\theta, \mathbf{x}) \leq \Delta\}.$$

We further assume that a function

$$f(\mathcal{M}) : \mathbb{X}^m \rightarrow \Theta^{\mathcal{O}(1)}$$

giving a finite set $\mathcal{T}$ of model parameters that fit a set $\mathcal{M}$ of m data points from $\mathbb{X}$ is given. The set $\mathcal{M}$ is called a sample. The size of the sample m must be sufficient to generate only a finite set of the model parameters 'fit' the sample. From computational reasons, discussed later, m is typically chosen to be the smallest possible.

Example. Consider line fitting to points in 2D. In this task, the universum of data points would be $\mathbb{X} = \mathbb{R}^2$. There are many possibilities, how to parameterize a line in 2D, for example $\Theta = \{(\phi, d), \phi \in \langle 0, \pi), d \in \mathbb{R}^+\}$, where ϕ is the angle between the line and the abscissa and d is the perpendicular distance between the line and the origin. The error function ρ would be the distance between a line and a point. The constant $m = 2$ as there is an infinite number of lines passing through a single point, and two distinct points define a line uniquely.

For epipolar geometry, the data points are correspondences in two images, $\mathbb{X} = \mathbb{P}^2 \times \mathbb{P}^2$. The parameter space Θ are fundamental matrices $\mathtt{F} \in \mathbb{R}^{3\times3}$ so that rank $\mathtt{F} = 2$. There are several choices of the error function, for example the algebraic error

$$\rho(\mathtt{F}, \mathbf{x} \leftrightarrow \mathbf{x}') = \left(\frac{1}{\|\mathtt{F}\|}\mathbf{x}'^{\top}\mathtt{F}\mathbf{x}\right)^2.$$

Other options are Sampson's or reprojection error [22, 21]. Seven point correspondences in general position define up to three fundamental matrices [22], therefore $m = 7$.

Mostly, a model of two-view geometry (epipolar geometry or homography) is considered in the thesis. The observations (data points) are called *correspondences* and are often denoted as $\mathbf{x} \leftrightarrow \mathbf{x}'$, where $\mathbf{x} \in \mathbb{P}^2$ is measured in one image and $\mathbf{x}' \in \mathbb{P}^2$ is measured in the other. Correspondences obtained by matching of local features are called *tentative* correspondences. Tentative correspondences are either matches or mismatches. Data points consistent with the optimal model parameters θ^* are called inliers and the rest of the data are called outliers. Typically, RANSAC is used to reveal correct matches, which implies assumption, that the set of correct matches is identical with the set of inliers and that mismatches do not conspire to mimic a competitive structure. The assumption is formalized and discussed in detail in section 3.1.

Symbol	Meaning	Epipolar geometry
$\mathbb{X}$	universum of data points	$\mathbb{X} = \mathbb{P}^2 \times \mathbb{P}^2$
$\mathcal{U}$	finite input data set $\mathcal{U} \subset \mathbb{X}$	tentative correspondences
Θ	parameter space	fundamental matrices $\mathtt{F}$, $\text{rank}(\mathtt{F}) = 2$, $\|\|\mathtt{F}\|\| = 1$
$\rho(\theta, \mathbf{x})$	error function $\Theta \times \mathbb{X} \rightarrow \mathbb{R}^+$	Sampson's error
$f(\mathcal{M})$	function that computes a set of parameters that exactly fit a set of data points $\exp(\mathbb{X}) \rightarrow \exp(\Theta)$	seven-point algorithm
m	minimal number of points that defines model parameters 'uniquely'	$m = 7$
Δ	error bound for inliers	user-defined threshold
$\mathcal{S}(\theta)$	a support of model parameters, i.e. data points with error under Δ	correspondences that are consistent with given $\mathtt{F}$
$J_{\mathcal{S}}(\theta)$	standard cost function $\Theta \rightarrow \mathbb{N}^+$, $J_{\mathcal{S}}(\theta) = \|\mathcal{S}(\theta)\|$	the number of correspondences consistent with given $\mathtt{F}$

Table 2.2.: Symbols used in the description of RANSAC and RANSAC-like algorithms and instances of the symbols for the case of epipolar geometry estimation.

3 RANSAC

The RANSAC algorithm has become one of the most popular robust estimators in computer vision community [22]. In this section, the basic algorithm introduced in 1981 by Fischler and Bolles [15] is reviewed.

The input of the RANSAC algorithm is a set $\mathcal{U}$ of measurements, here called *data points*. An unknown proportion of the data points is consistent with a model with unknown parameters from a parameter space Θ. Such data points are called inliers. The rest of the data points are erroneous (outliers). The goal is to find model parameters θ^* from a parameter space Θ that maximize a cost function $J_S(\theta, \mathcal{U}, \Delta)$. In the standard formulation, the cost function J_S is the size of the support of the model with parameters θ, i.e. how many data points from $\mathcal{U}$ are consistent with it. Data points with the error smaller than Δ are considered to be consistent or to support the model. An error function $\rho(\theta, \mathbf{x})$ representing a distance of a data point to a model is typically given. The threshold Δ is an *input parameter* to RANSAC. Notation is summarized in Tab. 2.2, and the precise definitions of the symbols can be found in chapter 2.

The RANSAC algorithm carries out the maximization of J_S as follows. Two steps are executed repeatedly: (i) a hypothesis generation step and (ii) the verification step. (i) In the *hypothesis generation* step, a hypothesis θ_k of the model parameters is computed from a subset $\mathcal{S}_k$ (so called sample) selected from the input data points $\mathcal{U}$ at random. The probability that a sample contains at least one outlier increases exponentially with the size of the sample – see equation (3.4). Hence, the size of the sample $m = |\mathcal{S}|$ is chosen to be as small as feasible to determine the model parameters *uniquely*. For example, a line is defined by two distinct points ($m = 2$) and a fundamental matrix[1] are defined by seven point-to-point correspondences in a general position ($m = 7$). (ii) In the *verification* step, the quality of the hypothesized model parameters is calculated. The cost function J_S is the cardinality of the support (consensus set) of the model with parameters θ_k. Under assumptions discussed in section 3.1, the maximum of the cost function J_S is recovered with a user predefined probability (confidence) $1 - \eta_0$ (typically set to 95%), where η_0 is the probability of returning a bad solution. The structure of the RANSAC algorithm is summarized in Alg. 1.

On the number of hypothesis tested. There are two types of samples: *contaminated*, those that contain at least one outlier, and *uncontaminated* (all-inlier or outlier-free) samples. Only the latter ones are of interest, as the model parameters computed from data points including outliers are arbitrary. Let P be the probability that an uncontaminated sample of size m is randomly

[1] In fact one to three fundamental matrices.

Input: $\mathcal{U}$, Δ, η_0
Output: θ^*, I^*

Repeat until the probability η (eq. (3.2)) of finding model with support larger than I^*_{k-1} in k-th step falls under threshold η_0:

1. Hypothesis generation
- a Select a random sample of minimum size m from $\mathcal{U}$.
- b Estimate model parameters θ_k fitting the sample.

2. Verification
- a Calculate the support $I_k = J(\theta_k)$ of the model.
- b Set $I^*_k = \max(I^*_{k-1}, I_k)$, and if $I^*_k > I^*_{k-1}$ (i.e. when a new maximum is reached), then store the current model parameters θ_k.

Algorithm 1: The structure of RANSAC (Fischler and Bolles [15]).

selected from a set $\mathcal{U}$ of N data points

$$P(I) = \frac{\binom{I}{m}}{\binom{N}{m}} = \prod_{j=0}^{m-1} \frac{I-j}{N-j} \leq \varepsilon^m, \tag{3.1}$$

where ε is the fraction of inliers $\varepsilon = I/N$.

The number of inliers I is not known beforehand. Let $\mathcal{I}^*_k$ be the largest support of a hypothesized model found up to k-th sample inclusively, $I^*_k = |\mathcal{I}^*_k|$. The sampling process is terminated [15, 58] when the likelihood of finding a better model (with larger support than I^*_k) falls under a threshold, i.e. when the probability η of missing a set of inliers $\mathcal{I}^+$ of size $|\mathcal{I}^+| \geq I^*_k$ within k samples falls under a predefined threshold η_0,

$$\eta = (1 - P(I^*_k))^k. \tag{3.2}$$

The number of samples that has to be drawn to satisfy $\eta \leq \eta_0$ is

$$k_{\eta_0}(I^*_k) = \frac{\ln(\eta_0)}{\ln(1 - P(I^*_k))}. \tag{3.3}$$

The approximate numbers of samples for various sample sizes m and different fractions ε of inliers are shown in Tab. 3.1.

Note: (On the sampling process) The uniform sampling of m correspondences can be also seen as a non-uniform sampling of the parameter space Θ. There is $\binom{N}{m}$ possible samples out of which $\binom{I}{m}$ are located close to the optimal model parameters θ^*, whereas the rest of the samples (contaminated samples) generate models with parameters scattered over the Θ. This observation is also exploited in Hough Transformation [23, 25], see also 4.7.

3.1. Assumptions

There are two tacit assumptions in derivation of the number of trials (eq. 3.3) needed to guarantee the confidence $1 - \eta_0$ in the RANSAC solution.

	15%	20%	30%	40%	50%	70%
2	132	73	32	17	10	4
4	5916	1871	368	116	46	11
7	$1.75 \cdot 10^6$	$2.34 \cdot 10^5$	$1.37 \cdot 10^4$	1827	382	35
8	$1.17 \cdot 10^7$	$1.17 \cdot 10^6$	$4.57 \cdot 10^4$	4570	765	50
12	$2.31 \cdot 10^{10}$	$7.31 \cdot 10^8$	$5.64 \cdot 10^6$	$1.79 \cdot 10^5$	$1.23 \cdot 10^4$	215
18	$2.08 \cdot 10^{15}$	$1.14 \cdot 10^{13}$	$7.73 \cdot 10^9$	$4.36 \cdot 10^7$	$7.85 \cdot 10^5$	1838
30	∞	∞	$1.35 \cdot 10^{16}$	$2.60 \cdot 10^{12}$	$3.22 \cdot 10^9$	$1.33 \cdot 10^5$
40	∞	∞	∞	$2.70 \cdot 10^{16}$	$3.29 \cdot 10^{12}$	$4.71 \cdot 10^6$

Table 3.1.: Number of samples for different model complexity m and the fraction of inliers ε.

Assumption 1 (A1) *An all-inlier sample generates model consistent with all inliers.*

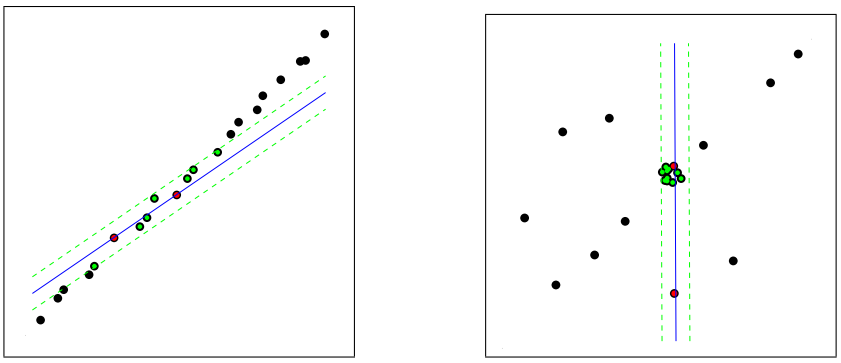
Assumption 2 (A2) *A model consistent with a sample contaminated by at least one outlier has small support.*

Assumption **A1** is used in equation (3.2) where the probability that a good model is not found is identified with the probability $1-P(I)$ that an all-inlier sample is not drawn. The assumptions **A2** assures that it is possible to distinguish between correct and incorrect solution based on the size of the support.

The RANSAC algorithm works as predicted by eq. (3.3) if the two assumptions **A1** and **A2** hold. In the case when one or both of the assumptions are violated, the quality of the solution is affected and/or the running time is longer. Such a situation is not a rare theoretical case. Fig. 3.1 shows the case of line estimation when the assumptions **A1** and **A2** are violated respectively.

The assumption **A1** does not hold in Fig. 3.1a. Due to the noise in the measured data points, even some all-inlier sample defines model that is not consistent with all inliers. This reduces the estimated number of inliers I^*. The number of samples is increased (the running time is longer) according to equations (3.2) and (3.3) as a consequence.

In Fig. 3.1b, the set of data points contains a cloud of points that are close to each other. All data points from the cloud are consistent with any line passing through it. Such a configuration of points that is consistent with infinite number of models (lines in this case) is called a *degenerate configuration*. A line defined by an outlier and a point from the degenerate configuration (as in Fig. 3.1b) is both well conditioned and supported by a high number of data points. Such a sample violates the assumption **A2**. Moreover, since the solution is well conditioned, it appears difficult to recognize such a case. Note also that due to the presence of the degenerate configuration in Fig. 3.1b, the assumption **A1** does not hold either. A line defined by two points from the degenerate configuration is ill-conditioned and can have arbitrary direction. As a result, wrong model (a line) that includes the degenerate configuration can be selected as an output by RANSAC.



(a) (b)

Figure 3.1.: Two examples of point configurations that cause RANSAC to output non-optimal result.

3.2. Detail Analysis of RANSAC

The probability of drawing (without replacement) an uncontaminated sample, eq. (3.1), is often approximated by $P(I) = \varepsilon^m$ for simplification. Such an approximation would be exact if the sampling is done with replacement, or with infinite number of data points (i.e. $\lim_{N\to\infty} P(\varepsilon N) = \varepsilon^m$). The equality $P(I) = \varepsilon^m$ is a good approximation for large values of N.

3.2.1. Time to the First Success

Under assumption **A1**, the probability of finding the optimal model parameters that are consistent with all inliers is the same as the probability $P(I)$ of drawing an uncontaminated sample. Since the samples are independent, the probability of drawing the first uncontaminated sample in k-th draw follows the geometric distribution

$$P_k = P(I)(1 - P(I))^{k-1}.$$

The average number of random samples drawn before the first uncontaminated sample occurs, i.e. the mean value $E(k) = \sum_{k=1}^{\infty} kP_k$, is [43]

$$\bar{k} = \frac{1}{P(I)} \approx \frac{1}{\varepsilon^m}. \tag{3.4}$$

The symbol k_{η_0} is used for the number of samples that have to be drawn to ensure a confidence $1 - \eta_0$ in the solution when there are I inliers $\mathcal{U}$, i.e. $k_{\eta_0}(I) = \ln(\eta_0)/\ln(1 - P(I))$. It is interesting to observe the relation between the average number of samples before the first uncontaminated sample is drawn $\bar{k}$ and the number of samples k_{η_0} needed to ensure the confidence $1 - \eta_0$ in the solution. Using the inequality $e^{-x} \geq 1 - x$ in the form of $-x \geq \ln(1 - x)$ we obtain

$$k_{\eta_0} \leq \frac{\ln(\eta_0)}{-P(I)},$$

and finally

$$\bar{k} \leq k_{\eta_0} \leq -\ln(\eta_0)\bar{k}. \tag{3.5}$$

Note, that the approximation $-x \approx \ln(1-x)$ is close for small positive x. Since $P(I) \approx \varepsilon^m$, the approximation

$$k_{\eta_0} \approx -\ln(\eta_0)\bar{k} \tag{3.6}$$

is close for small ε^m.

Since one uncontaminated sample is drawn within $\bar{k}$ samples on average, the number of uncontaminated samples within k_{η_0} samples is given by

$$k_{\eta_0}/\bar{k} \leq -\ln(\eta_0). \tag{3.7}$$

It follows from equation (3.7) that for for $\eta_0 = 0.05$ there are less than 3 uncontaminated samples drawn on average before the 95% confidence in the solution is obtained.

3.2.2. The Average Running Time

What is the average number of samples drawn before the RANSAC algorithm terminates? With the probability $1-\eta_0$ the set of inliers is found up to k_{η_0}-th sample. We will assume that any model that is not consistent with the set of inliers $\mathcal{I}$ will have significantly smaller support, and hence would not cause the termination of the algorithm. Then, after k_{η_0} steps, the algorithm is terminated as soon as an uncontaminated sample is drawn, in $\bar{k}$ samples on average. The relation for the average number $\hat{k}$ of samples drawn before termination is summarized in the following equation

$$\hat{k} = (1-\eta_0)k_{\eta_0} + \eta_0(k_{\eta_0} + \bar{k}) = k_{\eta_0} + \eta_0\bar{k} \approx (\eta_0 - \ln(\eta_0))\bar{k}. \tag{3.8}$$

3.2.3. The Best Sample So Far

RANSAC tries to find a single global maximum, i.e. a sample with the largest support. Therefore, samples with support lager than support of previously drawn samples (the-best-so-far samples) are the only candidates to become a final solution. How often may a new maximal support for an hypothesis during the sampling occur? Or how often does RANSAC store its the-best-so-far result in the 2b step of the algorithm (Alg. 1)? The number of data points consistent with a model from a randomly selected sample can be thought of as a random variable with an unknown density function. This density function is the same for all samples, so the probability that k-th sample will be the best so far is $1/k$. The average number of maxima reached within k samples is thus

$$\sum_{x=1}^{k} \frac{1}{x} \leq \int_1^k \frac{1}{x}\,dx + 1 = \ln k + 1. \tag{3.9}$$

3.3. Enhancing RANSAC

Two classes of enhancements to RANSAC are described in the rest of the thesis. First, the focus is turned towards the class of algorithms that improve the quality of the solution and/or make the applicability broader and more robust. Approaches that speed the RANSAC algorithm up while preserving the quality of the solution fall into the other class.

3.3.1. Robustifying RANSAC

As discussed in section 3.1, RANSAC relays on two assumptions. To make RANSAC robust to the violation of the assumptions, two approaches are proposed and discussed in the thesis. The assumptions are not fully removed, but replaced by a weaker assumptions.

Assumption 3 (A1') *All-inlier sample, even in the presence of noise, gives a model that is close to the optimal one. A local optimization of the model parameters reaches the optimal model that is supported by all inliers.*

Details of the LO-RANSAC algorithm that includes the Local Optimization step are discussed in chapter 5.

Violation of assumption **A2** is caused by the presence of a degenerate configuration (otherwise we speak of a multiple occurrence of the model in the data, such as two independent motions). With increasing dimension of the data points it becomes less likely that the degenerate configuration is consistent with an incorrect model, unless the data points from the degenerate configuration are included in the sample.

Assumption 4 (A2') *A model consistent with a sample that is contaminated by at least one outlier and* does not *contain a degenerate configuration as a subset has small support.*

An algorithm called DEGENSAC that examines samples with high support in order to avoid problems caused by degenerate configurations is described in chapter 6. The algorithm can, in fact, exploit the presence of degenerate configuration to speed the estimation up.

3.3.2. Beyond Random Sample and (Deterministic) Consensus

Two principal modification to the original algorithm are introduced in the thesis in order to reduce the time complexity of the procedure, one considering the hypothesis generation and the other the verification step.

The standard RANSAC algorithm treats all data points equally, since the samples are drawn uniformly from the input set of data points. However, a typical construction of data points provides / is based on some quality measure of the data points. In the correspondence problem, instances of such a quality measure are the cross-correlation score of the neighbourhoods of the interest points, or Euclidean (or Mahalanobis) distance of image transformation (affine) invariant descriptors of the interest points (regions). The quality measure implies ordering on the input data. Natural approach would be to concentrate only on the high-quality data points (this actually is the case of many algorithms for the correspondence problem: only correspondences with the

quality higher than a threshold are input to the robust geometry estimation stage). A hypothesis generator exploiting the ordering can significantly speed up the RANSAC procedure, as shown in chapter 7.

In chapter 8, the speed-up of RANSAC is achieved by randomization of the verification step. During the course of the algorithm, majority of the model parameters are influenced by outliers. Those 'bad' models typically have only small consensus set. Therefore, to reject such models, it is not necessary to execute the verification step on all data. It is sufficient to perform a statistical test on only a small number of data points.

4 State of the art

In this chapter, previous work on robust estimation algorithms similar to RANSAC is reviewed. We focus on methods that are generally applicable.

4.1. MLESAC

The MLESAC algorithm is an example of RANSAC that uses different cost function than the cardinality of the support. The algorithm was introduced by Torr and Zisserman [57] and further improvements were made by Tordoff and Murray [53] (see also section 4.2). Instead of maximizing the support of the model, the likelihood of the model is maximized. The error distribution is represented as a mixture of inlier and outlier distributions. The distribution of the residuals r_i is supposed to be Gaussian $N(0, \sigma)$ with known σ for the inlier data points, and uniform in $\langle 0, Z \rangle$ on outliers.

$$p(r_i|\theta) = \varepsilon \left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{r_i^2}{2\sigma^2}} \right) + (1 - \varepsilon)\frac{1}{Z} \tag{4.1}$$

The hypothesis generation phase is identical to RANSAC. The cost function evaluated in the verification steps is

$$J_{MLE}(\theta) = \sum_{i=1}^{N} \log p(r_i|\theta).$$

In [57], the $J_{MLE}(\theta)$ is calculated while the unknown mixing parameter ε is estimated using the EM algorithm [14, 48]. The mixing parameter ε describes how large fraction of data point came from the inlier distribution, i.e. gives the estimate of the fraction of inliers. This estimate is used to determine the number of samples as in standard RANSAC.

In tasks where the likelihood function on the model parameters has a strong peak, such as wide baseline matching, the cost functions J_S and J_{MLE} will be closely related. Since only three uncontaminated samples are drawn on average (see section 3.2.1) during the execution of RANSAC (as well as MLESAC), both cost functions are likely to select the same model as a result. On the other hand, the MLESAC algorithm benefits in tasks where the likelihood is flat and when it is not possible to sharply distinguish inliers and outlier. A typical instance of such a task is narrow baseline matching.

The following variant of MLESAC [57, 41] has been proposed for a narrow baseline matching, when a high fraction of correct matches is expected. The random sampling generates a fixed number of hypotheses. The number has to be high enough, so that the parameter space is sampled

with sufficient density around the optimal model parameters. Note that the fixed number of samples imposes a restriction on the fraction of inliers. Each hypothesized model is then scored by the J_{MLE} cost function and the model that maximizes the likelihood is selected as a result of the algorithm. This approach proved to be successful and suitable for a real time narrow baseline matching [40].

4.2. Guided Sample and Consensus

In [53], Tordoff and Murray introduced two extensions to the MLESAC algorithm which guide the selection of data points. It has been shown that the guided matching reduced the number of iterations required for a given confidence in the solution in a narrow baseline stereo matching.

In the paper, an observation that better hypothesis have better score J_{MLE} for any mixing parameter was presented. Hence, it is not necessary to solve for it using the EM algorithm. To select the best hypothesis, it is sufficient to use a fixed mixing parameter for all the hypotheses, $\varepsilon = 0.5$ for example.

Instead of using a global mixing parameter ε as in equation (4.1), it is preferable to use individual priors on a correspondence being correct if available. The unobservable property of correctness of i-th correspondence is denoted as v_i. Knowing the probabilities $p(v_i)$, the equation (4.1) is rewritten as

$$p(r_i|\theta) = p(v_i)\left(\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{r_i^2}{2\sigma^2}}\right) + (1 - p(v_i))\frac{1}{Z} \tag{4.2}$$

In a *narrow baseline* stereo matching problem, it is possible to learn the prior probabilities from the number of possible matches and the match scores. Let i-th feature have n_i possible matches with correctness denoted as $v_{ij} : j = 1 \dots n_i$ and quality (correlation score) s_{ij}. It was shown in [53], that in many narrow baseline image pairs the distribution of s_{ij} of mismatches ($\bar{v}_{ij}$) followed the following distribution

$$p(s_{ij}|\bar{v}_{ij}) \approx \frac{3(1 - s_{ij}^2)}{4}. \tag{4.3}$$

The correlations scores of correct matches (v_{ij}) were distributed as

$$p(s_{ij}|v_{ij}) = a\frac{1 - s_{ij}}{\alpha^2}e^{\frac{2(1-s_{ij})}{\alpha}}, \tag{4.4}$$

where α is a 'compactness' and a is a normalization constant, such that the area under the curve is unity. The probabilities are plotted in Fig. 4.1. Assuming that any of n_i possible matches of i-th feature is a priori equally likely (including the case that none is correct), the priors are set to $p(v_{ij}) = 1/(n_i + 1)$ and $p(\bar{v}_{ij}) = n_i/(n_i + 1)$. Finally, the probability that a tentative match ij is correct given the scores of all tentative matches is

$$p(v_{ij}|s_{i,1..n_i}) = \frac{p(v_{ij}|s_{ij})\prod_{k\neq j}^{n_i} p(\bar{v}_{ik}|s_{ik})}{\sum_{l=1}^{n_i}\left(p(v_{il}|s_{il})\prod_{k\neq l}^{n_i} p(\bar{v}_{ik}|s_{ik})\right) + \prod_{k=1}^{n_i} p(\bar{v}_{ik}|s_{ik})}. \tag{4.5}$$

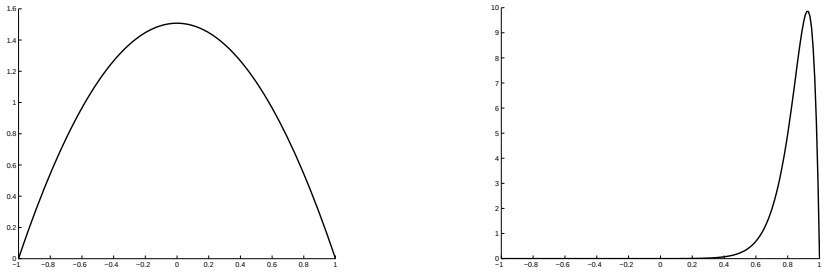


Figure 4.1.: Probabilities $p(s_{ij}|\bar{v}_{ij})$ and $p(s_{ij}|v_{ij})$.

The probability $p(v_{ij}|s_{i,1..n_i})$ is used as an estimate of the prior probability $p(v_i)$ in equation (4.2).

The hypothesis generation step is made more efficient by replacing the uniform random sampling of data by Monte-Carlo sampling with probabilities proportional to $p(v_i)$, i.e. that the correspondences with higher a priori probability to be inliers are sampled more often.

4.3. NAPSAC

The algorithm of NAPSAC [39] (N Adjacent Points SAmple Consensus) focused on the efficiency of the sampling strategy. The idea is that inliers tend to be closer to one another than outliers, and the sampling strategy can be modified to exploit the proximity of inliers. This might bring efficient robust recovery of high dimensional models, where the probability of drawing an uncontaminated sample becomes very low even for data sets with relatively low contamination of outliers.

Let us consider estimation of d-dimensional manifold in n-dimensional space. Let the measured data points $\mathcal{U}$ come from a bounded region only. Note that an epipolar geometry represents a bilinear 3-dimensional variety in a 4-dimensional joint image coordinate space, which is naturally bounded by the size of the images (not omnidirectional). Let the outliers be uniformly distributed within that region and the inliers be uniformly distributed on the manifold within the same region. Now, imagine a hyper-sphere with radius r centered at some point of the manifold. Within the hyper-sphere, the number of inliers is proportional to r^d where as the number of outliers to r^n, where $n > d$. Hence, with decreasing radius r the probability of outlier appearance in the hyper-sphere decreases faster than the probability of inlier appearance. However, estimation of the radius with higher fraction of inliers is not trivial. In [39], the derivation of optimal radius r is shown for two dimensional case.

The sampling phase of the algorithm is summarized in Alg. 2. First, an initial data point $\mathbf{x}_0$ is selected at random. Then, if possible, the sample is topped up by $m - 1$ data points selected at random from those, that are closer than r to $\mathbf{x}_0$.

a	Select initial data point $\mathbf{x}_0$ from $\mathcal{U}$ at random
b	Find points $\mathcal{R}_{\mathbf{x}_0} \subset \mathcal{U} \setminus \{\mathbf{x}_0\}$ lying within a hyper-sphere of radius r centered on $\mathbf{x}_0$
c	Fail if $\lvert\mathcal{R}\rvert < m - 1$ or form a sample $\mathcal{S}_k$ from $\mathbf{x}_0$ and $m - 1$ data points selected from $\mathcal{R}_{\mathbf{x}_0}$ at random

Algorithm 2: The sampling phase of NAPSAC.

4.4. Preemptive RANSAC

Preemtive RANSAC by Nistér [41] exploits the idea of randomizing the model verification step that was introduced in [5] and is described in detail in chapter 8.

The overall speed of the RANSAC algorithm is a product of the number of hypotheses generated and the time spent while processing each hypothesis. The number of hypotheses depends on the size of the sample and the fraction ε of inliers within the data. The time to process a hypothesis consists of the time needed to generate the hypothesis θ and from the evaluation of the cost function $J(\theta)$ over all data points in $\mathcal{U}$. The evaluation of the cost function is linearly dependent on the number N of data points. For large numbers of points, the algorithm may spent most of the time in the verification part. In RANSAC, the majority of the samples are contaminated, as shown in equation (3.4). Hence, if one was able to reject bad (contaminated) hypotheses as soon as possible then the whole process could be speeded up.

The preemptive RANSAC is based on comparing hypotheses after evaluating each hypothesis on just a fraction of the data points and rejecting those having worse scores. The algorithm uses a non increasing function $t(i)$, $i \in 0 \ldots N$ that determines how many hypotheses are left 'alive' after the cost function was evaluated on i data points. The data points to be verified are selected at random, with replacement. The algorithm proceeds as follows. First, $k = t(0)$ hypotheses are generated. Then, one data point is chosen for each active hypothesis, and the cost function is updated using that data point. After i data points were verified for each hypothesis, $t(i)$ best hypotheses are kept, and the others are rejected. The procedure stops when only a single hypothesis is kept, i.e. $t(i) = 1$, or after verifying all N data points for each surviving hypothesis.

1	Generate all $k = t(0)$ hypotheses
2	For $i = 1$ to N - evaluate the cost function for randomly chosen (with replacement) data point for all $t(i-1)$ hypotheses left - keep $t(i)$ hypotheses with the highest value of the cost function over i data points

Algorithm 3: The structure of preemptive RANSAC

Note: The algorithm described here is very efficient when one has to find some result within given fixed time. This is exactly the task the preemptive RANSAC was designed for [40]. Pre-

emptive RANSAC can fail, when the fraction of inliers is too low for the given k hypothesis to contain a uncontaminated one.

4.5. MINPRAN

The goal of RANSAC is to find a model that maximizes the number of data points with error smaller than user defined threshold Δ. In some problems, it might not be possible to know the error bound Δ for inlier data beforehand. In MINPRAN [51], this parameter of RANSAC is replaced by an assumption of uniform distribution of outliers. The MINPRAN algorithm selects model that MINimizes the Probability of RANdomness as a solution.

Let residuals of outliers be uniformly distributed within values $\langle 0, Z \rangle$ and let the residuals in k-th iteration be $r_k(i)$ for $i = 1 \dots N$ and let the residuals be sorted, i.e. let $r_k(i) \leq r_k(j)$ for $i < j$. Then the probability of randomness of first i residuals is expressed as

$$\mathcal{F}_k(i) = \sum_{j=i}^{N} \binom{N}{j} \left(\frac{r_k(i)}{Z}\right)^j \left(1 - \frac{r_k(i)}{Z}\right)^{N-j}. \tag{4.6}$$

From the definition (4.6), $\mathcal{F}_k(i) < \mathcal{F}_l(i)$ if and only if $r_k(i) < r_l(i)$ for any fixed $i \in \{1 \dots N\}$. Hence, to be able to determine $\min_{k,i} \mathcal{F}_k(i)$ it is sufficient to keep only values $\min_k r_k(i)$ for all i. The algorithm then proceeds as follows: in the verification step, the error function is evaluated in $\mathcal{O}(N)$. Then, the residuals are sorted ($\mathcal{O}(N \log N)$). For each $i = 1 \dots N$ minimal residual $r_k(i)$ is kept ($\mathcal{O}(N)$). Finally, the probability of randomness (4.6) is evaluated ($\mathcal{O}(N^2)$) and the solution with the minimal probability of randomness is selected. The total time complexity after k iterations of the algorithm is $\mathcal{O}(N^2 + kN \log N)$.
Note: (On the time complexity.) Consider the following sorting algorithm when uniform distribution of residuals is assumed. First, divide residuals into $\mathcal{O}(N)$ buckets and then use sorting algorithm in $\mathcal{O}(n \log n)$ for each bucket. When an contaminated sample was drawn, the probability $p(n)$ of falling n data points into a single bucket is driven by a multinomial distribution, which is marginalized by binomial distribution (inside / outside the bucket) as follows

$$p(n) \leq \binom{N}{n} \frac{1}{N^n} \left(1 - \frac{1}{N}\right)^{N-n} \leq \binom{N}{n} \frac{1}{N^n} = \frac{N!}{(N-n)!\, n!\, N^n} \leq \frac{1}{n!}.$$

The average time to sort data points in a single bucket is then

$$\sum_{n=1}^{N} p(n) n \log n \leq \sum_{n=1}^{N} \frac{n \log n}{n!} = \mathcal{O}(1).$$

Therefore, the sorting of all data points can be done in $\mathcal{O}(N)$ when a contaminated sample is drawn. Let c denote the number of times when an all-inlier sample was drawn. When an all-inlier sample is drawn, all data points can fall into a single bucket and the sorting will take $\mathcal{O}(N \log N)$. In practical situations, we can assume that $c \log N < k$ (see section 3.2.1). Then, the time complexity of the algorithm would be $\mathcal{O}(N^2 + kN)$.

4.6. Least Median of Squares

In LMS [46], it is assumed that at least half (or, more generally, a given α-quantile) of the data points are inliers. Therefore, the parameter Δ representing an error bound for inliers is not needed. The model with the smallest median (α-quantile) error is output as a result. The number of samples drawn must be sufficient to ensure certain confidence that in the worst case (if the proportion of inliers was exactly half (α) of the data) at least one all-inlier sample was drawn.

In some tasks, such as the correspondence problem, it is more reasonable to expect a knowledge of the error bound Δ (precision of the detector, maximal acceptable error for purposes of reconstruction, etc) than of the minimal fraction of inlier.

4.7. Hough Transform

Under two conditions, (the approximations of) all local extrema of the cost function $J(\theta)$ can be recovered efficiently. The first condition is that the cost function must be a sum of contributions of all data points. Second, there must be an efficient procedure enumerating all parameters θ consistent with a single data point. Hough transform [23, 25] then proceeds as follows. First, the permissible part of the parameter space Θ is discretised into a grid and an accumulator is allocated for each point of the grid. Then, every data point casts a vote for all bins that contain parameters of a model consistent with the data point. The votes from different data points are summed in the bins.

The Hough transform deals easily with the presence of multiple objects, which often happens in the detection of geometric primitives in the image. Conversely, Hough transformation is not suitable for detection of high-dimensional models, as the space needed to represent the discretised parameter space grows exponentially with the dimension of the model. Hence, the Hough transform is suitable for example for line detection (2 dimensional model), less suitable for the detection of an affine transformation (6 D) and intractable for the detection of a trifocal tensor (18 D).

A number of variants of Hough transform has been published. Randomized Hough Transform [63, 27] lies somewhere between Hough transform and RANSAC. Randomized Hough Transform generates samples in the same manner as RANSAC does. The parameters of all hypothesized models are stored in a data structure allowing finding similar parameters. When some of the parameters appeared a non-random number times, the inliers to those parameters are detected in the same was as in the verification step of RANSAC. The detected inliers of the model are removed from further sampling and the procedure continues.

5 LO-RANSAC

The termination criterion of RANSAC tacitly assumes (assumption **A1**, section 3.1), that every all-inlier (uncontaminated) sample is 'good', i.e. that every model computed from *an uncontaminated sample* is consistent with *all* inliers. This favorable situation is rare in practice, as shown in Fig. 3.1. The invalidity of assumption **A1** results in an instability (different model parameters and different inliers are output by the algorithm) of the solution, and in an increase in the number of samples drawn before the termination of the algorithm. Such a behavior of RANSAC in epipolar geometry estimation has been observed experimentally in [53].

The discrepancy between the predicted and observed RANSAC running times can be explained in terms of cardinalities of the set $\mathcal{B}$ of 'good' samples and the set $\mathcal{A}$ of all-inlier samples. Since RANSAC generates hypotheses from minimal sets, both noise in the data and poor conditioning of the model computation can 'remove' an uncontaminated sample from the set $\mathcal{B}$. Problems occur when the cardinality of $\mathcal{A}$ is significantly larger than the cardinality of $\mathcal{B}$, i.e. when the probability of drawing a sample from $\mathcal{B}$ becomes too small compared with the probability of drawing a sample from $\mathcal{A}$. A sample from $\mathcal{A} \setminus \mathcal{B}$ has smaller consensus set than a sample from $\mathcal{B}$, and hence the termination criterion based on the largest support so far requires higher number of samples. The effect is clearly visible in the histograms of the number of inliers found by standard RANSAC. The first column of Fig. 5.2 shows the histogram for five matching experiments. The number of inliers varies by about 20-30%, which leads to an increase in the number of RANSAC cycles by a factor of two to three. In the case of epipolar geometry estimation from three region-to-region correspondences (analyzed in Section 5.3), the ratio of cardinalities of $\mathcal{B}$ and $\mathcal{A}$ is so low that it renders the approach, in connection with standard RANSAC, impractical.

The LO-RANSAC is based on the observation (transferred into assumption **A1**', section 3.3.1) that virtually all models estimated from an uncontaminated minimal sample contain large fraction of inliers within their support. An optimization[1] process starting from the the-best-so-far hypothesized model is therefore inserted into RANSAC. *Applying the proposed optimization step produces an algorithm with a near perfect agreement with theoretical (i.e. optimal) performance*. In other words, LO-RANSAC makes the sets $\mathcal{A}$ and $\mathcal{B}$ almost identical. Therefore eq. (3.2) becomes valid for LO-RANSAC. Moreover, it is shown in section 5.1.1 that the cost of the additional step in negligible.

The potential of the proposed extension of the RANSAC algorithm is supported by two new matching algorithms exploiting LO-RANSAC:

First, an algorithm estimating epipolar geometry from three region-to-region correspondences

[1] Note, that the LO-RANSAC does *not* try to compete with the bundle adjustment methods. The aim is to provide a better starting point for the bundle adjustment than standard RANSAC in shorter time.

Figure 5.1.: Correspondences established by the proposed algorithm on the 'Flower' pair. Tentative correspondences contained only 7% of inliers. Some of the corresponding points are connected between the images to visualize the correspondence.

is introduced (section 5.3). Exploiting the affine-invariant local frames described in [36], three point-to-point correspondences are found for each region-to-region correspondence. The expected run-time then falls from $\mathcal{O}(\varepsilon^{-7})$ to $\mathcal{O}(\varepsilon^{-3})$. The straightforward consequence is a significant enlargement of the class of problems that are solvable. The idea of using multiple points in the estimation process is in principle simple. However, since the three points associated with a single region are in close proximity, the precision of the estimated epipolar geometry may be questioned. The experiments confirmed, that acquisition of a new local optimization step into the RANSAC algorithm was essential to solve the problem.

Second, a modified algorithm for simultaneous estimation of epipolar geometry (EG) and radial distortion (RD) [16] is described in section 5.4. Experimental validation shows that the new estimator is superior to known algorithms in quality (the number of detected matches), precision and speed of the matching.

5.1. LO-RANSAC Algorithm

The LO-RANSAC proceeds as RANSAC, only one (LO) step is added. The local optimization step is carried out only if a new maximum in the number of inliers from the current sample has occurred, i.e. when standard RANSAC stores its best result. For a detailed description of the locally optimized RANSAC see Alg. 4.

Repeat until the probability (3.2) of finding model with support larger than I^* in k-th step falls under predefined threshold η_0:

1. Select a random sample of minimum size m from $\mathcal{U}$.
2. Estimate model parameters consistent with the sample.
3. Calculate the support I_k of the model, i.e. the data points with error smaller than a predefined threshold θ. If $I_k > I_j$ for all $j < k$ (i.e. when a new maximum is reached), then run:
   LO step. Apply optimization. Store the best model found and its support I^* ($I^* \geq I_k$ due to the optimization).

Algorithm 4: The structure of LO-RANSAC. Note, that the algorithm is terminated based on the optimized support I^*, whereas execution of the LO step depends on supports of sampled hypotheses I_j.

5.1.1. The Additional Computational Cost

The LO step is carried out only if a new maximum in the number of inliers is reached. It was shown in section 3.2.3, that the best-so-far sample is only drawn $\log k$ times, where k is the number of samples drawn – eq. (3.9).

Note, that this is the upper bound as the number of correspondences is finite integer and so the same number of inliers will occur often. This theoretical bound was confirmed experimentally, the average numbers of local optimization over an execution of (locally optimized) RANSAC can be found in Table 5.3.

The logarithmic growth of the number of LO step invocations as a function of the number of hypothesize-and-verify cycles allows application of relatively computationally expensive optimization methods without an impact on the overall speed of the algorithm.

5.1.2. LO Methods

The choice of the optimization method depends on the type of the model that is being fit, on the (type of) error on inliers and possibly on other factors. We have tested the following methods of local optimization on EG and homography estimation. The re-sampling method (inner RANSAC) is generally applicable, since it uses the same principles as RANSAC.

1. Standard. The standard implementation of RANSAC without any local optimization.
2. Simple. Take all data points with error smaller than Δ and use a linear algorithm to hypothesize new model parameters.
3. Iterative Scheme. Take all data points with error smaller that $K \cdot \Delta$ and use linear algorithm to compute new model parameters. Reduce the threshold and iterate until the threshold is Δ.
4. Inner RANSAC. A new sampling procedure is executed. Samples are selected only form I_k data points consistent with the hypothesised model of k-th step of RANSAC. New models are verified against whole set of data points. As the sampling is running on inlier data, there is no

need for the size of sample to be minimal. On the contrary, the size of the sample is selected to minimize the error of the model parameter estimation. In our experiments the size of samples are set to $\min(I_k/2, 14)$ for epipolar geometry (see results in Appendix A) and to $\min(I_k/2, 12)$ for the case of homography estimation. The number of repetitions is set to ten in the experiments presented.

5. Inner RANSAC with iteration. This method is similar to the previous one, the difference being that each sample of the inner RANSAC is processed by method 3.

The choice of the local optimization method can be application dependent. The methods described in this section tend to be general methods. The justification of the chosen methods follows.

Iterative Scheme. It is well known from the robust statistic literature [46], that pseudo-robust algorithms that first estimate model parameters from all data by least squares minimization, then remove the data points with the biggest error (or residual) and iteratively repeat this procedure do not lead to correct estimates. It can be easily shown, that a single far–outlying data point, i.e. leverage point, will cause a total destruction of the estimated model parameters. That is because such a leverage point overweights even the majority of inliers in least-squares minimization. This algorithm works only well, when the outliers are not overbearing, so the majority of inliers have bigger influence on the least squares.

In local optimization method 3 (and 5) there are no leverage points, as each data point has error below $K \cdot \Delta$ subject to the sampled model.

Inner RANSAC. Let parameters of the best so far model are denoted θ_k and the consensus set of this model $\mathcal{S}_k$. If the sample giving the parameters θ_k to arise was an all-inlier sample then, according to the assumption **A1**', the optimal parameters θ^* should be located "close" to θ_k in the parameter space. A sampling of $\mathcal{S}_k$ efficiently generates hypothesis of model parameters "close" to θ_k for reasonable (continuous) models and sufficiently small threshold Δ (that decides whether a data point supports or not a given model).

Since $\mathcal{S}_k$ is a consensus set $\mathcal{S}_k$ (of the model with parameters θ_k), the fraction of outliers should be low. Therefore, samples of non-minimal size drawn from $\mathcal{S}_k$ will not suffer from the contamination by outliers. The advantage of drawing non-minimal samples is in higher precision of the estimated model parametres, the more data points are used the lower the influence of the noise. It has been mentioned in [54] that estimates of epipolar geometry from seven points are more precise than those from eight points. Our experiment in appendix A confirms this observation and shows that the eight point algorithm with nine or more correspondences yields better estimates of EG whan the seven point algorithm. For details see appendix A.

5.2. Experiments on EG and Homography Estimation

The proposed algorithm was extensively tested on the problem of estimation of the two view relations (epipolar geometry and homography) from image point correspondences. Five experiments are presented in this section, all of them on publicly available data, depicted in Figures 5.3

	A	B	C	D	E
# corr	94	94	1500	160	94
# inl	57	27	481	30	17
ε	61%	29%	32%	19%	18%
# sam	115	34529	8852	2873	3837

Table 5.1.: Characteristics of experiments A-E. Total number of correspondences, maximal number of inliers found within all tests, fraction of inliers ε and theoretically expected number of samples.

and 5.4. In experiments A and B, the epipolar geometry is estimated in a wide-baseline setting. In experiment C, the epipolar geometry was estimated too, this time from short-baseline stereo images. From the point of view of RANSAC use, the narrow and wide baseline problems differ by the number of correspondences and inliers (see Table 5.1), and also by the distribution of errors of outliers. Experiments D and E try to recover the homography. The scene in experiment E is the same as in experiment A and this experiment could be seen as a plane segmentation. All tentative correspondences were detected and matched automatically.

Algorithms were implemented in C and the experiments were run on AMD K7 1800+ MHz processor. The terminating criterion based on equation (3.2) was set to $\eta < 0.05$. The threshold θ was set to $\theta = 3.84\sigma^2$ for the epipolar geometry and $\theta = 5.99\sigma^2$ for the homography. In both cases the expected σ was set to $\sigma = 0.3$.

The characterization of the matching problem, such as number of correspondences, the total number of inliers and expected number of samples, are summarized in Table 5.1. The total number of inliers was set to the maximal number of inliers obtained over all methods over all repetitions. The expected number of samples was calculated according to the termination criterion mentioned above.

Performance of local optimization methods 1 to 5 was evaluated on problems A to E. The results for 100 runs are summarized in Table 5.2. For each experiment, a table containing the average number of inliers, average number of samples drawn, average time spent in RANSAC (in seconds) and efficiency (the ratio of the number of samples drawn and expected) is shown. Table 5.3 shows both, how many times the local optimization has been applied and the theoretical upper bound derived in Section 6.3.

The method 5 achieved the best results in all experiments in the number of samples and differs slightly from the theoretically expected number. On the other hand, standard RANSAC is 2.5 – 3.3 slower. In Fig. 5.2 the histograms of the sizes of the resulting inliers sets are shown. Each column shows results for one method, each row for one experiment. One can observe that the peaks are shifting to the higher values with the increasing identification number of method.

Method 5 reaches the best results in terms of sizes of inlier sets and consequently in number of samples before termination. This method should be used when the fraction of inliers is low. Resampling, on the other hand, might be quite costly in the case of high number of inliers, especially if accompanied by a small number of correspondences in total) as could be seen in

		1	2	3	4	5
A	inl	49.7	53.9	55.9	56.0	**56.2**
	sam	383	205	129	117	**115**
	time	0.018	0.010	**0.007**	0.010	0.019
	eff	3.35	1.79	1.12	1.02	**1.01**
B	inl	23.3	24.4	25.0	25.5	**25.7**
	sam	90816	63391	49962	44016	**39886**
	time	3.911	2.729	2.154	1.901	**1.731**
	eff	2.63	1.84	1.45	1.27	**1.16**
C	inl	423.5	446.2	467.5	468.9	**474.9**
	sam	25205	16564	11932	10947	**9916**
	time	4.114	2.707	1.971	1.850	**1.850**
	eff	2.85	1.87	1.35	1.24	**1.12**
D	inl	23.9	26.7	28.1	28.8	**29.0**
	sam	8652	5092	3936	3509	**3316**
	time	0.922	0.543	0.423	**0.387**	0.391
	eff	3.01	1.77	1.37	1.22	**1.15**
E	inl	13.5	14.6	15.3	15.7	**15.9**
	sam	12042	8551	6846	5613	**5254**
	time	0.979	0.696	0.559	0.463	**0.444**
	eff	3.14	2.23	1.78	1.46	**1.37**

Table 5.2.: The summary of local optimization experiments: average number of inliers (inl) and samples taken (sam), average time in seconds and efficiency (eff). The best values for each row are highlighted in bold. For more details see the description in text in Section 5.2.

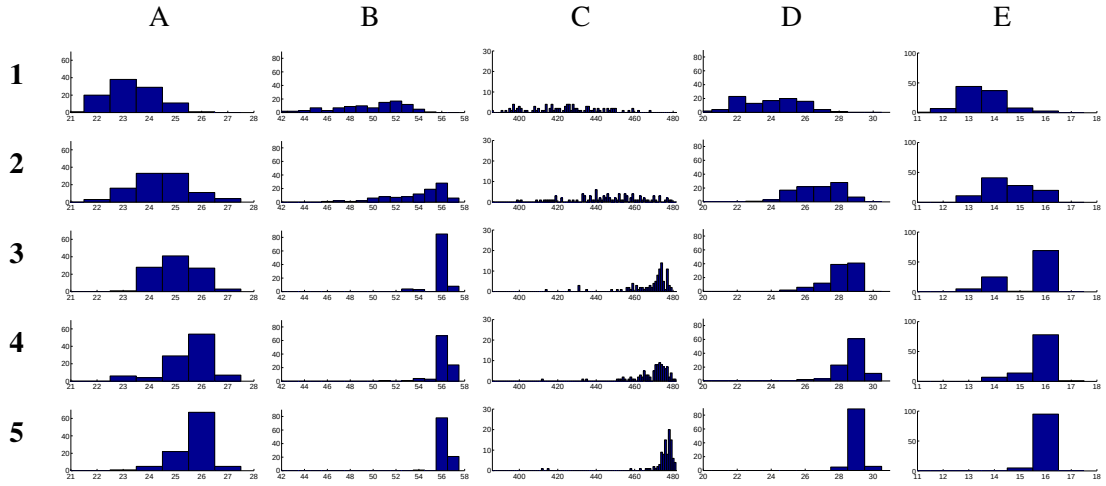


Figure 5.2.: Histograms of the number of inliers. The methods 1 to 5 (1 stands for standard RANSAC) are stored in rows and different dataset are shown in columns (A to E). On each graph, there is a number of inliers on the x-axis and how many times this number was reached within one hundred repetitions on the y-axis.

	1		2		3		4		5	
A	3.0	5.9	2.6	5.3	2.0	4.9	1.9	4.8	1.8	4.7
B	6.4	11.4	6.1	11.1	5.9	10.8	6.0	10.7	5.9	10.6
C	7.7	10.1	6.8	9.7	6.5	9.4	6.7	9.3	6.5	9.2
D	5.2	9.1	4.8	8.5	4.5	8.3	4.4	8.2	4.0	8.1
E	4.8	9.4	4.3	9.1	4.2	8.8	4.0	8.6	3.9	8.6

Table 5.3.: The average number of local optimizations ran during one execution of RANSAC and logarithm of average number of samples for comparison.

experiment A (61 % of inliers out of 94 correspondences). In this case, method 3 was the fastest. Method 3 obtained significantly better results than the standard RANSAC in all experiments, the speed up was about 100%, and slightly worse than for method 5. We suggest to use method 5. Method 3 might be used in real-time procedures when a high number of inliers is expected. Methods 2 and 4 are inferior to methods with iteration (3 and 5 respectively) without any time saving advantage.

5.3. 3LAF LO-RANSAC: EG from Three Correspondences

In wide-baseline matching, the process of selection of tentative correspondences often produces region-to-region mappings. However, the strong constraints the mappings provide are ignored in the EG estimation stage. The fundamental matrix is computed from seven point-to-point

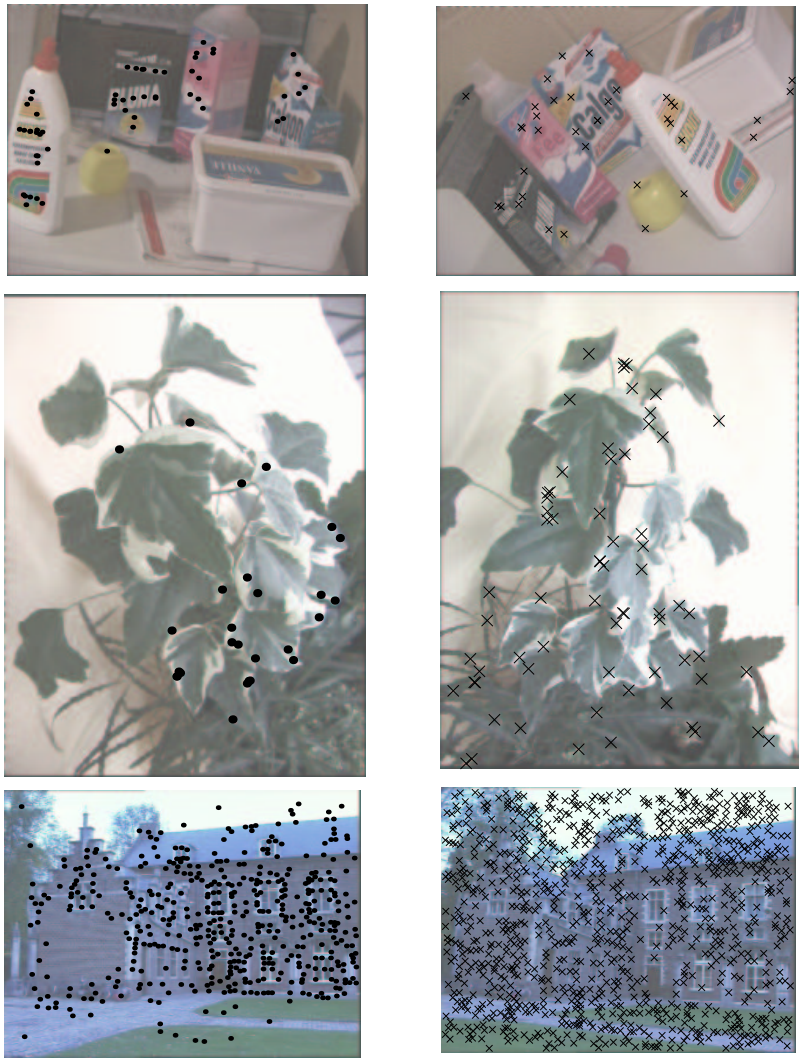


Figure 5.3.: Image pairs and detected points used in epipolar geometry experiments (A - C). Inliers are marked as dots in left images and outliers as crosses in right images.

Figure 5.4.: Image pairs and detected points used in homography experiments (D and E). Inliers are marked as dots in left images and outliers as crosses in right images.

correspondences, with each region-to-region mapping providing just a single point-to-point correspondence [60, 45, 36] .

In this section, we assume that a set of region-to-region tentative matches is available, and that three independent point-to-point matches can be obtained per each. In experiments, we used the output of the method published in [36], where the triplet of points originating from one region is called 'local affine frame' (LAF).

The algorithm is a straightforward application of LO-RANSAC. To hypothesize a model of epipolar geometry, random samples of three region correspondences are drawn. Three region correspondences give nine point correspondences. These are used to estimate the fundamental matrix F using the linear eight-point algorithm [19]. The LO step is applied (as always) only to the models that are so far the best and includes both the 'inner' RANSAC and iterative polishing, as described in Section 6.3. A region correspondence is consistent with a hypothesized epipolar geometry iff all three points are consistent.

5.3.1. Experiments

To highlight the advantage of 3LAF LO-RANSAC, tests were carried out on two image pairs (Fig. 5.1 and 5.6) with only about 7% and 20% of tentative correspondences correct respectively. The bookshelf test pair (Fig. 5.5) represents an indoor scene with large scale difference between the two views.

Results of the conducted experiments are summarized in Tab. 5.4. For the 'Flower' pair, the fraction of inliers is so low that the standard seven-point method failed. In the other two pairs, a significant speed-up measured by the number iterations was achieved.

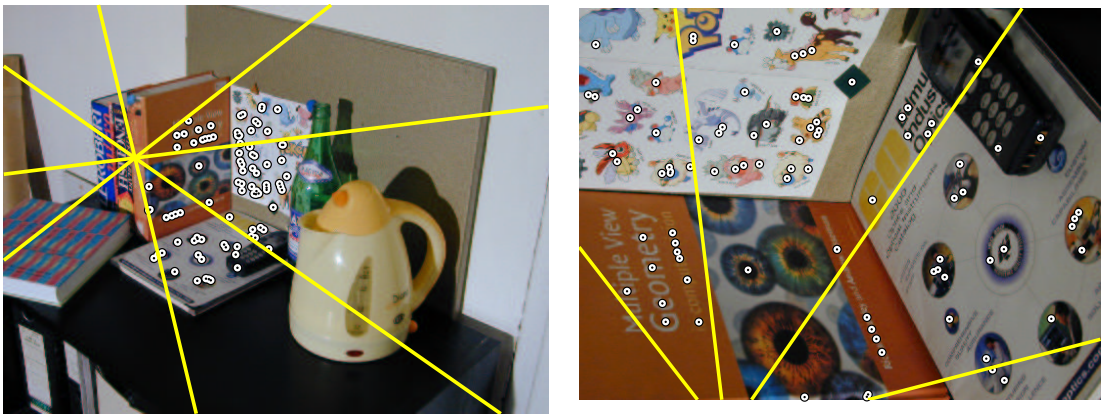


Figure 5.5.: Epipolar lines and 84 detected correspondences (markers) superimposed over the 'Bookshelf' pair.

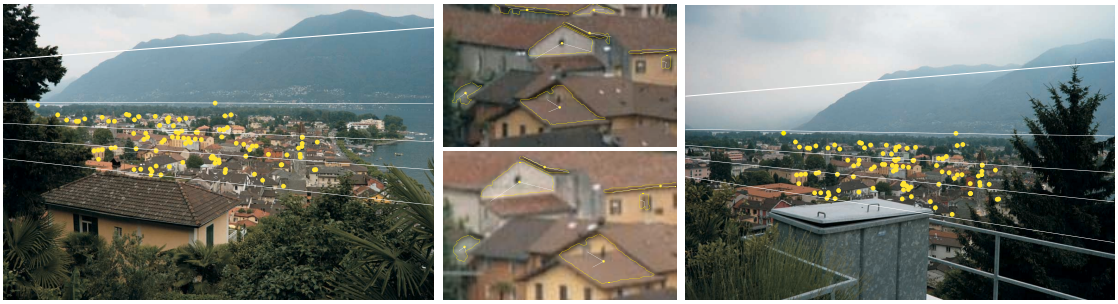


Figure 5.6.: Epipolar lines and correspondences superimposed over 'Ascona' pair. Close-ups with regions and LAFs highlighted are shown in the middle.

It is important to note that when applying the 3-frame RANSAC without the LO step (called 3LAF RANSAC), the set of detected inliers is significantly smaller, as shown in the middle column of Tab. 5.4. We believe that this is due to the fact that local affine frames are typically very small and the three points from a single region lie in near proximity. Consequently, the EG estimated from a minimal number of correspondences, as well as its support set, are unstable.

Clearly, application of the LO step is a very important ingredient of the newly proposed algorithm. As a final remark, we note that requiring all three point-to-point correspondence that form a region-to-region correspondence to obey the EG constraint also reduces the number of false positive matches, as the probability that a random correspondence will satisfy the epipolar constraint is decreased.

5.4. LO-RANSAC-RD: Estimating Radial Distortion

The benefits of exploiting a more complex model within the LO step can be demonstrated on the problem of simultaneous estimation of EG and radial distortion (RD). RD is a deviation

Method	EG consist.	iterations
Flower, **518** tentative corr., **7**% inliers		
7pt LO-RANSAC	N/A	$\approx 684\,000\,000$
3LAF RANSAC	25	47 668
3LAF LO-RANSAC	36	14 880
Bookshelf, **201** tentative corr., **42**% inliers		
7pt LO-RANSAC	83	1705
3LAF RANSAC	47	245
3LAF LO-RANSAC	84	41
Ascona, **584** tentative corr., **20**% inliers		
7pt LO-RANSAC	116	284 868
3LAF RANSAC	65	2 867
3LAF LO-RANSAC	116	384

Table 5.4.: Summary of experimental results. Number of correspondences found consistent with the epipolar geometry and the number of RANSAC iterations required to reach the solution. Note that all the numbers are random variables.

from the pinhole model commonly encountered in cameras, especially with wide-angle lenses. Fitzgibbon [16] introduced an algorithm for joint estimation of EG and RD given 9 point correspondences. It follows from eq. (3.1) that the price paid for the inclusion of Fitzgibbon's RD estimation is an increase in the number of samples drawn by a factor of $1/\varepsilon^2$. Since typically $\varepsilon \in [0.1, 0.5]$, the nine point RANSAC for EG and RD estimation (9pt RANSAC) is 4 to 100 slower than the standard 7pt RANSAC.

We present the simultaneous estimation of EG and RD as an algorithm in the LO-RANSAC framework. The algorithm draws 7 correspondence samples to estimate EG without RD in the hypothesize-verify loop and includes RD model in the LO step. Such sampling strategy ensures that LO-RANSAC-RD has the same time complexity $\mathcal{O}(\varepsilon^{-7})$ as the 7pt RANSAC. To parameterize the RD model, we have chosen the division model [16]

$$\mathbf{p} = \frac{1}{1 + \lambda|\mathbf{x}|^2}\mathbf{x},$$

where λ is the only model parameter, $\mathbf{x}$ stands for the measured point, $|\mathbf{x}|$ for the distance of $\mathbf{x}$ to the optical center (center of the image) and $\mathbf{p}$ for the undistorted point so that the epipolar constraint can be written as $\mathbf{p}'\mathbf{F}\mathbf{p} = 0$. EG and RD were simultaneously estimated in the LO step solving quadratic eigenvalue problem as in [16].

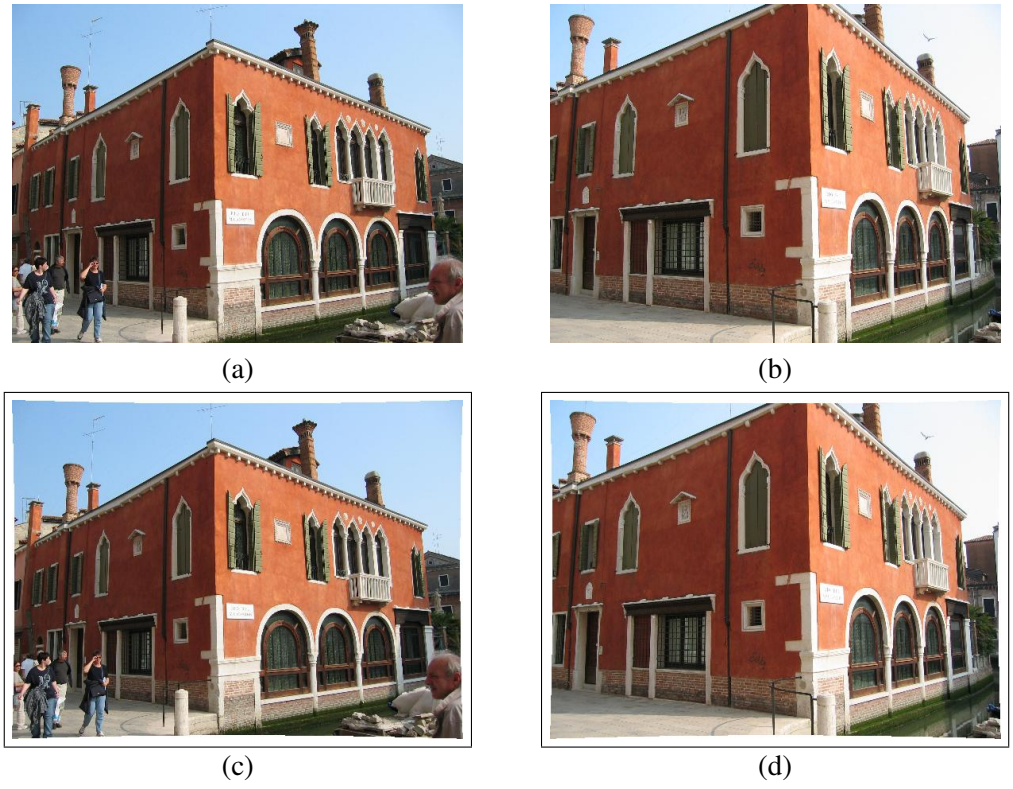
5.4.1. Experiments

Performance of three algorithms, 7pt LO-RANSAC (A), 9pt RANSAC (B) and LO-RANSAC-RD (C), was compared on image pairs with low RD (Orange house, Fig. 5.7) and high RD (Courtyard QY, Fig. 5.8). The number of detected inliers is shown in Fig. 5.9. Alg. B finds more inliers

RANSAC	7pt LO	9pt	LO–RD
Orange house	5 528	31 456	790
Courtyard QY	1 861	6 863	432

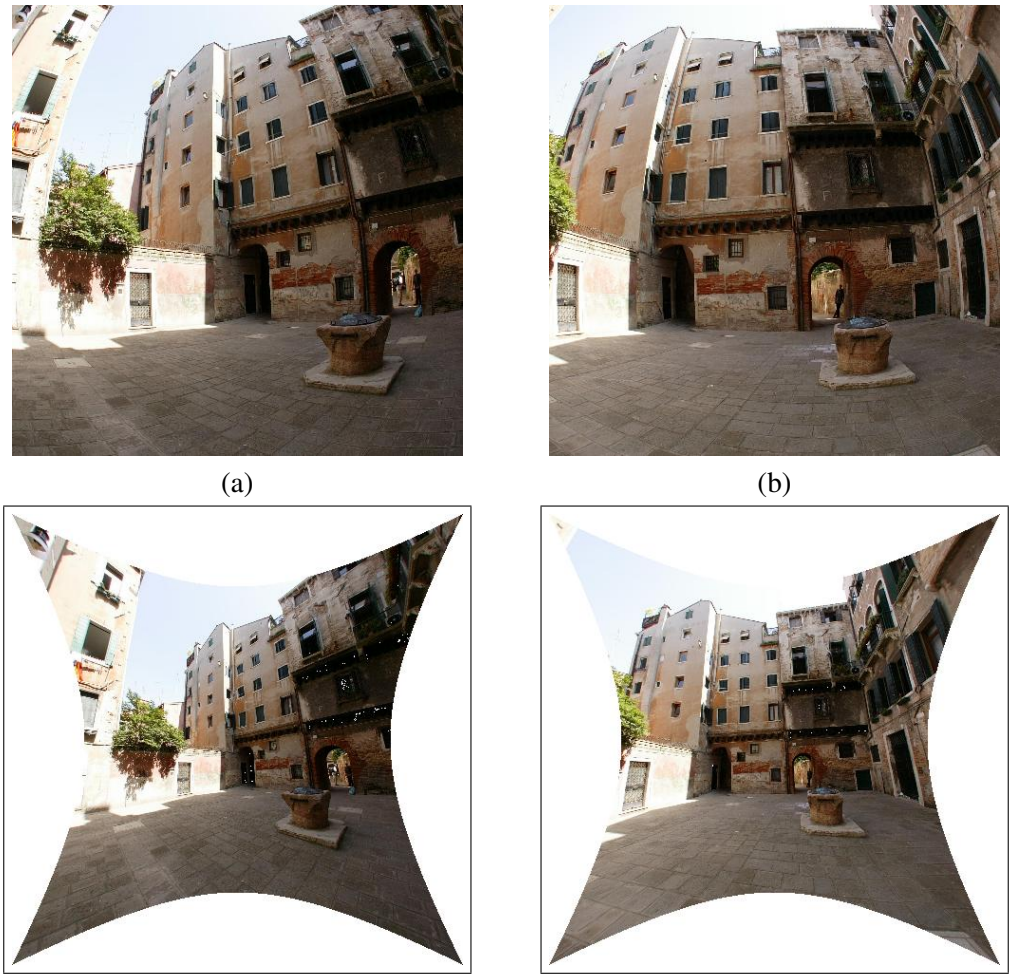
Table 5.5.: The average number of samples drawn over 100 runs of different RANSAC algorithm. Almost 40 times speed-up between 9pt RANSAC and LO-RANSAC-RD was observed.

than A because it uses a more precise model. Alg. C finds more inliers than B due to the LO step. The speed of A,B and C is measured by the number of samples drawn (Tab. 5.5). Alg. B is the slowest, as its time complexity is $\mathcal{O}(\varepsilon^{-9})$, compared with $\mathcal{O}(\varepsilon^{-7})$ of A and C. As a consequence of eq. (3.3), C terminates much earlier than A since it finds a higher number of inliers. Finally, the stability of the radial distortion estimation was measured. The graphs of the distribution of estimated parameter λ, depicted in Fig. 5.10, show that C is more stable than B – the variation of λ is smaller.



(a) (b) (c) (d)

Figure 5.7.: A low radial distortion example, 'Orange house' in Venice, 45% of inliers. Original image pair (top row) was taken by a compact digital camera. The radial distortion (with parameter $\lambda = -0.27$) was removed (bottom row) from the images.



(a) (b)

(c) (d)

Figure 5.8.: High radial distortion example: 'Courtyard QY' image pair, 48% of inliers. Central part (1400 x 1400 pixels, view field 100°) of the fish-eye images (top row). The radial distortion (with parameter $\lambda = -3.6$) was removed (bottom row) from the images.

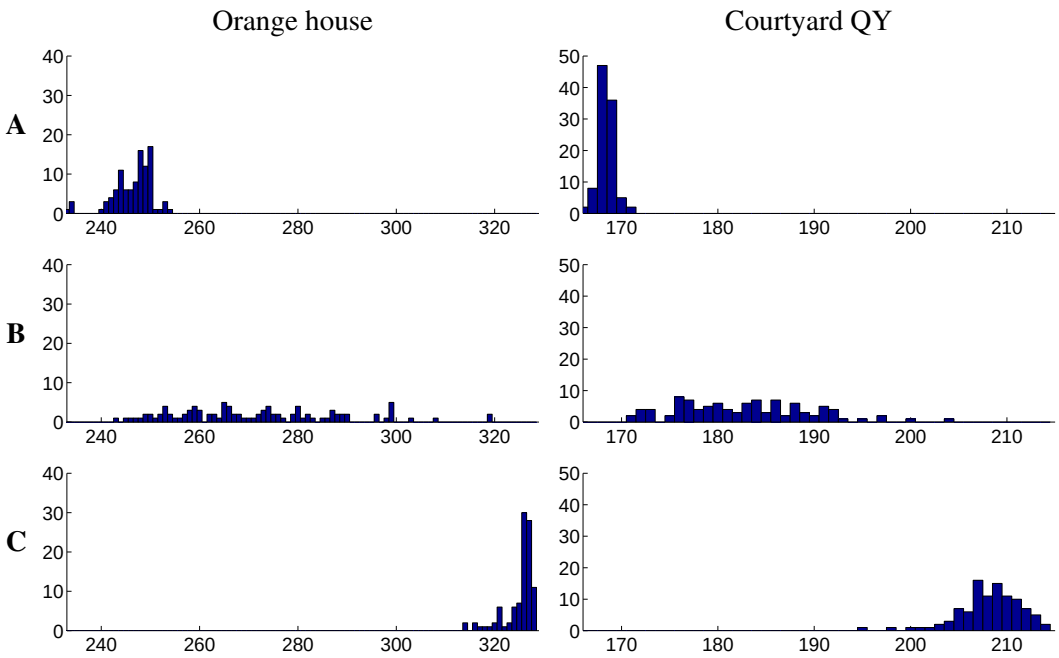


Figure 5.9.: Histograms of the numbers of inliers detected by different RANSAC algorithms: A - 7pt LO-RANSAC with no RD estimation, B - 9pt RANSAC, C - LO-RANSAC-RD.

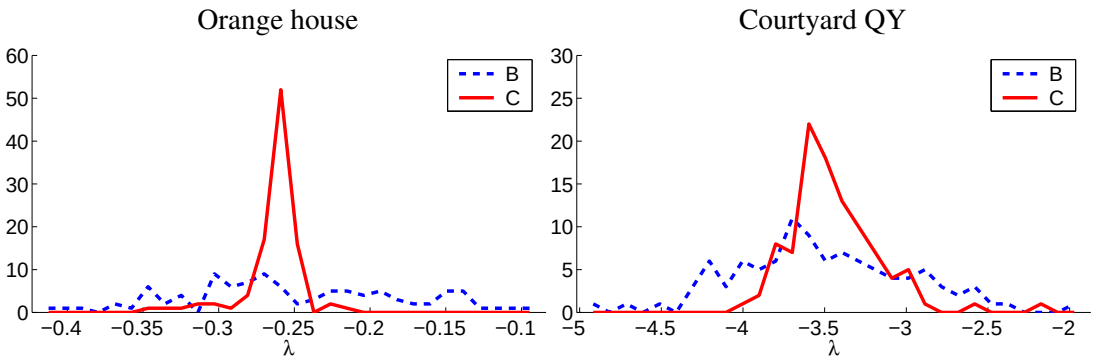


Figure 5.10.: The distribution of estimated parameter λ of radial distortion for 9pt RANSAC (B dashed) and LO-RANSAC-RD (C solid).

Figure 5.11.: The 'Lobby' pair, 28% of inliers. Central parts of the fish-eye images. Original first image on the left, the second image with radial distortion removed on right. Correspondences detected by 3LAF LO-RANSAC-RD superimposed over the images (markers).

5.5. 3LAF LO-RANSAC-RD: All-in-one

Estimation of RD and the idea of using LAFs for EG estimations can be applied at the same time. The 3LAF LO-RANSAC-RD algorithm estimates EG from three region-to-region correspondences as 3LAF LO-RANSAC, then uses the same LO step as LO-RANSAC-RD. The Courtyard QY problem Fig. 5.8 is solved by 3LAF LO-RANSAC-RD after less then 30 samples (compare with Tab. 5.5) with the same precision as when using LO-RANSAC-RD. A more complicated matching with only 28% of inliers (see Fig. 5.11) was solved after only 133 RANSAC cycles.

5.6. Conclusions

In the chapter, the LO in RANSAC algorithm was introduced. The number of detected inliers increased due to the local optimization, and consequently the number of samples drawn decreased. In all experiments, the running-time is reduced by a factor of at least two, which may be very important in real-time application incorporating a RANSAC step. It has been shown and experimentally verified that the number of local optimization steps is lower than logarithm of the number of samples drawn, and thus local optimization does not slow the procedure down. Three applications demonstrating the properties of LO-RANSAC framework were presented.

First, LO-RANSAC-RD, an algorithm for joint estimation of the epipolar geometry and radial distortion was presented. We showed, that the algorithm: 1. has the same complexity as the 7pt RANSAC, i.e. $\mathcal{O}(\varepsilon^{-7})$, 2. produces more inliers than the 7pt LO-RANSAC and hence can be terminated earlier, and 3. is more stable than the 9pt RANSAC (both the number of detected inliers and the estimated parameter of radial distortion have smaller variance).

Second, 3LAF LO-RANSAC – a new algorithm for the correspondence problem – was described. Exploiting output of the processes proposed in [36] for computation of affine-invariant local frames, three point-to-point correspondences were found for each region-to-region correspondence and used in epipolar geometry estimation. We have experimentally shown that: 1. 3LAF LO-RANSAC estimates epipolar geometry in time that is orders of magnitude faster than the standard method, 2. that the precision of the 3LAF LO-RANSAC and the standard method are comparable, and 3. that RANSAC without the LO step applied to triplets of points from a single region is significantly less precise than the new 3LAF LO-RANSAC algorithm. The presented matching method is pushing the limit of solvable problems, allowing EG estimation in correspondence problems with the ratio of inliers below 10%.

Finally, the combination of the previous two algorithms was tested. The 3LAF LO-RANSAC-RD algorithm has advantages of both LO-RANSAC-RD and 3LAF LO-RANSAC. The simultaneous estimation of EG and RD increases the precision and the number of correct matches, the time complexity is reduced to $\mathcal{O}(\varepsilon^{-3})$. This compares favourably with the $\mathcal{O}(\varepsilon^{-9})$ complexity of the state-of-the-art.

6 Degenerate Configurations

The topic of this chapter is robust estimation of epipolar geometry (EG) from image point correspondences in the possible presence of a dominant scene plane. A novel RANSAC-based algorithm is presented that handles in a unified manner the following three classes of scenes:

1. all points belong to a single scene plane,
2. majority of points belong to a dominant plane and the rest is off the plane,
3. minority or no points lie in a scene plane (a general scene).

In the first case, only a plane homography is computed, in the other cases, a correct EG is computed. It need not be known *a priori* which class the input scene belongs to, and still the computations are not slower than in EG estimation by plain RANSAC for a general scene.

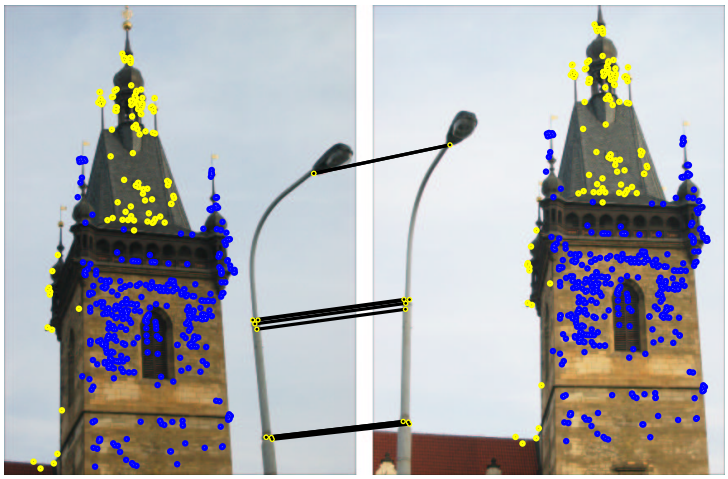


Figure 6.1.: The LAMPPOST scene with 97% of correct tentative correspondences lying in or near a dominant plane. In 100 runs, RANSAC fails to find a single inlier on the lamp 83 times; in the remaining 17, no more than 4 out of the 10 correspondences on the lamppost are found. Points on the lamppost are far from the dominant plane and therefore critically influence the precision of epipolar geometry and egomotion estimation. The DEGENSAC algorithm, with the same computational complexity as RANSAC, found the 10 lamppost inliers in all runs. Corresponding points lying in the dominant plane are dark, off-the-plane points are light, and the points on the lamp are highlighted by line segments.

The algorithm is based on the following theorem from two-view geometry, proved in the chapter. If five or more correspondences in a sample of seven correspondences are related by a homography (we refer to such a sample as H-*degenerate*), there always exists an EG consistent with both the seven correspondences and all correspondences related by the homography,

including those not in the seven-tuple (that is, not used in the EG estimation). For example, given five correspondences related by a homography and two other *arbitrary* correspondences, there is an EG consistent with all correspondences related by the homography and with the two other correspondences. H-degeneracy is not necessarily indicated by ill-conditioned fundamental matrix estimation in the seven-point algorithm.

In robust EG estimation, it is standard practice to measure the quality of the EG by the number of inliers [22] or some closely related quantity such as robust likelihood [57]. It has been observed that in scenes where most points lie in a plane, standard EG estimation algorithms often return an EG with a high number of inliers that is however totally incorrect. This behavior is explained by the proved geometric theorem, since a high inlier count can be obtained even if the seven-point sample includes two outliers. Such a number of inliers can cause the termination of RANSAC before a *non-degenerate all-inlier* sample is drawn. In general, the problem cannot be overcome by ex-post local optimization such as bundle adjustment, since the solution does not include off the plane correspondences, and therefore is likely to converge to a wrong local optimum.

Exploiting properties of H-degenerate samples, we also present a RANSAC-based algorithm for homography estimation that draws non-minimal samples. Contrary to the common practice of drawing minimal samples in RANSAC, the expected running time of the algorithm is lower than if minimal samples of four correspondences are drawn. This rather counter-intuitive result is explained by statistical analysis of the algorithm.

The planar degeneracy problem has been addressed before. In [55], different criteria for model selection are studied. The PLUNDER algorithm, which estimates multiple models separately and then performs model selection, was proposed in [58]. A method for correct feature tracking in a video sequence containing degenerate subsequences is designed in [56]. The method was extended in [44], where complete metric reconstruction is obtained. The proposed algorithm is novel since, unlike the previous approaches, it does not separately search for the two models (homography, epipolar geometry). Instead, *it detects both competing models simultaneously*, and the detection is *not slower* than direct estimation of the correct (but unknown) model.

The rest of the chapter is organized as follows. In section 6.1, degenerate samples are classified by the number of correspondences consistent with a homography, and properties of each class are discussed. The H-degeneracy test is developed in section 6.2. A novel DEGENSAC algorithm for EG estimation unaffected by a dominant plane is presented (section 6.3). Section 6.3.3 shows how homography can be estimated efficiently exploiting H-degeneracy. The performance of DEGENSAC is evaluated on two views of a scene with a dominant plane (section 6.4).

6.1. H-degenerate Configurations

Corresponding point pairs $(\mathbf{x}_1, \mathbf{x}'_1), \ldots, (\mathbf{x}_n, \mathbf{x}'_n)$ from two images taken by pinhole cameras satisfy the EG constraint [22]

$$(\forall i = 1, \ldots, n) \quad \mathbf{x}_i'^{\top} \mathtt{F} \mathbf{x}_i = 0, \tag{6.1}$$

where the 3×3 fundamental matrix F has rank 2.

Up to scale, there are 3 fundamental matrices consistent with $n = 7$ correspondences in general position, of which 2 may be complex conjugates. They are computed by the 7 point algorithm [22] as follows. Matrices $\mathtt{F}$ satisfying seven equations (6.1) form a 2-dimensional linear space, $\mathtt{F} \sim \lambda\mathtt{F}_1 + (1-\lambda)\mathtt{F}_2$. Requiring $\mathtt{F}$ to have rank 2 imposes an additional cubic constraint

$$|\lambda\mathtt{F}_1 + (1-\lambda)\mathtt{F}_2| = 0. \tag{6.2}$$

A homography, represented by a 3×3 matrix $\mathtt{H}$, is said to be consistent with EG represented by $\mathtt{F}$ if all point pairs satisfying $\mathbf{x}' \sim \mathtt{H}\mathbf{x}$ simultaneously satisfy $\mathbf{x}'^{\top}\mathtt{F}\mathbf{x} = 0$. This happens if and only if matrix $\mathtt{H}^{\top}\mathtt{F}$ is skew-symmetric [22].

The contribution of the chapter is based on the following fact from geometry of two uncalibrated images. Let a set X of n correspondences be related by a unique homography, but otherwise in general position. If $n = 7$ or $n = 6$, then any EG consistent with X is also consistent with the homography. If $n = 5$, one of the three EGs defined by X and two more correspondences is consistent with the homography. A configuration of 7 correspondences, of which 5, 6 or 7 are homography related, is called an $\mathtt{H}$-*degenerate* configuration[1].

To prove this, we start with two rather general theorems.

Theorem 1 *Let $\mathbf{x}_1, \ldots, \mathbf{x}_4$ be points in a plane. The locus of such points $\mathbf{e}$ that the cross-ratio of the line pencil $(\mathbf{e} \times \mathbf{x}_1, \ldots, \mathbf{e} \times \mathbf{x}_4)$ is constant is a conic passing through $\mathbf{x}_1, \ldots, \mathbf{x}_4$.*

Proof. This is a well-known theorem from the projective geometry, sometimes called Chasles' theorem [13]. □

Theorem 2 *Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be 5 or more points in a plane, no three of them collinear. Let $\mathtt{H}$ be a non-singular homography. The set of epipole pairs consistent with correspondences $\{(\mathbf{x}_1, \mathtt{H}\mathbf{x}_1), \ldots, (\mathbf{x}_n, \mathtt{H}\mathbf{x}_n)\}$ is given by $E = E^p \cup E^c$ where*

$$E^p = \{\, (\mathbf{e}, \mathtt{H}\mathbf{e}) \mid \mathbf{e} \textit{ is arbitrary} \,\},$$
$$E^c = \{\, (\mathbf{e}, \mathbf{e}') \mid \textit{points } \mathtt{H}\mathbf{x}_1, \ldots, \mathtt{H}\mathbf{x}_n, \mathtt{H}\mathbf{e}, \mathbf{e}' \textit{ are conconic} \,\}.$$

Proof. Assume without loss of generality $\mathtt{H} = \mathtt{I}$, that is, $\mathbf{x}'_i \sim \mathbf{x}_i$. Then the theorem is a trivial consequence of what is called Chasles' theorem in [22, Theorem 22.3]. However, since Chasles' theorem is usually stated only for four points (Theorem 1 in this chapter), we will give an explicit proof.

An epipole pair $(\mathbf{e}, \mathbf{e}')$ is consistent with the correspondences if line pencil $\mathbf{e}(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ is projectively related to line pencil $\mathbf{e}'(\mathbf{x}_1, \ldots, \mathbf{x}_n)$. That is, the pencil joining $\mathbf{e}$ with any four points of $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ must have the same cross-ratio as the pencil joining $\mathbf{e}'$ with the same four points.

Let us choose $\mathbf{e}$ arbitrarily. By Theorem 1, $\mathbf{e}'$ must lie on the conic passing through points $\{\mathbf{x}_1, \ldots, \mathbf{x}_4, \mathbf{e}\}$ and, at the same time, on the conic passing through $\{\mathbf{x}_2, \ldots, \mathbf{x}_5, \mathbf{e}\}$. If these two conics are different, their common points are $\{\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{e}\}$. There cannot be more common

[1] $\mathtt{H}$-degeneracy should not be confused with the term *degenerate* configuration, which we use in the usual meaning to denote a configuration of 7 correspondences consistent with an infinite class of EGs.

points since two different non-singular conics have at most four intersections. By induction on other 4-tuples of $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$, we obtain that either $\mathbf{e}' \sim \mathbf{e}$ or, if points $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ are conconic, that $\mathbf{e}$ and $\mathbf{e}'$ both lie on this common conic. □

Theorem 2 states that a set of 6 or more homography-related point pairs not lying on a single conic is consistent with the same class of EGs as is the whole homography. However, if the points are conconic (which is always the case for 5 pairs), there is another class of EGs besides the one consistent with the homography, namely the class defined by epipoles lying anywhere on the respective conics.

In the theorem, the former class is denoted by E^p. This stands for the *planar part* of E, since given $(\mathbf{e}, \mathbf{e}') \in E^p$, 3D reconstruction of the correspondences lies in a plane. The latter part is denoted by E^c. This stands for *cubic part* of E since the 3D reconstruction of the correspondences lies on a twisted cubic along with the two camera centers. The cubic part is closely related to the horopter theorem [22].

If the points are conconic, it can be shown that set E corresponds to the set of fundamental matrices given by

$$\mathcal{F} = \{\, \mathtt{F} \mid \mathtt{F} = \mathtt{H}^{-\top}([\mathbf{v}]_\times + \lambda \mathtt{Q}),\ |\mathtt{F}| = 0,\ \mathbf{v} \in \mathbb{R}^3,\ \lambda \in \mathbb{R} \,\},$$

where matrix $\mathtt{Q}$ represents the Steiner conic [22, Section 9.4] passing through the points. Taking $\lambda = 0$ yields the planar part E^p, whereas $\lambda \neq 0$ yields the cubic part E^c.

The rest of this section applies Theorem 2 to prove the above facts about H-degenerate configurations.

6.1.1. All 7 Pairs Related by Homography

If all 7 out of 7 correspondences are related by a homography, the class of EGs consistent with them is the same as the class consistent with the homography. The EGs are parametrized by the position of one epipole.

Theorem 3 *Let points $\mathbf{x}_1, \ldots, \mathbf{x}_7$ contain no collinear triplet and not lie on a conic. Let $\mathtt{H}$ be a non-singular homography. Then all fundamental matrices consistent with correspondences $\{(\mathbf{x}_1, \mathtt{H}\mathbf{x}_1), \ldots, (\mathbf{x}_7, \mathtt{H}\mathbf{x}_7)\}$ are consistent with $\mathtt{H}$. In detail, these fundamental matrices form a set*

$$\mathcal{F}_7 = \{\, [\mathbf{e}']_\times \mathtt{H} \mid \mathbf{e}' \in \mathbb{R}^3 \,\}.$$

Proof. Straightforward by Theorem 2. Since the points are not conconic, it is $E^c = \emptyset$ and $E = E^p$. □

For this configuration, the linear space of matrices $\mathtt{F}$ satisfying (6.1) has dimension 3 (rather than 2 as for a 7-tuple in a general position) and all these matrices have rank 2. It follows that $\mathcal{F}_7$ is a 3-dimensional linear space.

6.1.2. 6 of 7 Pairs Related by Homography

If 6 of 7 pairs are related by a homography, the class of EGs consistent with them is the class consistent with the homography and one additional pair off the homography. The class is parametrized by the position of one epipole which has to be collinear with the last pair.

Theorem 4 *Let points $\mathbf{x}_1, \ldots, \mathbf{x}_6$ contain no collinear triplet and not lie on a conic. Let $\mathtt{H}$ be a non-singular homography. All fundamental matrices consistent with correspondences $\{(\mathbf{x}_1, \mathtt{H}\mathbf{x}_1), \ldots, (\mathbf{x}_6, \mathtt{H}\mathbf{x}_6)\}$ are consistent with $\mathtt{H}$. In detail, for any $(\mathbf{x}_7, \mathbf{x}'_7)$, fundamental matrices consistent with correspondences $\{(\mathbf{x}_1, \mathtt{H}\mathbf{x}_1), \ldots, (\mathbf{x}_6, \mathtt{H}\mathbf{x}_6), (\mathbf{x}_7, \mathbf{x}'_7)\}$ form set*

$$\mathcal{F}_6 = \{\, [\mathbf{e}']_\times \mathtt{H} \mid \mathbf{e}' \in \mathbb{R}^3,\ \mathbf{e}'^\top (\mathtt{H}\mathbf{x}_7 \times \mathbf{x}'_7) = 0 \,\}.$$

Proof. Straightforward by applying Theorem 2 on the first 6 points. This yields $E = E^p$. Part of E^p consistent with $(\mathbf{x}_7, \mathbf{x}'_7)$ contains pairs $(\mathbf{e}, \mathtt{H}\mathbf{e})$ for $\mathbf{e}$ being collinear with the 7th point pair. □

For this configuration, the linear space of matrices $\mathtt{F}$ satisfying (6.1) has dimension 2. However, unlike for a 7-tuple in a general position, all matrices in this space have rank 2. In other words, coefficients in the polynomial (6.2) are identically zero. Therefore, $\mathcal{F}_6$ is a 2-dimensional linear space.

6.1.3. 5 of 7 Pairs Related by Homography

The class of EGs consistent with the configuration of 5 homography related pairs and 2 more pairs off the homography always contains an EG consistent with the homography. This follows straightforwardly from the plane-and-parallax algorithm [26, 22], which linearly computes EG from a homography and additional 2 point pairs off the homography.

Not so obviously, this configuration is non-degenerate in the usual sense because, like a general configuration of 7 points, it yields 3 solutions for EG. In contrast, the configurations in the previous two subsections are degenerate.

By Theorem 2, the class of EGs consistent with this configuration is the same as the class consistent with a point-wise corresponding pair of conics and two additional point pairs off the homography fixed by the conics.

Theorem 6 bellow will summarize these facts. Before stating it, we give another theorem needed for its proof. It defines the class of EGs consistent with two point-wise corresponding conics and a single additional point pair (the theorem assumes $\mathtt{H} = \mathtt{I}$, hence only one conic is mentioned).

Theorem 5 *Let $\mathtt{Q}$ be a conic and $\mathbf{x}$ and $\mathbf{x}'$ two points not on $\mathtt{Q}$. Let $(\mathbf{e}, \mathbf{e}')$ be a pair of epipoles lying on $\mathtt{Q}$ and consistent with correspondences $\{\, (\mathbf{y}, \mathbf{y}) \mid \mathbf{y} \in \mathtt{Q} \,\} \cup (\mathbf{x}, \mathbf{x}')$. Then the line $\mathbf{e} \times \mathbf{e}'$ is tangent to the conic (see Figure 6.2)*

$$\mathtt{R} = (\mathbf{x}^\top \mathtt{Q} \mathbf{x})(\mathbf{x}'^\top \mathtt{Q} \mathbf{x}')\mathtt{Q} - |\mathtt{Q}|(\mathbf{x} \times \mathbf{x}')(\mathbf{x} \times \mathbf{x}')^\top. \tag{6.3}$$

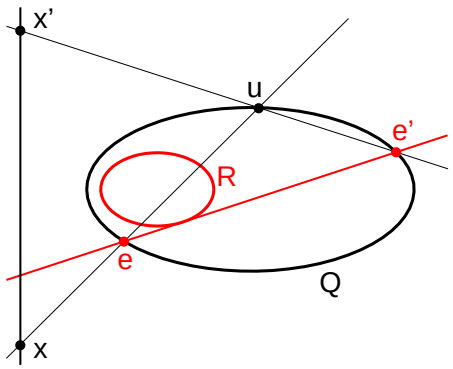


Figure 6.2.: Conic R is the envelope of lines joining epipole pairs consistent with conic Q and the point pair $(\mathbf{x}, \mathbf{x}')$.

Proof. A pair of epipoles lying on Q and consistent with $\{\,(\mathbf{y},\mathbf{y}) \mid \mathbf{y} \in \mathtt{Q}\,\} \cup (\mathbf{x},\mathbf{x}')$ has the property that the point $\mathbf{u} = (\mathbf{x} \times \mathbf{e}) \times (\mathbf{x}' \times \mathbf{e}')$ lies on Q. Thus, all such epipole pairs are parameterized by a point $\mathbf{u}$ moving on Q.

Given a point $\mathbf{u}$ on Q, the second intersection of the line $\mathbf{u} \times \mathbf{x}$ with Q is $\mathbf{e} = \mathtt{A}\mathbf{u}$ where $\mathtt{A} = \mathtt{I} - 2\mathbf{x}\mathbf{x}^\top\mathtt{Q}/(\mathbf{x}^\top\mathtt{Q}\mathbf{x})$ is the harmonic homology [22] with the vertex $\mathbf{x}$ and the axis $\mathtt{Q}\mathbf{x}$. Similarly, $\mathbf{e}' = \mathtt{A}'\mathbf{u}$ where $\mathtt{A}' = \mathtt{I} - 2\mathbf{x}'\mathbf{x}'^\top\mathtt{Q}/(\mathbf{x}'^\top\mathtt{Q}\mathbf{x}')$.

If Q is real non-degenerate, we can without loss of generality assume $\mathtt{Q} = \mathrm{diag}(1,1,-1)$ and $\mathbf{u} = [\cos t, \sin t, 1]^\top$. It can be verified that the expression $(\mathbf{e} \times \mathbf{e}')^\top \mathtt{R}^* (\mathbf{e} \times \mathbf{e}')$ vanishes identically. Hence the line $\mathbf{e} \times \mathbf{e}'$ is tangent to R for any $\mathbf{u}$ lying on Q. Here, $\mathtt{R}^*$ denotes the matrix of cofactors of R, representing the conic dual to R.

If the signature σ of Q is different from $(1,1,-1)$, meaning than Q is complex and/or degenerate, we proceed the same way using $\mathtt{Q} = \mathrm{diag}(\sigma)$ and the appropriate parameterization of $\mathbf{u}$ moving on Q. □

Alternatively, R is obtained by transforming Q by the homology [22] B with the axis $\mathbf{x} \times \mathbf{x}'$, the vertex $\mathtt{Q}\mathbf{x} \times \mathtt{Q}\mathbf{x}'$, and the characteristic ratio $\mu = (\mathbf{x}^\top\mathtt{Q}\mathbf{x}')/\sqrt{(\mathbf{x}^\top\mathtt{Q}\mathbf{x})(\mathbf{x}'^\top\mathtt{Q}\mathbf{x}')}$ being the projective invariant of a conic and two points [22, Exercise 2.10.2]. Unlike equation (6.3) however, this construction is indefinite if the line $\mathbf{x} \times \mathbf{x}'$ is tangent to Q.

Let $\mathbf{j}$ and $\mathbf{j}'$ denote the common points of Q and the line $\mathbf{x} \times \mathbf{x}'$, and let γ denote the cross-ratio $\langle \mathbf{x}, \mathbf{x}'; \mathbf{j}, \mathbf{j}' \rangle$. Then it can be shown that $4\mu^2 - 2 = \gamma + \gamma^{-1}$.

Assuming non-degeneracy of Q, conic R degenerates if either $(\mathbf{x}^\top\mathtt{Q}\mathbf{x})(\mathbf{x}'^\top\mathtt{Q}\mathbf{x}') = 0$ (in that case, R is the double line $\mathbf{x} \times \mathbf{x}'$), or $\mathbf{x}^\top\mathtt{Q}\mathbf{x}' = 0$ (then homology B is singular and R is the pair of tangents to Q from the point $\mathtt{Q}\mathbf{x} \times \mathtt{Q}\mathbf{x}'$).

Theorem 6 *Let no three of points $\mathbf{x}_1, \ldots, \mathbf{x}_5$ be collinear. Let* H *be a non-singular homography. Let $(\mathbf{x}_6, \mathbf{x}'_6)$ and $(\mathbf{x}_7, \mathbf{x}'_7)$ be two correspondences in general position w.r.t. the first five ones and* H*. There is s finite number of EGs consistent with correspondences $\{(\mathbf{x}_1, \mathtt{H}\mathbf{x}_1), \ldots, (\mathbf{x}_5, \mathtt{H}\mathbf{x}_5), (\mathbf{x}_6, \mathbf{x}'_6), (\mathbf{x}_7, \mathbf{x}'_7)\}$ of which one EG is consistent with* H*.*

Proof. Assume $\mathtt{H} = \mathtt{I}$. By Theorem 2, the first five correspondences restrict the epipole pairs to be in $E^p \cup E^c$ where $E^p = \{(\mathbf{e}, \mathbf{e}) \mid \mathbf{e} \text{ arbitrary}\}$ and $E^c = \{(\mathbf{e}, \mathbf{e}') \mid \mathbf{x}_1, \ldots, \mathbf{x}_5, \mathbf{e}, \mathbf{e}' \text{ conconic}\}$.

By the plane-and-parallax algorithm [22], the subset of E^p consistent with $\{(\mathbf{x}_6, \mathbf{x}'_6), (\mathbf{x}_7, \mathbf{x}'_7)\}$ is given by $\mathbf{e} = (\mathbf{x}_6 \times \mathbf{x}'_6) \times (\mathbf{x}_7 \times \mathbf{x}'_7)$.

Let $\mathtt{Q}$ be the conic through $\mathbf{x}_1, \ldots, \mathbf{x}_5$. By Theorem 5, the subset of E^c consistent with $(\mathbf{x}_6, \mathbf{x}'_6)$ is given by requiring the line $\mathbf{e} \times \mathbf{e}'$ to be tangent to the conic which we denote by $\mathtt{R}_6$. The similar conic for $(\mathbf{x}_7, \mathbf{x}'_7)$ is denoted by $\mathtt{R}_7$.

An epipole pair lying on $\mathtt{Q}$ and consistent with $\{(\mathbf{x}_6, \mathbf{x}'_6), (\mathbf{x}_7, \mathbf{x}'_7)\}$ must lie on a common tangent to $\mathtt{R}_6$ and $\mathtt{R}_7$. In general case, there is a finite number (up to four) of these tangents, hence there is a finite number of feasible epipole pairs and the configuration is non-degenerate.

Note, not every common tangent intersects $\mathtt{Q}$ in a feasible epipole pair. Some do in epipole pairs consistent with $\{(\mathbf{x}_6, \mathbf{x}'_6), (\mathbf{x}'_7, \mathbf{x}_7)\}$ rather than with $\{(\mathbf{x}_6, \mathbf{x}'_6), (\mathbf{x}_7, \mathbf{x}'_7)\}$. A feasible epipole pair must satisfy that the points $(\mathbf{e} \times \mathbf{x}_6) \times (\mathbf{e}' \times \mathbf{x}'_6)$ and $(\mathbf{e} \times \mathbf{x}_7) \times (\mathbf{e}' \times \mathbf{x}'_7)$ lie on $\mathtt{Q}$. □

6.2. Detection of H-degenerate Samples

This section describes an efficient test whether seven correspondences $\{(\mathbf{x}_i, \mathbf{x}'_i)\}_{i=1}^7$ are H-degenerate. The input is not only the seven correspondences but also a fundamental matrix F consistent with them (one of the F output by the 7-point algorithm). The test verifies whether there exist five correspondences related by a homography H which is compatible with F.

Given F, only three correspondences $\{(\mathbf{x}_i, \mathbf{x}'_i)\}_{i=1}^3$ are sufficient to compute a plane homography H as [22]

$$\mathtt{H} = \mathtt{A} - \mathbf{e}'(\mathtt{M}^{-1}\mathbf{b})^\top, \tag{6.4}$$

where $\mathtt{A} = [\mathbf{e}']_\times \mathtt{F}$ and $\mathbf{b}$ is a 3-vector with components

$$b_i = (\mathbf{x}'_i \times (\mathtt{A}\mathbf{x}_i))^\top (\mathbf{x}'_i \times \mathbf{e}') \|\mathbf{x}'_i \times \mathbf{e}'\|^{-2},$$

and $\mathtt{M}$ is a 3×3 matrix with rows $\mathbf{x}_i^\top$.

A triplet from each five-tuple defines a homography and the other four correspondences are checked for consistency with this homography. Any of $\binom{7}{5} = 21$ five-tuples contains at least one of the triplets $\{1, 2, 3\}$, $\{4, 5, 6\}$, $\{1, 2, 7\}$, $\{4, 5, 7\}$ and $\{3, 6, 7\}$. Hence, at most five homographies have to be tested.

In general, up to three fundamental matrices are consistent with the seven correspondences and the test should be carried for all of them. In practice, it is sufficient to check H-degeneracy of the F consistent with the largest number of tentative correspondences, for details see Section 6.3.

Note that the test is also applicable to samples containing six or seven H-related correspondences, as every fundamental matrix from the sets $\mathcal{F}_6$ and $\mathcal{F}_7$ is consistent with H.

6.2.1. The Stability of the Test

The test is derived for corresponding points that are exactly related by H. In real tasks, inliers are typically perturbed by noise. The fundamental matrix F estimated from seven correspondences

corrupted by noise is not precise and a search for a homography compatible with F may be too restrictive. The following experiment has been carried out to assess the stability of the test. Sets of seven correspondences, five correspondences related by H and two not related by H, were generated at random. The seven-point algorithm was used to produce up to three fundamental matrices. The fundamental matrix with the highest support was then checked by the H-degeneracy test. The experiment on the LAMPPOST scene Fig. 6.1 showed, that the test correctly recognizes about 75% of the H-degenerate samples. The following variations of the H-degeneracy test were also tested.

For each of the five homographies checked in the test, errors on the seven correspondences were computed. For every homography, a new homography was re-estimated from five correspondences with the lowest error. The new 'direct' homography is more precise, since it is computed directly from the correspondences, not involving F. Note, that the 'direct' homography is generally not compatible with the original fundamental matrix F. If the five correspondences are fit well by the 'direct' homography, the sample is labelled as a H-degenerate sample. This variation of the test correctly recognized over 99% of the H-degenerate samples.

As an alternative, the followng (faster) method was tested. The 'direct' homography is not estimated for each of the five tested homographies. Instead, only the homography with the smallest error on a correspondence with the fourth (fifth respectively) smallest error is re-estimated by a 'direct' homography. Such an approach correctly recognized about 97% of the H-degenerate samples.

6.3. Two-view Geometry Estimation Unaffected by a Dominant Plane

In this section, we show how the results on degeneracies of seven point correspondences (Section 6.1) can be exploited to design an EG estimation algorithm that is robust to the presence of a dominant plane in the scene. Epipolar geometry estimators of practical importance must be able to handle outliers among tentative correspondences. RANSAC, which achieves robustness by drawing independent samples repeatedly, is the most commonly used EG estimator and we focus on this algorithm.

Frequently, an incorrect EG is output by RANSAC if the dominant plane is present in the scene, i.e., if a large fraction of correct tentative correspondences are related by a homography. The behavior, demonstrated in Fig. 6.5, is not difficult to explain. In standard RANSAC [15], it is assumed that once an all-inlier sample is drawn, the correct model is recovered. However, this is a necessary but not a sufficient condition. To estimate the EG correctly, the sample from which the EG is estimated must also be non-degenerate, i.e. at most five of its correspondences may be related by a single homography (as shown in Sections 6.2).

Let us analyze the RANSAC stopping criterion in detail. Let $P_{\mathtt{F}}$ be the probability that a *good sample* is drawn, i.e. a sample that enables estimation of the correct EG. The probability η that RANSAC failed to find the correct solution after k independent draws, i.e. no *good* sample has been drawn, is

$$\eta(P_{\mathtt{F}}) = (1 - P_{\mathtt{F}})^k. \tag{6.5}$$

In standard RANSAC, an all-inlier sample is assumed to be a *good* sample. The probability of drawing an all-inlier (seven-tuple containing seven inliers) is $P_{7/7} = \varepsilon^7$, where ε is the fraction of inliers in the set of tentative correspondences. However, for the case of a dominant plane, the probabilities $P_{7/7}$ and $P_{\mathtt{F}}$ differ. The correct EG can only be estimated from an all-inlier sample with no more than five correspondences from the plane and 2 or more inliers off the plane. Let $\varepsilon_{\mathtt{H}}$ be the fraction of homography consistent correspondences. Then the probability $P_{\mathtt{F}}$ of drawing a *good* sample is

$$P_{\mathtt{F}} = \sum_{i=0}^{5} \binom{7}{i} \varepsilon_{\mathtt{H}}^{i} (\varepsilon - \varepsilon_{\mathtt{H}})^{7-i}. \tag{6.6}$$

Note that $P_{\mathtt{F}} = P_{7/7}$ if $\varepsilon_{\mathtt{H}} = 0$, and $P_{\mathtt{F}} < P_{7/7}$ otherwise. If $\varepsilon_{\mathtt{H}}$ were known (and it never is), the probability $P_{\mathtt{F}}$ could be used to calculate the confidence $1 - \eta(P_{\mathtt{F}})$. Such a stopping rule would be very inefficient, as the $P_{\mathtt{F}}$ drops to zero when $\varepsilon_{\mathtt{H}}$ approaching ε.

6.3.1. The DEGENSAC Algorithm

The DEGENSAC algorithm that correctly estimates epipolar geometry even for high values of $\varepsilon_{\mathtt{H}}$ is introduced in this section. The first two steps of the algorithm, summarized in Alg. 5, are identical to standard RANSAC [22, p.291]. The algorithm repeatedly draws samples $\mathcal{S}_k$ of seven correspondences. The seven-point algorithm produces up to three fundamental matrices. Let $\mathtt{F}_k$ be the fundamental matrix with the largest support I_k of the fundamental matrices calculated from $\mathcal{S}_k$. If $\mathtt{F}_k$ is the best so far ($I_k > I_j, j = 1 \dots k-1$) the fundamental matrix and the size of its support is stored.

When the best sample so far is drawn, the H-degeneracy test involving $\mathtt{F}_k$ and the seven-tuple $\mathcal{S}_k$ is performed (step 4) as described in Section 6.2. If no five correspondences are H-related, the inliers are considered to be in general position and the algorithm continues with the step 1. Otherwise, an H-degenerate sample is detected which means that a homography $\mathtt{H}_k$ consistent with at least five correspondences from $\mathcal{S}_k$ exists. Using $\mathtt{H}_k$, EG is estimated by the plane-and-parallax algorithm (step 6). In this case, $\mathtt{H}_k$ and two additional correspondences are sufficient to define an EG. The EG with the largest support found in step 6 is stored and the algorithm proceeds with the next iteration.

6.3.2. DEGENSAC – the Probability of Success

There are two cases how the correct EG is recovered; either from an all-inlier non-degenerate sample, or by the plane-and-parallax algorithm after a homography is detected. The former case occurs with probability $P_{\mathtt{F}}$, see (6.6). The latter case happens when a sample containing five or more H-related correspondences is drawn. The probability $P_{5/7}$ of drawing an H-degenerate sample is

$$P_{5/7} = \sum_{i=5}^{7} \binom{7}{i} \varepsilon_{\mathtt{H}}^{i} (1 - \varepsilon_{\mathtt{H}})^{7-i}. \tag{6.7}$$

The sample of five H-related and two EG consistent but not H-related correspondences is both a *good* and H-degenerate sample. The probability of drawing such a sample, $21\varepsilon_{\mathtt{H}}^{5}(\varepsilon - \varepsilon_{\mathtt{H}})^2$, is in-

Input: The set of N point-to-point correspondences.
Output: The fundamental matrix $\mathtt{F}$ with the largest support; optionally the homography $\mathtt{H}$ with the largest support.

Set the lower bound on the number of inliers $I^* := 0$.

Repeat until the probability $\eta(P_{7/7}) = (1 - (I^*/N)^7)^k$ of finding EG with support larger than I^* in k-th step falls under threshold η_0:

1 Select a random sample of 7 correspondences, calculate up to three fundamental matrices consistent with the sample. Compute the support for each of them. Let $\mathtt{F}_k$ be the one with the largest support (size I_k).
2 If $I_k > I^*$:

Novelty —

3 Store $\mathtt{F}_k$ and let $I^* := I_k$.
4 Evaluate the H-degenerate sample test (see Section 6.2). If 5 or more correspondences from the sample consistent with homography $\mathtt{H}_k$ are found:
5 Compute the size $I_k^{\mathtt{H}}$ of the support of $\mathtt{H}_k$. Store $\mathtt{H}_k$ if it has the biggest support between homographies so far.
6 Execute robust estimation of fundamental matrix $\mathtt{F}_k^{\mathtt{H}}$ based on plane-and-parallax algorithm and calculate its support size $I_k^{\mathtt{F}}$.
7 If $I_k^{\mathtt{F}} > I_k$ store $\mathtt{F}_k^{\mathtt{H}}$ and let $I^* := I_k^{\mathtt{F}}$.

Algorithm 5: The structure of the DEGENSAC algorithm.

cluded in both probabilities $P_{\mathtt{F}}$ and $P_{5/7}$. Therefore, following the inclusion-exclusion principle, the probability that the correct EG is recovered by drawing a single sample is

$$P = P_{\mathtt{F}} + P_{5/7} - 21\varepsilon_{\mathtt{H}}^5(\varepsilon - \varepsilon_{\mathtt{H}})^2. \tag{6.8}$$

Note, that $P \geq P_{7/7}$ as an all-inliers sample is either *good* or H-degenerate sample, or both.

Probabilities $P_{\mathtt{F}}$, $P_{5/7}$ and P are plotted in Fig. 6.3 (top). As $P_{\mathtt{F}}$ and $P_{5/7}$ are functions of both ε and $\varepsilon_{\mathtt{H}}$, $\varepsilon = 0.67$ was fixed in the figure to keep the plot one-dimensional. The figure shows, that the probability of drawing a *good* sample is almost unchanged up to approximately 30% of inliers lying on a single plane. The probability drops off for higher values of $\varepsilon_{\mathtt{H}}/\varepsilon$. For more than 52% of coplanar inliers, it becomes more likely to draw an H-degenerate sample than a *good* sample. Note that $P \geq \max(P_{\mathtt{F}}, P_{5/7})$. Qualitatively, the dependency is the same for all values of ε. The $\log(P_{5/7}/P_{\mathtt{F}})$ function for different ε is drawn in Fig. 6.3-bottom. The plot shows when it is more likely to draw a non-degenerate all-inliers sample ($\log(P_{5/7}/P_{\mathtt{F}}) < 0$) or H-degenerate sample.

The termination criterion. The two algorithms, DEGENSAC and RANSAC, have equal probability of finding a correct solution after the same number of steps if $\varepsilon_{\mathtt{H}} = 0$. With increasing $\varepsilon_{\mathtt{H}}$, the probability of success is increasing for DEGENSAC and decreasing for RANSAC, i.e. $\varepsilon_{\mathtt{H}} = 0$ is the worst case for DEGENSAC. To ensure the required confidence in the solution even in the

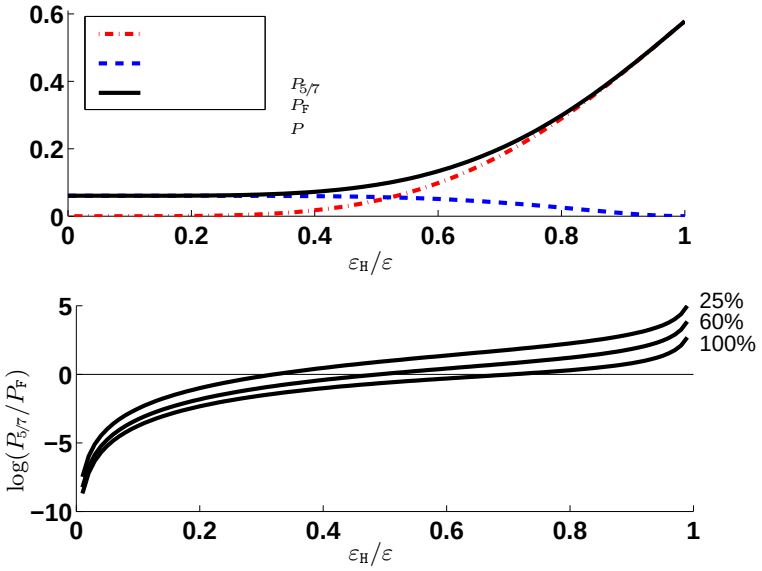


Figure 6.3.: Probabilities $P_{5/7}$, $P_{\mathtt{F}}$ and P (eqs. (6.6-6.8)), for inlier percentage $\varepsilon = 67\%$ (top) and $\log(P_{5/7}/P_{\mathtt{F}})$ for values 25%, 60% and 100% of ε respectively (bottom).

worst case, DEGENSAC is terminated when $\eta(P_{7/7})$ falls under predefined threshold (typically 5%).

Computational complexity. The computations carried out in DEGENSAC and RANSAC are identical with one exception – when the best sample so far is drawn. Note that other than the best samples so far are of no interest, as EG fitting as many correspondences as possible is sought for. It is not difficult to show [7], that the best sample so far occurs on average only $\log(k)$ times, where k is the number of samples drawn. Thus, the procedure is executed rarely and its time complexity is not critical.

If the best sample so far is drawn, the H-degeneracy test is performed (step 4). It takes constant (and in practice insignificant) time to evaluate the test. A more complex procedure (step 6), based on the plane-and-parallax algorithm, is executed if the test detects an H-degenerate sample. In this step, another RANSAC-like strategy that draws samples of two correspondences not related by the detected homography $\mathtt{H}_k$ is performed. Note that the time complexity of algorithm searching for a two inlier sample is significantly lower then the complexity of the main loop.

Homography or epipolar geometry? Besides returning the EG, the DEGENSAC algorithm also outputs the homography with the largest support. The EG is valid unless a planar scene is observed or if there is a zero translation of the camera center. Some model selection strategy [55] can be applied to check the validity of the EG. The dominant plane homography may or may not be of direct interest. It can be used e.g. as a strong geometric constraint in guided matching [22].

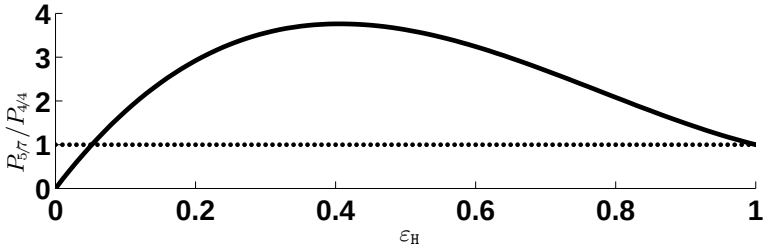


Figure 6.4.: The ratio of probabilities of drawing an H-degenerate sample and an all-inlier (four out of four) sample ($P_{5/7}/P_{4/4}$).

6.3.3. Homography Estimation through EG

Consider the problem of homography estimation from data containing outliers. The standard approach (of RANSAC) is to draw samples of four correspondences to generate a hypotheses of homography. Surprisingly, a modified DEGENSAC is almost always able to estimate the homography in a smaller number of iterations.

Instead of drawing minimal samples of four correspondences and estimating homography directly, samples of seven correspondences are drawn and epipolar geometry is estimated instead. If the support of the EG is the best so far, then the H-degenerate sample test is carried out to recover the homography. In other words, the DEGENSAC algorithm is run and the plane-and-parallax step is not executed and the estimated EG is ignored. The probability of drawing an H-degenerate sample is $P_{5/7}$ and therefore the algorithm is terminated when the confidence $\eta(P_{5/7})$ falls under a predefined threshold.

Let us compare the probability that the solution is found in a single sample by standard RANSAC, $P_{4/4} = \varepsilon_{\mathtt{H}}^4$, and by the DEGENSAC based algorithm $P_{5/7}$ (6.7). Fig. 6.4 shows the ratio of the two probabilities $P_{5/7}/P_{4/4}$. For $\varepsilon_{\mathtt{H}} \in (0.052, 1)$ the DEGENSAC method finds the solution in less iterations than the standard RANSAC! Note, that the range of values of $\varepsilon_{\mathtt{H}}$ where $P_{5/7}$ is bigger than $P_{4/4}$ covers all practical situations.

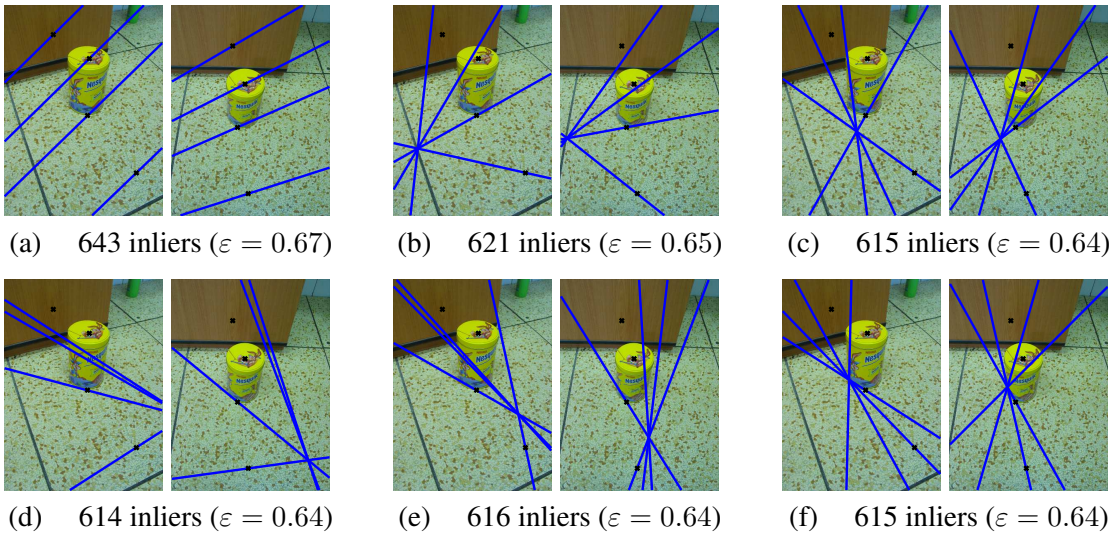
However, the estimation of homography through EG is unlikely to bring a significant speed-up[2]. Nevertheless, this is an interesting situations where a model of lower dimension can be estimated indirectly through an estimation of a model of higher dimension.

6.4. Experiments

Properties of EG estimation by the DEGENSAC and RANSAC algorithms are demonstrated on the BOX (Fig. 6.5) and the LAMPPOST (Fig. 6.1) scenes.

In the BOX experiment, both algorithms processed 958 tentative correspondences established by wide-baseline matching [35]. The set of correspondences includes 643 inliers ($\varepsilon = 67\%$) and 613 correct matches related by a homography ($\varepsilon_{\mathtt{H}} = 64\%$). The fraction of inliers that lie

[2]There are up to three fundamental matrices defined by seven points and hence the decrease in the number of samples will be offset by increased number of verified fundamental matrices.



(a) 643 inliers ($\varepsilon = 0.67$) (b) 621 inliers ($\varepsilon = 0.65$) (c) 615 inliers ($\varepsilon = 0.64$)

(d) 614 inliers ($\varepsilon = 0.64$) (e) 616 inliers ($\varepsilon = 0.64$) (f) 615 inliers ($\varepsilon = 0.64$)

Figure 6.5.: The BOX experiment. Examples of three EGs estimated during hundred executions of standard RANSAC (b-f); all 100 EGs estimated by DEGENSAC (a) were qualitatively the same. Epipolar lines $\mathtt{F}^\top \mathbf{x}'$ and $\mathtt{F}\mathbf{x}$ for four ground truth correspondences, two off the plane and two in the plane, are superimposed over the images. The EGs on (b-d) are consistent with the floor plane, but not with the correspondences off the plane.

on a plane is $\varepsilon_{\mathtt{H}}/\varepsilon = 95\%$. For such values of ε and $\varepsilon_{\mathtt{H}}$, the probabilities $P_{\mathtt{F}}$ of drawing a good sample (6.6) and a H-degenerate sample $P_{5/7}$ (6.7) are $P_{\mathtt{F}} = 0.003$ and $P_{5/7} = 0.5$ respectively.

Both the RANSAC and DEGENSAC algorithms, were executed one hundred times. The percentage of runs where a given correspondence was labeled as an inlier is plotted in Fig. 6.6, for both RANSAC (top) and DEGENSAC (bottom). The same correspondence occupies the same column in both plots and the correspondences are sorted according to the DEGENSAC results.

In each of the hundred executions, DEGENSAC detected the same set of 613 H-related correspondences. From the set of the remaining 345 correspondences, 329 were classified constantly in every single execution of DEGENSAC: 29 as inliers and 300 as outliers. This demonstrates the stability of the estimated EG. The remaining 16 correspondences consist of 8 correct matches and 8 mismatches that lie on the boundary of the error threshold. On average, of the 16 borderline correspondences, 4 correct matches and 2.8 mismatches were labeled as inliers per execution of DEGENSAC.

On the other hand, RANSAC often failed to find correct off-plane matches and returned an incorrect EG defined by in-plane matches and some random mismatches. To visualize the resulting epipolar geometries, four "ground truth" correspondences were established manually ex-post, two off the plane and two in the plane[3].

The EG detected by DEGENSAC is shown in Fig. 6.5a. All four "ground truth" correspondences, both in and off the plane, satisfy the epipolar constraint. Examples of EGs estimated

[3]The manually established correspondences were, of course, not part of the input to RANSAC and DEGENSAC.

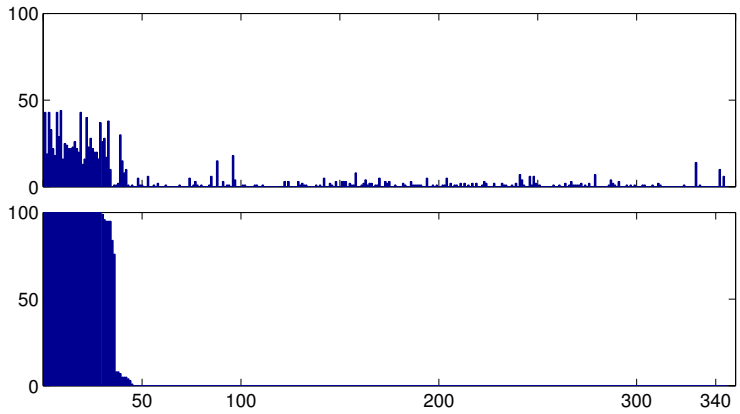


Figure 6.6.: Each column in the graph represents a correspondence off the plane. The bars show the number of runs when the correspondence was labeled as inlier during 100 repetitions of RANSAC (top) and DEGENSAC (bottom). Note the bimodality of the DEGENSAC output.

by RANSAC are depicted in Fig. 6.5b-d. Note, that even though the number of correspondences labeled as inliers in Fig. 6.5b-d are non random and close to the number of inliers in Fig. 6.5a, the resulting EG is incorrect. Only the in-plane correspondences satisfy the epipolar constraint.

Similar results were obtained for the LAMPPOST scene Fig. 6.1. In this experiment, the fraction of inliers was $\varepsilon = 86\%$ and the fraction of in-plane correspondences was $\varepsilon_{\mathtt{H}} = 65\%$; $\varepsilon_{\mathtt{H}}/\varepsilon = 76\%$. Most of the off-plane inliers lie close to the dominant frontal plane (correspondences on the roof). Together, 97% of correspondences were in or near the plane. Ten inliers most distant from the plane, critical for the well-posedness of the solution, are located on the lamppost. In all of the one hundred DEGENSAC executions, all ten correspondences on the lamppost were labelled as inliers. RANSAC selected as inlier none of them in 83 executions. In the remaining 17 executions, RANSAC labelled no more than 4 of the correspondences as inliers.

6.4.1. Limitations of the Method

The DEGENSAC algorithm alleviates the problems caused by the presence of a dominant plane in the scene. There are some more treacherous configurations, that are not handled by the algorithm. Consider a plane with a repeated pattern whose elements repeat along lines that intersect in a single (finite or infinite) point $\mathbf{E}$. In two views, let the plane induce a homography $\mathtt{H}$ and let the projection of $\mathbf{E}$ into the second image be $\mathbf{e}'$.

The automatic matching method will match some of the features on the plane correctly and some incorrectly to another appearance of the feature in the repeated pattern. All correct matches from the plane are linked by the homography $\mathtt{H}$. All mismatched points $\mathbf{x}'$ in the second image (mismatched due to the repeated structure) lie on a line $[\mathbf{e}']_\times \mathtt{H}\mathbf{x}$. Therefore, the epipolar geometry represented by a fundamental matrix $\mathtt{F} = [\mathbf{e}']_\times \mathtt{H}$ is consistent with all matches, both good and mismatches, from the plane. The fundamental matrix $\mathtt{F}$ has possibly large support on the plane (i.e. linked by the homography $\mathtt{H}$) and off the plane (i.e. not linked by the homography $\mathtt{H}$).

from a set of correspondences alone, unless some constraints are imposed on the 3D scene. More information, typically available in the earlier stages of matching, have to be exploited. Two approaches to overcome the repeated pattern problem follow.

Possible solutions. An intra-image matching can be done in the images in order to detect repeated structures along lines that intersect in a single point. Epipolar geometry with an epipole in such a point will be treated as suspicious.

Second method assumes, that the (dominant) plane homography H was already recovered. The correspondences that are not consistent with H can be checked, whether there is an alternative match on a location predicted by the homography H. Correspondences with such a property should not be used in the EG estimation, since they possibly are mismatches caused by the repeated pattern.

Relatively common instance of the situation described in this section are images of an urban scene. Consider two images of a house facade (with repeated horizontal pattern) taken with arbitrary camera motion. The mismatches cause the estimated camera motion to be parallel (horizontal) to the repeated pattern, regardless of the true motion.

6.5. Conclusions

The concept of H-degeneracy of a sample of seven correspondences was defined. The cases with five, six and seven H-related correspondences were analyzed separately and a single test of H-degeneracy for all three cases was proposed.

Exploiting the results on H-degeneracy, a novel algorithm, DEGENSAC, was designed. If no large plane is present in the scene, DEGENSAC works exactly as RANSAC. If a dominant plane is detected, DEGENSAC switches to EG estimation using the plane-and-parallax strategy. The computational cost of the H-degeneracy test and, potentially, plane-and-parallax is not significant, since the steps not present in RANSAC are executed only in $\log(k)$ times, where k is the total number of samples drawn.

The EG estimation process was analyzed and we showed that the larger the number of H-related correspondences, the higher the probability that DEGENSAC finds the solution. As a consequence, with the increase in the number of points in a dominant plane the running time of DEGENSAC decreases. It was demonstrated experimentally that, unlike RANSAC, DEGENSAC estimates both the EG and the homography correctly and reliably in the presence of a dominant plane.

7 PROSAC

Standard RANSAC algorithm does not model the local matching process. It is viewed as a black box that generates N tentative correspondences, i.e. the error-prone matches established by comparing local descriptors. The set $\mathcal{U}$ of tentative correspondences contains an *a priori* unknown number I of correct matches (inliers). The inliers are consistent with a global geometric model that is found by fitting a model to a randomly selected subset of $\mathcal{U}$. The hypothesize-and-test loop is terminated when the probability of finding a superior solution falls below a pre-selected threshold. The time complexity of RANSAC depends on N, I, and the complexity m of the geometric model. The average number of samples drawn is proportional to $(N/I)^m$ [22].

In this chapter, we introduce a new hypothesize-and-verify (sample-and-test) matching ap-

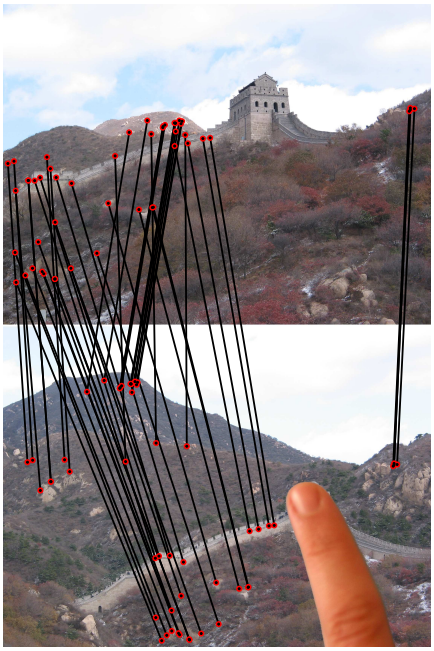


Figure 7.1.: The Great Wall image pair with an occlusion. Given 250 tentative correspondences as input, both PROSAC and RANSAC found 57 correct correspondences (inliers). To estimate the epipolar geometry, RANSAC tested 106,534 seven-tuples of correspondences in 10.76 seconds while PROSAC tested only 9 seven-tuples in 0.06 sec (on average, over hundred runs). Inlier correspondences are marked by a line segment joining the corresponding points.

proach called PROSAC (PROgressive SAmple Consensus). The method achieves large computational savings (with speed-up factors of the order of 10^2 compared to RANSAC) by exploiting the linear ordering structure of $\mathcal{U}$. The ordering is defined at least implicitly in all commonly used local matching methods, because the set of tentative correspondences is obtained by first evaluating a real-valued similarity function (or "quality") $q(\cdot)$ that is subsequently thresholded to obtain the N correspondences. Correlation of intensities around points of interest [64], Mahalanobis distance of invariant descriptors [60] or the ratio of distances in the SIFT space of the first to second nearest neighbor [32] are commonly used examples of $q(\cdot)$.

In PROSAC, samples are semi-randomly drawn from progressively larger sets of tentative correspondences. The improvement in efficiency rests on the mild assumption that tentative correspondences with high similarity are more likely to be inliers. More precisely, we assume that the ordering defined by the similarity used during the formation of tentative matches is not worse than random ordering. The assumption was found valid in our experiments, for all quality functions and for all tested image pairs. Experiments presented in Section 7.2 demonstrate that the fraction of inliers among the top n sorted correspondences falls off fairly rapidly and consequently PROSAC is orders of magnitude faster than the worst-case prediction.

The PROSAC process is in principle simple, but to fully specify it, two problems must be addressed. First, the growth function $n = g(t)$ that defines the set $\mathcal{U}_n$ of n top-ranked correspondences that is sampled after t trials must be selected. Second, a stopping criterion giving guarantees similar to RANSAC about the optimality of the obtained solution must be found. We propose a growth function $g(t)$ guaranteeing that PROSAC is at least equally likely to find the optimal solution as RANSAC. However, we have not been able to prove analytically that PROSAC and RANSAC have the same performance for the worst-case situation, i.e. when the correspondences are ordered randomly. Nevertheless, the comparison of PROSAC and RANSAC on randomly ordered sets of correspondences showed that their performance was effectively identical.

The PROSAC algorithm has two other desirable features. The size N of the set of tentative correspondences has limited influence on its speed, since the solution is typically found early, when samples are taken from a smaller set. One parameter of the matching process is thus effectively removed. Instead, the user controls the behavior of PROSAC by specifying the time when the sampling distribution of PROSAC and RANSAC become identical. For the growth function $g(t)$ selected according to the above-mentioned criteria, PROSAC can be interpreted as a process running RANSAC processes in parallel for all $\mathcal{U}_n, n \in \{m \dots N\}$. In experiments presented in Section 7.2, PROSAC speed was close to that of RANSAC that would operate on (the *a priori* unknown) set of correspondences with the highest inlier ratio.

Related work. Tordoff and Murray [53] combine the MLESAC [57] algorithm with non-uniform (guided) sampling of correspondences. This is the published work closest to PROSAC, that differs in two important aspects. First, guided sampling requires estimates of the probability of correctness of individual correspondences while here we only assume that some quantity monotonically related to the probability is available. Second, PROSAC dynamically adapts the sampling strategy to the information revealed by the sampling process itself. The hypothesize-and-verify loop is a series of incorrect guesses until the first success. Each failure decreases the likelihood that the correspondences used to estimate the model parameters are correct. Grad-

ually, the observed evidence against *a priori* preferred correspondences should result in the reduction of their preference. PROSAC can be viewed as an instance of a process that starts by deterministically testing the most promising hypotheses and than converging to uniform sampling as the confidence in the "quality" of the *a priori* sorting declines after unsuccessful tests.

The objective of PROSAC is to find inliers in the set of all tentative correspondences $\mathcal{U}_N$ in the shortest possible time and to guarantee, with a certain probability, that all inliers from $\mathcal{U}_N$ are found. Issues related to the precision of the model that is computed from the set of inliers are not discussed since they are not directly related to the problem of efficient sampling. Bundle adjustment [59] can be performed *ex post*.

Notation. The set of N data points (tentative correspondences) is denoted as $\mathcal{U}_N$. The data points in $\mathcal{U}_N$ are sorted in descending order with respect to the quality function q

$$\mathbf{u}_i, \mathbf{u}_j \in \mathcal{U}_N : i < j \Rightarrow q(\mathbf{u}_i) \geq q(\mathbf{u}_j).$$

A set of n data points with the highest quality is denoted $\mathcal{U}_n$. A sample $\mathcal{M}$ is a subset of data points $\mathcal{M} \subset \mathcal{U}_N$, $|\mathcal{M}| = m$ where m is the size (cardinality) of the sample. The quality function on samples is defined as the lowest quality of a data point included in the sample

$$q(\mathcal{M}) = \min_{\mathbf{u}_i \in \mathcal{M}} q(\mathbf{u}_i).$$

7.1. Algorithm

The structure of the PROSAC algorithm is similar to RANSAC. First, hypotheses are generated by random sampling. The samples, unlike in RANSAC, are not drawn form all data, but from a subset of the data with the highest quality. The size of the hypothesis generation set is gradually increased. The samples that are more likely to be uncontaminated are therefore examined early. In fact, PROSAC is designed to draw the same samples as RANSAC, only in a different order. The hypotheses are verified against all data. As in RANSAC, the algorithm is terminated when the probability of the existence of solution that would be better than the best so far becomes low (smaller than 5%). Two important issues, the choice of the size of the hypothesis generation set and the termination criterion of the sampling process, are discussed below.

7.1.1. The Growth Function and Sampling

The design of the growth function defining the $\mathcal{U}_n$ must find a balance between the over-optimistic reliance on the pre-sorting by the quality and the over-pessimistic RANSAC approach that treats all correspondences equally. If the probabilities $P\{\mathbf{u}_i\} = P\{\text{corresp. } \mathbf{u}_i \text{ is correct}\}$ were known, it would be in principle possible to adopt a Bayesian approach. After each sample-and-test cycle, the posterior probability would be re-computed for all correspondences included in the sample. The correspondences would be sorted by their posterior probabilities and samples with the highest probability would be drawn. We pursued this line of research, but abandoned it for two reasons. Firstly, probabilities $P\{\mathbf{u}_i\}$ of correspondences tested are not independent after a test and it is not feasible to represent the joint probability for all but the simplest models. Secondly, errors in estimates of $P\{\mathbf{u}_i\}$ propagate through the Bayes formula and are accentuated.

So if the initial estimate of $P\{\mathbf{u}_i\}$ based on the similarity of the correspondence is incorrect, the posterior probability becomes worthless soon.

The alternative, pursued here, is to make minimal assumptions about the link between $P\{\mathbf{u}_i\}$ and the similarity function $q(\mathbf{u}_j)$. In particular, we assume monotonicity, i.e.

$$q(\mathbf{u}_i) \geq q(\mathbf{u}_j) \Rightarrow P\{\mathbf{u}_i\} \geq P\{\mathbf{u}_j\}. \tag{7.1}$$

Sequences of correspondences satisfying

$$i < j \Rightarrow P\{\mathbf{u}_i\} \geq P\{\mathbf{u}_j\} \tag{7.2}$$

will be called not-worse-than-random.

Note that we are searching for a *single* growth function. It seems possible to adapt the growth function to reflect the result of previous sample-and-test cycles. However, all PROSAC (and RANSAC) runs are alike: a sequence of failures followed by a 'hit' due to an all-inlier sample. The history of the sampling process is thus fully captured by t, the number of tests carried so far.

The sampling strategy. Imagine standard RANSAC drawing T_N samples of size m out of N data points. Let $\{\mathcal{M}_i\}_{i=1}^{T_N}$ denote the sequence of samples $\mathcal{M}_i \subset \mathcal{U}_N$ that are uniformly drawn by RANSAC, and let $\{\mathcal{M}_{(i)}\}_{i=1}^{T_N}$ be sequence of the same samples sorted in descending order according to the sample quality

$$i < j \Rightarrow q(\mathcal{M}_{(i)}) \geq q(\mathcal{M}_{(j)}).$$

If the samples are taken in order $\mathcal{M}_{(i)}$, the samples that are more likely to be uncontaminated are drawn earlier. Progressively, samples containing data points with lower quality function are drawn. After T_N samples, exactly all RANSAC samples $\{\mathcal{M}_i\}_{i=1}^{T_N}$ were drawn.

Let T_n be an average number of samples from $\{\mathcal{M}_i\}_{i=1}^{T_N}$ that contain data points from $\mathcal{U}_n$ only

$$T_n = T_N \frac{\binom{n}{m}}{\binom{N}{m}} = T_N \prod_{i=0}^{m-1} \frac{n-i}{N-i}, \text{ then}$$

$$\frac{T_{n+1}}{T_n} = \frac{T_N}{T_N} \prod_{i=0}^{m-1} \frac{n+1-i}{N-i} \prod_{i=0}^{m-1} \frac{N-i}{n-i} = \frac{n+1}{n+1-m}.$$

Finally, the recurrent relation for T_{n+1} is

$$T_{n+1} = \frac{n+1}{n+1-m} T_n. \tag{7.3}$$

There are T_n samples containing only data points from $\mathcal{U}_n$ and T_{n+1} samples containing only data points from $\mathcal{U}_{n+1}$. Since $\mathcal{U}_{n+1} = \mathcal{U}_n \cup \{\mathbf{u}_{n+1}\}$, there are $T_{n+1} - T_n$ samples that contain a data point $\mathbf{u}_{n+1}$ and $m-1$ data points drawn from $\mathcal{U}_n$. Therefore, the procedure that for $n = m \ldots N$ draws $T_{n+1} - T_n$ samples consisting of a data point $\mathbf{u}_{n+1}$ and $m-1$ data points drawn from $\mathcal{U}_n$ at random efficiently generates samples $\mathcal{M}_{(i)}$.

As the values of T_n are not integer in general, we define $T'_m = 1$ and

$$T'_{n+1} = T'_n + \lceil T_{n+1} - T_n \rceil. \tag{7.4}$$

The growth function is then defined as

$$g(t) = \min\{n : T'_n \geq t\}. \tag{7.5}$$

In PROSAC, the t-th sample $\mathcal{M}_t$ consists of

$$\mathcal{M}_t = \{\mathbf{u}_{g(t)}\} \cup \mathcal{M}'_t, \tag{7.6}$$

where $\mathcal{M}'_t \subset \mathcal{U}_{g(t)-1}$ is a set of $|\mathcal{M}'_t| = m - 1$ data points drawn from $\mathcal{U}_{g(t)-1}$ at random. The parameter T_N defines after how many samples the behavior of PROSAC and RANSAC becomes identical. In our experiments, the parameter was set to $T_N = 200000$.

$t := 0$, $n := m$, $n^* := N$

Repeat until a solution satisfying eqs. (3.3), (7.9) is found.

1. Choice of the hypothesis generation set

$t := t + 1$

if $(t = T'_n)$ & $(n < n^*)$ then $n := n + 1$ (see eqn. 7.4)

2. Semi-random sample $\mathcal{M}_t$ of size m

if $T'_n < t$ then

The sample contains $m - 1$ points selected from $\mathcal{U}_{n-1}$ at random and $\mathbf{u}_n$

else

Select m points form $\mathcal{U}_n$ at random

3. Model parameter estimation

Compute model parameters p_t from the sample $\mathcal{M}_t$

4. Model verification

Find support (i.e. consistent data points) of the model with parameters p_t

Select termination length n^* if possible according to Section 7.1.2

Algorithm 6: Outline of the PROSAC algorithm.

7.1.2. Stopping Criterion

The PROSAC algorithm terminates if the number of inliers I_{n^*} within the set $\mathcal{U}_{n^*}$ satisfies the following conditions:

- *non-randomness* – the probability that I_{n^*} out of n^* data points are by chance inliers to an arbitrary incorrect model is smaller than Ψ (typically set to 5%)
- *maximality* – the probability that a solution with more than I_{n^*} inliers in $\mathcal{U}_{n^*}$ exists and was not found after k samples is smaller than η_0 (typically set to 5%).

From all such solutions the one that causes the termination first is chosen.

The **non-randomness** requirement prevents PROSAC from selecting a solution supported by outliers that are by chance consistent with it. The constraint is typically checked ex-post in standard approaches [2]. The distribution of the cardinalities of sets of random 'inliers' is binomial

$$P_n^R(i) = \beta^{i-m}(1-\beta)^{n-i+m}\binom{n-m}{i-m}, \tag{7.7}$$

where β is the probability, that an incorrect model calculated from a random sample containing an outlier is supported by a correspondence not included in the sample.

We set β pessimistically based on geometric considerations. If needed, the estimate of β can be made more precise during the sampling of PROSAC.

For each n, the minimal number of inliers $I_n^{\min}$ is calculated so that the probability of size of such support being random is smaller than Ψ

$$I_n^{\min} = \min\left\{ j : \sum_{i=j}^{n} P_n^R(i) < \Psi \right\}. \tag{7.8}$$

A non-random solution found on $\mathcal{U}_{n^*}$ must satisfy

$$I_{n^*} \geq I_{n^*}^{\min}. \tag{7.9}$$

A **maximality** constraint defines how many samples are needed to be drawn to ensure the confidence in the solution and is the (only) termination criterion of RANSAC [22].

For a hypothesis generation set $\mathcal{U}_n$, the probability P_{I_n} that an uncontaminated sample of size m is randomly selected from a set $\mathcal{U}_n$ of n correspondences is

$$P_{I_n} = \frac{\binom{I_n}{m}}{\binom{n}{m}} = \prod_{j=0}^{m-1} \frac{I_n - j}{n - j} \approx \varepsilon_n^m, \tag{7.10}$$

where I_n is the number of inliers in $\mathcal{U}_n$ and $\varepsilon_n = I_n/n$ is the fraction of inliers. The probability η of missing a set of inliers of the size I_n on the set $\mathcal{U}_n$ after k samples of PROSAC, where $g(k) \leq n$, is

$$\eta = (1 - P_{I_n})^k. \tag{7.11}$$

The number of samples that have to be drawn to ensure the probability η falls under the predefined threshold η_0 is

$$k_{n^*}(\eta_0) \geq \log(\eta_0)/\log(1 - P_{I_{n^*}}). \tag{7.12}$$

The termination length n^* is chosen to minimize $k_{n^*}(\eta_0)$ subject to $I_{n^*} \geq I_{n^*}^{\min}$.

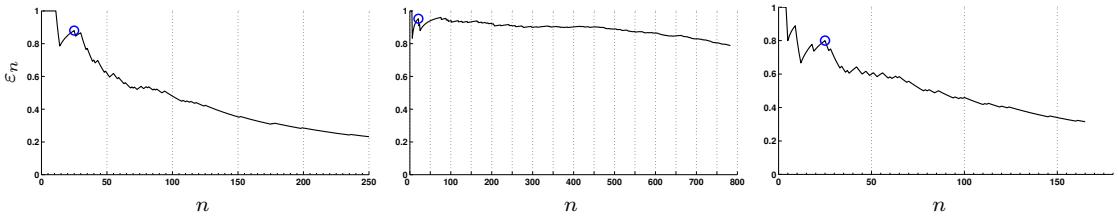


Figure 7.2.: The fraction of inliers ε among top n correspondences sorted by quality for the GREAT WALL (left), MUG background (center) and MUG foreground (right) scenes. The circles mark the (average) size of the largest set of correspondences that PROSAC sampled, i.e. the size it sampled when it stopped. The circles are close to the optimal stopping size.

7.2. Experiments

The not-worse-than-random assumption about the ordering of correspondences was tested for two different similarity functions.

Matching based on SIFT descriptors [32] was used to obtain tentative correspondences in PLANT and MUG experiments[1]. The similarity was defined as the ratio of the distances in the SIFT space of the best and second match. The threshold for the similarity is set to 0.8 as suggested in [32]. This setting has been shown experimentally [32, 50] to provide a high fraction of inliers in the tentative correspondences. However, this thresholding also leads to small absolute number of tentative correspondences.

In the GREAT WALL experiment, the Euclidean distance of the first fifteen Discrete Cosine Transform (DCT) coefficients was used as a similarity function [42, 7]. The DCT was computed on normalized, affine invariantly detected, parallelograms. As tentative correspondences, points with mutually best similarity were selected.

Figures 7.2 and 7.4 show the dependence of the fraction of inliers ε on the order of a tentative correspondence induced by the similarity function. In all experiments, regardless of the similarity function used, the assumption of not-worse-than-random ordering held. The fraction of inliers ε decreased almost monotonically as a function of the number of tentative correspondences n.

Comparison of efficiency. The number of samples drawn by RANSAC and PROSAC as well as wall clock time of both algorithms were measured on three scenes.

For **the GREAT WALL scene** (Fig. 7.1) both PROSAC and RANSAC algorithms found 57 inliers among the 250 tentative correspondences. RANSAC needed on average (over hundred runs) 106,534 samples which took 10.76 sec. PROSAC estimated the 57-inlier epipolar geometry after 9 samples (0.06 sec) on average.

The PLANT scene is challenging due to a large number of depth discontinuities and the presence of both repetitive (floor) and locally similar (leafs) patterns. Tentative correspondences are obtained by SIFT matching [32] computed on MSERs [35]. For the 0.8 threshold, the tentative correspondences include only 12 inliers. The epipolar geometry could not be estimated reliably

[1]The code was kindly provided by David Lowe, UBC, Vancouver, Canada.

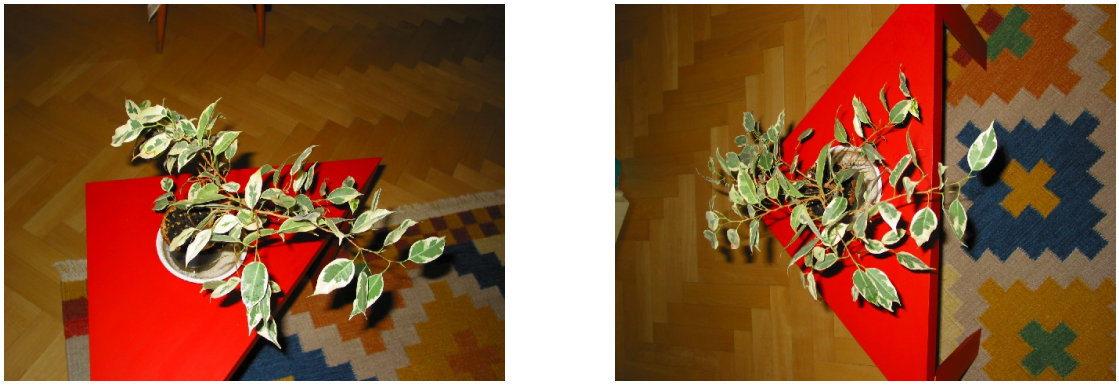


Figure 7.3.: The PLANT scene. Depth discontinuities, self-occlusions and repetitive patterns reduce the probability that a correspondence with high quality (similarity) is indeed an inlier. RANSAC fails to estimate epipolar geometry on this image pair.

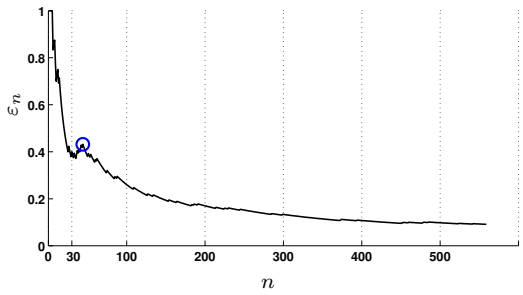


Figure 7.4.: The dependence of the fraction of inliers ε on the ordering by the SIFT similarity on the PLANT scene. The first 30 correspondences have similarity below the threshold 0.8. The circle shows the optimal stopping length n^* chosen by PROSAC.

Figure 7.5.: Motion segmentation. The motion of the background and the foreground (the mug) are denoted by light and black lines respectively.

Background		$N = 783$, $\varepsilon = 79\%$	
	I	k	time [sec]
PROSAC	617	1.0	0.33
RANSAC	617	15	1.10
Mug		$N = 166$, $\varepsilon = 31\%$	
	I	k	time [sec]
PROSAC	51.6	18	0.12
RANSAC	52.3	10,551	0.96

Table 7.1.: The number of inliers (I) detected, samples (k) and the time needed in the motion estimation of the background (top) and the foreground (bottom) in the MUG experiment.

by either RANSAC or PROSAC. When the SIFT threshold was set to 0.95, there were $N = 559$ tentative correspondences and $I = 51$ inliers. In this case, RANSAC fails due to low fraction of inliers $\varepsilon = 9.2\%$; on average, RANSAC would need $8.43 \cdot 10^7$ samples to find the solution, which is not realistic. PROSAC, on the other hand, found all the inliers and estimated the correct epipolar geometry after 3,576 samples in 0.76 sec on average (over 100 execution).

The **MUG scene**, (Fig. 7.5) is non-rigid. The mug moved between the acquisition of the two images. Tentative correspondences were obtained matching MSER [35] and affine-invariant [38] regions using SIFT descriptors. First, epipolar geometry was estimated on all tentative correspondences. Then, inliers to the first epipolar geometry were removed and another EG was estimated to segment the motion in the scene. The results are summarized in Tab. 7.1.

In the Fig. 7.2, the dependency of ε on the ordering is shown for the background segmentation (center) and the foreground (the mug) segmentation (right). All the correct correspondences on the mug are outliers to the background motion while having high similarity score. This can be observed in Fig. 7.2(center) as dips in the plot. This also shows that the probability of a

	k	$\min k$	$\max k$	time [sec]
RANSAC	106,534	97,702	126,069	10.76
PROSAC	9	5	29	0.06
PROSAC OR	61,263	1,465	110,727	6.28

Table 7.2.: The comparison of the number of samples drawn by RANSAC, PROSAC, and PROSAC with random ordering on the GREAT WALL experiment. The values of k are average, minimum, and maximum over 100 runs respectively.

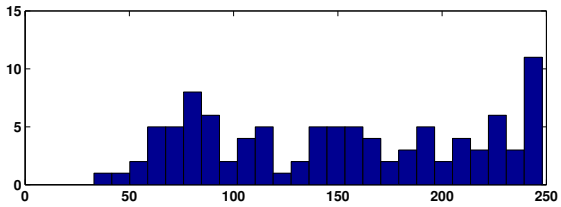


Figure 7.6.: Histogram of the stopping length n^* of PROSAC for 100 random orderings in the GREAT WALL scene.

correspondence being an inlier to a given geometric model depends on other factors besides the similarity of local descriptors.

Testing robustness to **the worst case situation**, i.e. to random ordering of tentative correspondences. To compare RANSAC and PROSAC in the least favorable situation for PROSAC, an experiment was carried out on sequences of randomly permuted correspondences. The 250 correspondences from the GREAT WALL experiment were used. For PROSAC, the average number of samples taken over 100 different random permutations was $61,263$ (standard deviation $2.94 \cdot 10^4$). The number of samples ranged from $1,465$ to $110,727$. For RANSAC, the average and standard deviation were $130,419$ and $6.55 \cdot 10^3$ respectively. The results, together with results of PROSAC on tentative correspondences sorted by the similarity function are shown in Tab. 7.2.

PROSAC drew less samples and was faster than RANSAC in this experiment. The difference in the average number of samples can be attributed to the fact that even in a randomly permuted sequence there are sub-sequences with higher than average inlier fractions, allowing PROSAC to terminate earlier. The histogram of PROSAC termination lengths n^* is plotted in Fig. 7.6. Only a fraction of PROSAC executions were terminated on the full set of the tentative correspondences, where RANSAC is terminated. However, the maximal number of samples drawn are comparable for PROSAC and RANSAC, i.e. in the worst case PROSAC behaves as RANSAC.

The stopping criterion. In all experiments, the optimal stopping length n^* calculated ex-post was identical with the length automatically selected by PROSAC. The values of n^* are depicted in Figs 7.2 and 7.4 as circles.

7.3. Conclusions

PROSAC – a novel robust estimator of the hypothesize-and-verify type was introduced. The PROSAC algorithm exploits the ordering structure of the set of tentative correspondences, assuming that the ordering by similarity computed on local descriptors is better than random. The assumption held in all our experiment for both quality measures that were tested.

The sampling on progressively larger subsets consisting of top-ranked correspondences brings very significant computational savings. In comparison to RANSAC, PROSAC was more than hundred time faster on non-trivial problems. Using synthetically generated sequences of correspondences, we showed that the worst-case behavior of PROSAC and RANSAC are effectively identical.

PROSAC removes one parameter of the matching process – the threshold on the similarity function for selection of tentative correspondences. Thus robustness against either having too few correspondences or a large number of tentative correspondences with high outlier percentage is achieved. In one tested problem, PROSAC found a solution of the matching problem that cannot be solved by RANSAC.

8 Randomized RANSAC

The speed of RANSAC depends on two factors. First, the percentage of outliers determines the number of random samples needed to guarantee the $1 - \eta_0$ confidence in the optimality of the solution. Second, the time needed to assess the quality of a hypothesized model parameters is proportional to the number N of the input data points. The total running time t of RANSAC can be expressed as

$$t = k(t_M + \overline{m}_S \, t_V), \tag{8.1}$$

where k is the number of samples drawn, t_M is time needed to instantiate a model hypotheses given a sample, $\overline{m}_S$ is an average number of models per sample and t_V is average time needed to evaluate the quality of the sample. We choose the time needed to verify a single correspondence as the unit of time for t_M, t_V and t (in standard RANSAC, $t_V = N$).

The core idea of the Randomized (hypothesis evaluation) RANSAC is that most model parameter hypotheses evaluated are influenced by outliers. To reject such erroneous models, it is sufficient to perform a statistical test on only a small number of data points. The test can be formulated as follows. The hypothesis generation step proposes a model. It is either 'good', i.e. it leads to the optimal solution (the solution with maximal support), or it is 'bad', i.e. one of the data points in the sample is an outlier. The property 'good' is a hidden state that is not directly observable but is statistically linked to observable events. The observable events are "data point (correspondence) is/is-not consistent with the model".

The statistical test reduces the number of verified correspondences (and thus time complexity of the verification step) on one hand, and brings in a probability α of rejection (overlooking) of a good sample on the other. The probability α is a significance of the test and influences (increases) the number of samples drawn before the $1 - \eta_0$ confidence is ensured. The genuine model parameters are recovered if an uncontaminated sample is drawn and passes the test. This happens with probability

$$P = P_g(1 - \alpha).$$

Typically, the more data points are processed by the test the lower probability that a 'good' model will be rejected. The task is to find the balance between the number of correspondences processed by the test and the increase in the number of samples, so that the total running time t (8.1) is minimised. To simplify the derivation, we show that minimizing time to the first all-inlier sample is approximately equivalent. It was shown in section 3.2.1 that $k_{\eta_0} \approx -\ln \eta_0 \bar{k}$, where $\bar{k}$ is the average number of samples before first uncontaminated sample is drawn. Therefore the total time and the time to the first all-inlier sample differ only by a multiplicative constant and minimizing (8.1) with $k = \bar{k}$ and $k = k_{\eta_0}$ is equivalent. Since the average time to draw an

uncontaminated model that passes the test is $\bar{k} = 1/(P_g(1-\alpha))$, we have

$$t = \frac{1}{P_g(1-\alpha)}(t_M + \overline{m}_S\, t_V). \tag{8.2}$$

The design of the test generally depends on two probabilities ε and δ, where ε denotes the fraction of inliers within the set of data points and δ is a probability that a data point is consistent with a model with arbitrary parameters. These probabilities are typically unknown beforehand and have to be either estimated during the course of the algorithm, or the test must be efficient for very large range of the values.

Two different test are described in the chapter: $T_{(d,d)}$ test and a test based on Wald's Sequential Probability Ratio Test (SPRT) [61].

The $T_{(d,d)}$ test (section 8.1) is based on testing d data points, the test is passed when all d out of d randomly selected data points are consistent with the model. This test is not optimal, however it is simple and the derivation of it properties is mathematically tractable. Experiments demonstrate, that even such a simple test can increase the speed of the RANSAC procedure.

The RANSAC with SPRT provides an optimal sequential test assuming that the parameters ε and δ are known beforehand, as shown in section 8.2. An algorithm with near optimal behaviour, that estimates the probabilities ε and δ is discussed in section 8.3.

8.1. The $T_{d,d}$ Test

In this section we introduce a simple and thus mathematically tractable class of preverification tests. Despite its simplicity, the experiments show its potential. The test we analyze is defined as follows:

Definition (the T(d,d)) The T(d,d) is passed if all d data points out of d randomly selected are consistent with the hypothesized model.

In the $T_{d,d}$ test, the number of verified correspondences per a test depends on d and is expressed as

$$t_V(d) = P_g\big(\alpha N + (1-\alpha)\bar{t}_\alpha\big) + (1-P_g)\big(\beta N + (1-\beta)\bar{t}_\beta\big). \tag{8.3}$$

Here, β stands for the probability, that a bad sample passes the preverification test. Note that it is important that $\beta \ll \alpha$, i.e. a bad (contaminated) sample is consistent with a smaller number of data points than a good sample. In the rest of this section we derive the optimal value for d. First of all we express constants as introduced in the previous section as

$$\alpha = \varepsilon^d \qquad \text{and} \qquad \beta = \delta^d,$$

where δ is the probability that a data point is consistent with a "random" model. Since we do not need to test all d points (since single failure means that the pre-test failed), the average time spent in the preverification test is

$$\bar{t}_\alpha = \sum_{i=1}^{d} i\,(1-\varepsilon)\,\varepsilon^{i-1} \qquad \text{and} \qquad \bar{t}_\beta = \sum_{i=1}^{d} i\,(1-\delta)\,\delta^{i-1}$$

Since

$$\sum_{i=1}^{d} i(1-x)x^{i-1} \leq \sum_{i=1}^{\infty} i(1-x)x^{i-1} = \frac{1}{1-x}, \tag{8.4}$$

we have

$$\bar{t}_{\alpha} \leq \frac{1}{1-\varepsilon} \qquad \text{and} \qquad \bar{t}_{\beta} \leq \frac{1}{1-\delta}.$$

The approximation we get after substituting (8.4) into (8.3)

$$t_V(d) \approx \varepsilon^m \left(\varepsilon^d N + \frac{1-\varepsilon^d}{1-\varepsilon} \right) + (1-\varepsilon^m) \left(\delta^d N + \frac{1-\delta^d}{1-\delta} \right)$$

is too complicated for finding optimal d. Therefore, we incorporate the following approximations

$$(1-\varepsilon^m)\frac{1-\delta^d}{1-\delta} \approx 1,$$

$$(1-\varepsilon^m)\delta^d N \approx \delta^d N, \text{ and}$$

$$\varepsilon^d N \gg \frac{1-\varepsilon^d}{1-\varepsilon},$$

which are sufficiently accurate for commonly encountered values of ε, δ and N. After applying these approximations, we have

$$t_V(d) \approx N\,\delta^d + 1 + \varepsilon^{m+d}\,N \tag{8.5}$$

The average time spent in R-RANSAC in number of verified data points is then approximately

$$t(T_{d,d}) \approx \frac{1}{\varepsilon^m\,\varepsilon^d} \left(t_M + \overline{m}_S (N\,\delta^d + \varepsilon^{m+d}\,N + 1) \right) \tag{8.6}$$

We are looking for the minimum of $t(T_{d,d})$ which is found by solving for d in $\frac{\partial t(T_{d,d})}{\partial d} = 0$. The optimal length of the $T_{d,d}$ test is

$$d^* \approx \frac{\ln\left(\frac{\ln \varepsilon (t_M + \overline{m}_S)}{\overline{m}_S N\,(\ln \delta - \ln \varepsilon)} \right)}{\ln \delta}. \tag{8.7}$$

The value of d_{opt} must be an integer greater or equal to zero, so it could be expressed as

$$d_{opt} = \max(0, \arg \min_{d \in \{\lfloor d^* \rfloor, \lceil d^* \rceil\}} t(T_{d,d})). \tag{8.8}$$

Since the cost function $t(T_{d,d})$ has only one extremum and for $d \to \pm\infty$ we have $t(T_{d,d}) \to \infty$, we can say that R-RANSAC is faster than the standard RANSAC if

$$t(T_{0,0}) > t(T_{1,1}).$$

From this equation we get (after approximation)

$$N > \frac{t_M + \overline{m}_S}{\overline{m}_S} \cdot \frac{1-\varepsilon}{\varepsilon - \delta}. \tag{8.9}$$

For broad range of values of ε and δ, the optimal value of d is $d = 1$. Therefore, without any prior knowledge of ε and δ, the suggested test (from the $T_{d,d}$ class of tests) is $T_{1,1}$.
Note: Instead of drawing m data points and verifying only d data points in the test, the following approach can be adopted in the $T_{d,d}$ test when efficient. Draw randomly $m + d$ data points and fit model parameters to the sample. Measure error on the $m + d$ data points in the sample. If the error is smaller than predefined threshold Δ, i.e. there is a model that fits all the data well, proceed with the verification step, otherwise generate a new hypothesis. The advantage of such an approach is, that a single model can be obtained from $m + d$ data points (in contrary to $\overline{m}_S$ models from m data points only) and also a model of higher accuracy is typically obtained if more points are used. This is specific to the $T_{d,d}$ class of tests, where the consistency of all d data points is required.

8.2. The Optimal Sequential Test

Thus far, a special type (simple and mathematically tractable) of test was described and its property were derived. In this section, the theory of sequential testing is exploited.

In sequential testing, as applied e.g. in industrial inspection, the problem is to decide whether the model (or the batch of products) is 'good' or 'bad' in the shortest possible time (i.e. making the smallest number of observations) and yet satisfying the predefined bounds on the probabilities of the two possible errors – accepting a 'bad' model as 'good' and vice versa. Wald proposed the sequential probability ratio test (SPRT) and showed [61] that, given errors bound on the errors of the first and second kind, it minimizes the number of observations (time to decision) [1].

Wald's SPRT test is a solution of a *constrained optimization* problem. The user supplies the acceptable probabilities of the errors of the first and the second kind and the resulting optimal test is a trade-off between time to decision (or cost of observations) and the errors committed. However, when evaluating RANSAC, the situation is different. First of all, a 'good' model is always evaluated for all data points (correspondences) since the number of inliers is one of the outputs of the algorithms. So the only error that can be committed is an early rejection of a 'good' model (error of the first kind). But this only means that more samples have to be drawn to achieve the required confidence $1 - \eta_0$ of finding the optimal solution. So unlike in the classical setting, we are solving a *global optimization* problem, minimizing a single real number – the time to decision, since the consequence of an error is also a loss of time.

The model evaluation step of the optimal R-RANSAC proceeds as Wald's sequential probability ratio test (SPRT) with the probability α of rejecting a 'good' sample set to achieve maximum speed of the whole RANSAC process. To understand the operation of R-RANSAC with SPRT,

[1] Precisely speaking, the SPRT is only approximately optimal. However, the approximation has been shown by Wald to be so close to the optimum that, for practical purposes, it is considered the optimal test.

some familiarity with Wald's decision theory is required. We therefore introduce its relevant parts. Some of the results are presented in a form that is not fully general, but sufficient for the derivation of the R-RANSAC with SPRT algorithm. Some of Wald's terminology is modified in order to make the exposition more accessible.

In the model evaluation step, our objective is to decide between the hypothesis H_g that model is 'good' and the alternative hypothesis H_b that the model is 'bad'. A 'good' model is computed from an all-inlier sample. The Wald's SPRT is based on the likelihood ratio [61]

$$\lambda_j = \prod_{r=1}^{j} \frac{p(x_r|H_b)}{p(x_r|H_g)} = \lambda_{j-1} \cdot \frac{p(x_j|H_b)}{p(x_j|H_g)}, \tag{8.10}$$

a ratio of two conditional probabilities of an observation x_r under the assumptions of H_g and H_b respectively. In RANSAC, x_r is equal to 1 if the r-th data point is consistent with a model with parameters θ and 0 otherwise. For example, a correspondence is consistent with (i.e. supporting) an epipolar geometry represented by a fundamental matrix F if its Sampson's error is smaller than some predefined threshold [22]. The probability $p(1|H_g)$ that any randomly chosen data point is consistent with a 'good' model is approximated by the fraction of inliers ε among the data points[2]. The probability of a data point being consistent with a 'bad' model is modeled as a probability of a random event with Bernoulli distribution with parameter δ: $p(1|H_b) = \delta$. The process of estimation of δ and ε is discussed in section 8.3.

Output: model accepted/rejected, number of tested data points, a fraction of data points consistent with the model

Set $j = 1$

1 Check whether j-th data point is consistent with the model
2 Compute the likelihood ratio λ_j eq. (8.10)
3 If $\lambda_j > A$, decide the model is 'bad' (model "rejected"), else increment j or continue testing
4 If $j = N$ the number of correspondences decide model "accepted"

Algorithm 7: The adapted sequential probability ratio test (Adapted SPRT).

After each observation the standard Wald's SPRT makes one of three decisions: accept a 'good' model, reject a 'bad' model, or continue testing. Since in RANSAC the total number of inliers is needed to decide on termination, nothing is gained by an early decision in favor of a 'good' model. Therefore the option of an early acceptance of the model has been removed in the adapted SPRT (Alg. 7). The full SPRT is described e.g. in Wald [61] and, in a more accessible form, in Lee [30].

[2]The probability ε would be exact if the data points were selected with replacement. Since the objective of the verification is to count the size of the support of the model, the correspondences are drawn without replacement. However, the approximation is close.

8.2.1. The Optimal Value of the Decision Threshold

The decision threshold A is the only parameter of the Adapted SPRT. We show how to set it to achieve optimal performance, i.e. minimal average RANSAC running time given the probabilities δ and ε. We use the following theorems (for proofs, see [61]).

Theorem 7 *The probability α of rejecting a 'good' model in SPRT $\alpha \leq 1/A$.*

Proof: Wald's theorem [61, p. 41] states $\alpha \leq (1-\beta)/A$, where β stands for the probability that a 'bad' model is incorrectly accepted as 'good'. In the adapted SPRT, since the only decision of the test can be "reject", $\beta = 0$ and thus $\alpha \leq 1/A$. □
The approximation $\alpha \approx 1/A$ is close and is often used.

Theorem 8 (Wald's lemma) *The average number of observations (checked data points) carried out while testing a 'bad' model is $C^{-1} \log A$, where*

$$C = p(0|H_b) \log \frac{p(0|H_b)}{p(0|H_g)} + p(1|H_b) \log \frac{p(1|H_b)}{p(1|H_g)}. \tag{8.11}$$

Proof: According to [61, p. 53]

$$C = E\left(\log \frac{p(x|H_b)}{p(x|H_g)}\right). \tag{8.12}$$

The value of x is from $\{0, 1\}$. The expectation E is a sum of two terms weighted by probability $p(x|H_b)$. Equation (8.11) follows. □

In the particular case of RANSAC, $p(1|H_b) = \delta$, $p(0|H_b) = 1-\delta$, $p(0|H_g) = 1-\varepsilon$, and $p(1|H_g) = \varepsilon$. Therefore the average number of verified correspondences per model is:

$$\frac{\log A}{C} = \left((1-\delta) \log \frac{1-\delta}{1-\varepsilon} + \delta \log \frac{\delta}{\varepsilon}\right)^{-1} \log A. \tag{8.13}$$

The value of A influences the total running time in two opposing ways. The larger the value of A, the smaller the probability of rejection of a 'good' model. On the other hand, the number of correspondendes verified per model increases with $\log A$ (eq (8.13)). We wish to set A to achieve minimal average time needed to find the solution.

The average time-to-solution in R-RANSAC is $t = \bar{k}\bar{t}_s$, where $\bar{k}$ is the average number of samples drawn until a 'good' model and $\bar{t}_s$ is the average testing time per sample. In the following, the time unit will be the time needed to check one data point. The probability P_g of drawing a 'good' model is $P_g = \varepsilon^m$, where m is the number of data points in the RANSAC sample. The number of tested samples before a 'good' one is drawn and not rejected is a random variable with geometric distribution and mean $\bar{k} = 1/(P_g(1-\alpha)) \approx 1/(P_g(1-1/A))$. The average time $\bar{t}_s$ of processing a sample consists of two components: time t_M needed to instantiate a

model hypotheses given a sample[3], and the average time of testing each hypothesis. Let $\overline{m}_S$ be the number of models that are verified per sample[4] and $C^{-1}\log A$ be the average length of the SPRT (Theorem 8). The average time to the solution expressed as a function of A is

$$t(A) = \frac{1}{P_g(1-1/A)}(t_M + \overline{m}_S \frac{\log A}{C}). \tag{8.14}$$

The formula (8.14) can be simplified to

$$t(A) = \frac{K_1 + K_2 \log A}{1 - 1/A},$$

where $K_1 = t_M/P_g$ and $K_2 = \overline{m}_S/(P_g C)$. We are interested in the optimal value of A, i.e.

$$A^* = \arg\min_A t(A).$$

The minimum is found by solving

$$\frac{dt}{dA} = -\frac{K_1 + K_2 - K_2 A + K_2 \log A}{(A-1)^2} = 0.$$

After rearrangements, we have

$$A^* = \frac{K_1}{K_2} + 1 + \log A^* = \frac{t_M C}{\overline{m}_S} + 1 + \log A^*. \tag{8.15}$$

Equation (8.15) has two real solutions for positive K_1/K_2, $A_1^* < 1 < A_2^*$. Since $\delta < \varepsilon$, the contribution to the likelihood ratio (eq. (8.10)) of a correspondence that is not consistent with the model is greater than 1, therefore the solution of interest is $A^* > 1$. This solution can be obtained as $A^* = \lim_{n\to\infty} A_n$, where $A_0 = K_1/K_2 + 1$ and $A_{n+1} = K_1/K_2 + 1 + \log(A_n)$. The series converges fast, typically within four iterations.

8.2.2. Extension to MLE Framework

In RANSAC, it is observed whether a data point supports the model or not. In practice, an error function ρ is evaluated and data points with the error function under a threshold are thought to support the model. In MLESAC [57, 53], it is assumed that the error $x = \rho(\theta, \mathbf{x})$ of a data point $\mathbf{x}$ with respect to a model with parameters θ is distributed as a mixture of Gaussian error distribution for inliers and uniform error distribution for outliers for a 'good' model and as a uniform distribution for a 'bad' model

$$\begin{aligned} p(x|H_g) &= \varepsilon\left(\frac{1}{\sigma\sqrt{2\pi}} e^{\frac{x^2}{2\sigma^2}}\right) + (1-\varepsilon)\frac{1}{Z} \\ p(x|H_b) &= \frac{1}{Z} \end{aligned}$$

[3]Computing model parameters from a sample takes the same time as verification of t_M data points.
[4]In the 7-pt algorithm for epipolar geometry estimation, 1 to 3 models have to be verified.

Then, the likelihood ratio λ_j is expressed as

$$\lambda_j = \lambda_{j-1} \left(Z\varepsilon \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{x_j^2}{2\sigma^2}} + (1-\varepsilon) \right)^{-1}, \tag{8.16}$$

where x_j is an error of j-th data point.

The term C defining the average number $C^{-1} \log A$ of observations carried out while testing a 'bad' model is derived, following equation (8.12), as follows

$$C = \int_0^Z p(x|H_b) \log \frac{p(x|H_b)}{p(x|H_g)} dx,$$

and finally

$$C = \frac{1}{Z} \int_0^Z -\log \left(Z\varepsilon \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{x^2}{2\sigma^2}} + (1-\varepsilon) \right) dx. \tag{8.17}$$

Unfortunately, the integral in (8.17) has to be either approximated or evaluated numerically. The rest of the derivation is identical with RANSAC.

8.3. R-RANSAC with SPRT

The R-RANSAC with SPRT algorithm is outlined in Alg. 8. To fully specify the details of the algorithm, two issues have to be addressed. First, the estimation of parameters δ and ε; second, the termination criterion guaranteeing $1-\eta$ confidence in the solution has to be derived.

Initialize ε_0, δ_0, calculate A_0 and set $i = 0$.

Repeat until the probability η (eq. (8.20)) of finding a model with support larger than $\hat{\varepsilon}$ falls under a user defined confidence value η_0:

1. Hypothesis generation

- Select a random sample of minimum size m from the set of data points.
- Estimate model parameters θ fitting the sample.

2. Verification

Execute the SPRT (Alg. 7) and update the estimates if

a Model rejected: re-estimate δ. If the estimate $\hat{\delta}$ differs from δ_i by more than 5% design (i+1)-th test ($\varepsilon_{i+1} = \varepsilon_i$, $\delta_{i+1} = \hat{\delta}$, $i = i+1$)

b Model accepted and the largest support so far: design (i+1)-th test ($\varepsilon_{i+1} = \hat{\varepsilon}$, $\delta_{i+1} = \hat{\delta}$, $i = i+1$). Store the current model parameters θ.

Algorithm 8: The structure of R-RANSAC with SPRT.

The algorithm proceeds like standard RANSAC [15, 22], only instead of checking all data points in the model verification step, the data points are evaluated sequentially and hypotheses with low support are rejected early. After a hypothesis is rejected, δ is re-estimated (Alg. 8, step

2a). Accepted hypotheses are candidates for the RANSAC outcome (see below). The overhead of the evaluation of the likelihood ratio λ_j eq. (8.10) is negligible compared to the evaluation of the model versus data point error function.

The optimal test derived in section 8.2 requires the knowledge of two parameters, ε and δ. These probabilities are different for different data sets and we assume they are unknown. The proposed algorithm uses values of ε and δ that are estimated during the sampling process and the test is adjusted to reflect the current estimates.

If the probabilities ε and δ are available a-priori, e.g. in some standard setting where the algorithm is run repeatedly, they can be used in the initialisation of the algorithm.

Estimation of δ. Since almost all tested models are 'bad'[5], the probability δ can be estimated as the average fraction of consistent data points in rejected models. When current estimate δ differs from the estimate used to design the SPRT (by more than 5%, for example), new (i+1)-th test is designed. The initial estimate δ_0 is obtained by geometric considerations, i.e. as a fraction of the area that supports a hypothesised model (a strip around an epipolar line in case of epipolar geometry) to the area of possible appearance of outlier data (the area of the search window). Alternatively, a few models can be evaluated without applying SPRT in order to obtain an initial estimate of δ.

Estimation of ε. In general, it is not possible to obtain an unbiased estimate of ε, since this would require the knowledge of the solution to the optimization problem we are solving. The tightest lower bound on ε is provided by the size of the largest support so far. It was shown in [34] that a sample with the largest support so far appears $\log k$ times, where k is the number of samples drawn. When such a sample (with support of size I_{i+1}) appears, new test is designed for $\varepsilon_{i+1} = I_{i+1}/N$. Throughout the course of the algorithm, a series of different tests with

$$\varepsilon_0 < \cdots < \varepsilon_i < \cdots < \varepsilon$$

are performed. The initial value of ε_0 can be derived from the maximal time the user is willing to assign to the algorithm.

The termination criterion. The algorithm is terminated, when the probability η of missing a set of inliers larger than the largest support found so far falls under a predefined threshold η_0. In standard RANSAC, where the probability of rejection of a 'good' model is zero, the probability is equal to

$$\eta_R = (1 - P_g)^k .$$

In R-RANSAC, the probability of hypothesising and not rejecting a 'good' model is $P_g(1-\alpha)$ and the probability η becomes as

$$\eta = (1 - P_g(1-\alpha))^k .$$

In R-RANSAC with SPRT, the SPRT is adjusted to current estimates of δ_i and ε_i, so α is no more constant. Theorem 7, which gives the probability α of rejecting a 'good' model for *the test designed for optimal value of* ε, does not cover this situation. The following theorem is needed:

[5] RANSAC verifies, on average, $-\log(\eta_0)$ 'good' models, e.g. for the typical $\eta_0 = 0.05$ a 'good' model is hypothesised three times prior to termination of the algorithm.

Theorem 9 *The probability of rejecting a 'good' model with fraction of inliers ε in a SPRT designed for ε_i and δ_i with threshold A_i is*

$$\alpha_i = A_i^{-h_i}, \tag{8.18}$$

where h_i is given by

$$\varepsilon \left(\frac{\delta_i}{\varepsilon_i}\right)^{h_i} + (1-\varepsilon)\left(\frac{1-\delta_i}{1-\varepsilon_i}\right)^{h_i} = 1. \tag{8.19}$$

Proof: For proof see [61, p. 50]. □

Equation (8.19) has two solutions, one being $h_i = 0$. Since $\varepsilon_i < \varepsilon$, $h_i > 1$ holds for other solution. This solution is found numerically.

Let for each of l tests the following values be stored: the expected fraction of inliers ε_i, the SPRT threshold A_i, the number of samples k_i processed by the test, and h_i satisfying (8.19). Then, the probability η is given by

$$\eta(l) = \prod_{i=0}^{l} \left(1 - P_g(1 - A_i^{-h_i})\right)^{k_i}. \tag{8.20}$$

The number k_l of samples that are needed to be drawn with current (i.e. l-th) SPRT follows from (8.20) as

$$k_l = \frac{\log \eta_0 - \log\left(\eta(l-1)\right)}{\log(1 - P_g A_l^{-1})} \tag{8.21}$$

Implementation note: since $\eta > \eta_R$ the equation (8.21) does not have to be evaluated before $\eta_R < \eta_0$ is satisfied.

8.4. Experiments

Several experiments were performed comparing the proposed R-RANSAC with SPRT with three other RANSAC algorithms: (1) standard RANSAC that verifies all correspondences for every model, (2) R-RANSAC with the $T_{d,d}$ test [34] that rejects the model when the first checked correspondence is not consistent with it ($d=1$), and (3) R-RANSAC with the a priori SPRT, i.e. the R-RANSAC with SPRT designed for the true values of ε and δ (labelled SPRT*). The results achieved with a priori SPRT show the best achievable performance of RANSAC with a randomized verification step for a problem characterized by given δ and ε.

For epipolar geometry estimation, the time needed to compute model parameters $t_M = 200$ was set within the range observed in a large number of experiments (i.e. in our implementation, checking whether a correspondence is consistent with a fundamental matrix is 200 times faster than estimating the matrix). The exact value depends on many factors including the CPU speed and type. The constant $\overline{m}_S = 2.38$ was set to the experimentally observed average of the number

of models generated by the 7-point algorithm per sample [6]. The initial values of δ and ε were set to $\delta_0 = 0.05$ and $\varepsilon = .2$ respectively.

For homography estimation, the values were set as follows $t_M = 200$, $\overline{m}_S = 1$, $\delta_0 = 0.01$, $\varepsilon_0 = 0.1$.

The experimental image pairs are displayed in Fig. 8.1. The number N of correspondences as well as the true values of ε and δ estimated by evaluation 100,000 verifications of random models are summarized in Tab. 8.2. The results of compared algorithms are shown in Table 8.1.

As a consequence of the randomization of model verification that erroneously rejects some 'good' models, on average, the randomized algorithms must draw a larger number of samples than standard RANSAC. This is confirmed in the first column of Table 8.1. This small increase is more than compensated by the reduction in the number of data points (correspondences) checked on average per model. The running time of RANSAC is reduced by factors ranging from 2.8 to 10.9 In all experiments the SPRT outperforms the $T_{d,d}$ test.

8.5. Conclusions

We have derived an optimal sequential strategy for randomised evaluation of model quality in RANSAC. A method for estimating the two probabilities characterising the problem and critically influencing the design of the optimal strategy was proposed. A termination criterion maintaining the $1 - \eta$ confidence in the solution was derived.

Properties of R-RANSAC with SPRT were tested on wide range of standard data. Tests included epipolar geometry estimation in both wide baseline setting and narrow baseline settings and homography estimation.

The method was 2.8 to 10 times faster than the standard RANSAC and up to 4 times faster than R-RANSAC with $T_{d,d}$ test.

[6] It is known that the 7-point algorithm produces 1 to 3 potential models. In experiments, the average number of models per sample – 2.38 – has been observed consistently in a number of scenes. No theoretical justification of the stability of this average is known to the authors.

The ROTUNDA A experiment					
	k	models	vpm	time	spd-up
RANSAC	195	470	886.0	4.3	1.0
$T_{d,d}$	345	834	49.0	1.4	3.1
SPRT*	201	486	29.1	0.8	5.4
SPRT	200	483	41.5	0.9	4.9

The GREAT WALL experiment					
	k	models	vpm	time	spd-up
RANSAC	48085	122223	507.0	667	1.0
$T_{d,d}$	224865	571059	6.9	642	1.0
SPRT*	56678	143993	14.4	178	3.8
SPRT	55796	141761	16.3	178	3.7

The CORRIDOR experiment					
	k	models	vpm	time	spd-up
RANSAC	145	344	600.0	2.3	1.0
$T_{d,d}$	243	576	86.6	1.2	2.0
SPRT*	154	366	32.9	0.6	3.7
SPRT	153	364	77.2	0.8	2.8

The LEUVEN experiment					
	k	models	vpm	time	spd-up
RANSAC	1855	4434	786.0	35.5	1.0
$T_{d,d}$	4490	10730	27.5	14.8	2.4
SPRT*	2025	4839	20.4	6.8	5.3
SPRT	1982	4736	23.2	6.8	5.2

The GRAFFITI experiment					
	k	models	vpm	time	spd-up
RANSAC	121	121	405.0	3.0	1.0
$T_{d,d}$	287	287	16.7	1.0	2.8
SPRT*	132	132	35.1	0.6	4.6
SPRT	129	129	47.1	0.7	4.1

The homography LEUVEN experiment					
	k	models	vpm	time	spd-up
RANSAC	1203	1203	789.0	53.1	1.0
$T_{d,d}$	5323	5323	3.4	15.0	3.5
SPRT*	1351	1351	14.5	4.7	11.2
SPRT	1266	1266	20.7	4.9	10.9

Table 8.1.: The comparison of RANSAC, R-RANSAC with $T_{d,d}$ test, a priori SPRT* and SPRT: the number of samples (k), the number of models (models), the number of checked correspondences per model (vpm), time in ms (time), and relative speed-up with respect to standard RANSAC (spd-up). Averaged over 500 runs.

	corr	ε	δ
ROTUNDA	893	0.60	0.052
THE GREAT WALL	514	0.28	0.014
CORRIDOR	607	0.56	0.142
LEUVEN	793	0.47	0.032
GRAFFITI	409	0.51	0.018
LEUVEN H	793	0.29	0.004

Table 8.2.: Number of correspondences (corr), fraction of inliers (ε), the probability of a correspondence being consistent with bad model (δ).

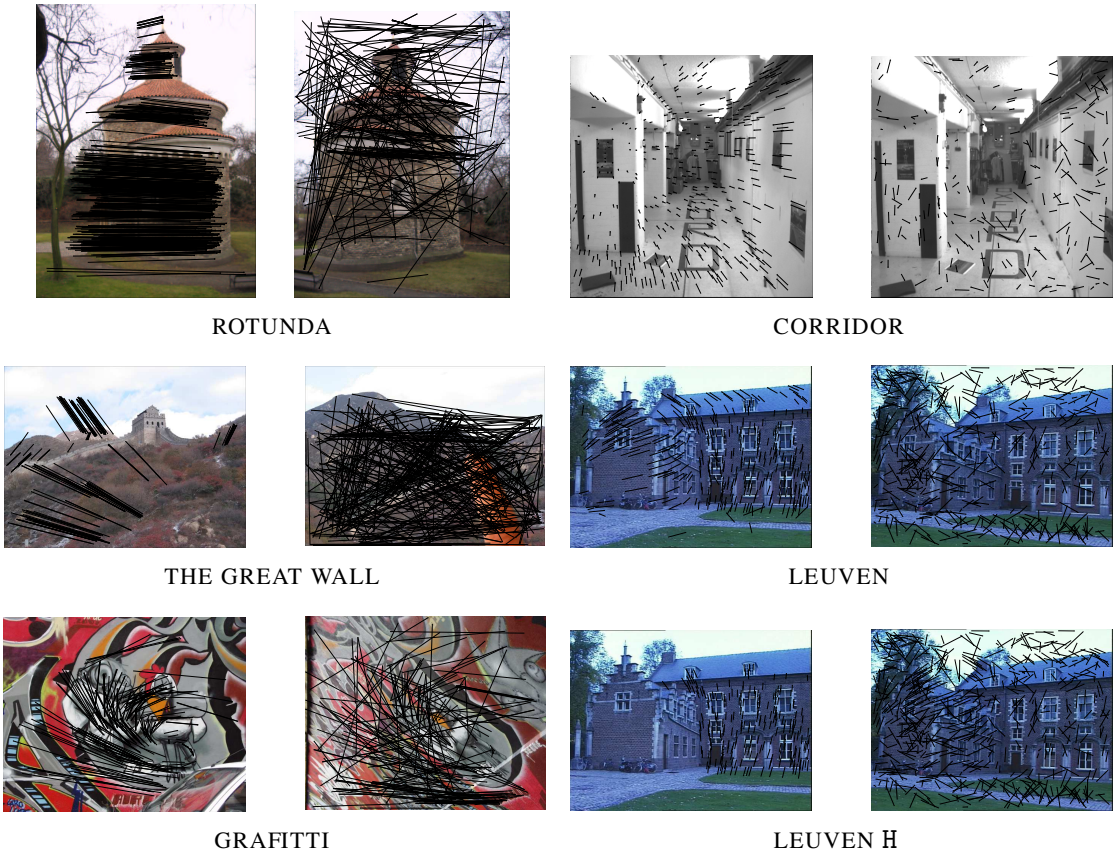


Figure 8.1.: The experimental image pairs with inliers (left) and outliers (right) superimposed. Two wide-baseline epipolar geometry experiments ROTUNDA and the GREAT WALL; two narrow-baseline EG experiments LEUVEN and CORRIDOR; two homography experiments GRAFITTI and LEUVEN H.

9 Oriented Constraints

Some applications impose additional constraints on the model parameters that do not reduce the model complexity (the number of data points that define the model 'uniquely'). For example, consider fitting circles of radius r, $r_{\min} \leq r \leq r_{\max}$, to points in a plane. One possible approach would be to modify the hypothesis generator (e.g. by using some additional data structures) to generate only allowed hypotheses. However, this might render the hypothesis generation time demanding or might be even intractable. Another, rather passive, option is just to ignore the additional constraint and count on the fact that the models not satisfying the constraint are arbitrary and will not have sufficient support. This is often true and therefore an incorrect solution that does not satisfy the additional constrains is not observed as a problem. Nevertheless, such an approach may waste a significant proportion of time verifying many (even a majority) of samples that do not satisfy the additional constraints. This can be avoided in a very simple way – by testing the validity of the parameters before the verification step. In this chapter, it is shown that checking the orientation constraint often saves significant amount of computational time compared with the standard approaches that ignore those constraints.

9.1. Oriented Projective Geometry

For "real" cameras (i.e. physical devices), only points in front of the camera are visible. This is modeled in the framework of the oriented projective geometry [52], where cameras form images by projecting along *half-lines* emanating from a projection center. Points in two views taken by a camera satisfy, besides the epipolar constraint, some additional constraints [29, 62]. The constraints have been used before for outlier removal after the epipolar geometry was recovered [29], but not directly in the process of epipolar geometry estimation.

In this chapter we show that the use of oriented constraints within RANSAC brings significant computational savings. The approach has no negative side-effects, such as limited applicability or poor worst-case performance, and thus should become a part of any state-of-the-art RANSAC implementation of epipolar geometry estimation.

In RANSAC, epipolar geometry is obtained by repeating a hypothesize-and-verify loop. If the hypothesized epipolar geometry violates the oriented constraint, the verification step does not have to be carried out. Since the orientation test takes negligible time compared to both the epipolar geometry computation and the verification, a speed-up may be achieved virtually for free. The overall impact of the orientation constraint on RANSAC depends on the fraction of hypothesized models that can be rejected, without verification, solely on the basis of failing the orientation constraint test. We empirically measure this fraction in a number of real scenes, both

in a wide-baseline and narrow-baseline stereo settings. The performance evaluation carried out shows that the the impact of exploiting the orientation constraint on RANSAC running time is in many circumstances significant.

The rest of the chapter is structured as follows. First, the derivation of the oriented epipolar constraint is reviewed in section 9.2. RANSAC with the orientation constraint is introduced in section 9.3. Next, in section 9.4, performance of the improved RANSAC constraint is evaluated. Two quantities are measured: the fraction of hypothesis that fail the orientation test and the time saved as a consequence of inserting the orientation test into the RANSAC loop.

9.2. Oriented Epipolar Constraint

Let[1] a camera with 3×4 projection matrix $\mathtt{P}$ observe a rigid scene. An image point represented by a homogeneous 3-vector $\mathbf{x}$ is a projection of a scene point represented by a homogeneous 4-vector $\mathbf{X}$ if and only if $\mathbf{x} \sim \mathtt{P}\mathbf{X}$ [20].

Following the classical (i.e. unoriented) projective geometry, the homogeneous quantities $\mathbf{x}$, $\mathbf{X}$, and $\mathtt{P}$ represent the same geometric objects if multiplied by a non-zero scale. E.g., the homogeneous vectors $\mathbf{x} = (x, y, 1)^\top$ and $-\mathbf{x} = (-x, -y, -1)^\top$ represent the same image point with affine coordinates $(x, y)^\top$.

It has been noticed that the *oriented projective geometry* [52, 28] is a more appropriate model for multiple view geometry as it can represent ray orientations. In oriented geometry, vectors $\mathbf{x}$ and $-\mathbf{x}$ represent two different image points, differing by whether the corresponding scene point lies in front of or behind the camera. The (oriented) relation between scene point $\mathbf{X}$ and its image $\mathbf{x}$ is $\mathbf{x} \overset{+}{\sim} \mathtt{P}\mathbf{X}$, as opposed to unoriented relation $\mathbf{x} \sim \mathtt{P}\mathbf{X}$.

The orientation of image points is known from the fact that all visible points lie in front of the camera [28]. Formally, image points lie on the positive side of the image line at infinity $\mathbf{l}_\infty$. For the usual choice $\mathbf{l}_\infty = (0, 0, 1)^\top$, the correctly oriented homogeneous vector representing an image point with affine coordinates $(x, y)^\top$ is $(x, y, 1)^\top$ or its positive multiple.

Let two cameras with projection matrices $\mathtt{P}$ and $\mathtt{P}'$ observe a rigid scene. It is well-known [20] that there exists a 3×3 *fundamental matrix* $\mathtt{F}$ of rank 2 such that any pair $\mathbf{x} \leftrightarrow \mathbf{x}'$ of corresponding image points satisfies the *epipolar constraint*

$$\mathbf{x}'^\top \mathtt{F}\mathbf{x} = 0. \tag{9.1}$$

The oriented version of the epipolar constraint is [62]

$$\mathbf{e}' \times \mathbf{x}' \overset{+}{\sim} \mathtt{F}\mathbf{x}. \tag{9.2}$$

It is implied by the following lemma (we omit the proof).

Lemma 1 *Any 3×4 full-rank matrices $\mathtt{P}$ and $\mathtt{P}'$ and a 4-vector $\mathbf{X}$ satisfy*

$$\mathbf{e}' \times (\mathtt{P}'\mathbf{X}) = \mathtt{F}(\mathtt{P}\mathbf{X}), \tag{9.3}$$

[1] In this section, $\mathbf{a} \sim \mathbf{b}$ denotes equality of two vectors up to a non-zero scale and $\mathbf{a} \overset{+}{\sim} \mathbf{b}$ equality up to a positive scale. Vector product of two 3-vectors is $\mathbf{a} \times \mathbf{b}$. Symbol $[\mathbf{a}]_\times$ denotes the matrix such that $[\mathbf{a}]_\times \mathbf{b} = \mathbf{a} \times \mathbf{b}$. Matrix pseudoinverse is denoted $\mathtt{P}^+$ and vector norm $\|\mathbf{a}\|$.

	corrs	inliers	ε [%]
Juice	447	274	61.30
Shelf	126	43	34.13
Valbonne	216	42	19.44
Great Wall	318	68	21.38
Leuven	793	379	47.79
Corridor	607	394	64.91

Table 9.1.: The number of correspondences ('corrs'), inliers ('inliers') and the fraction of inliers ('ε') in the experiments.

where $\mathtt{F} = [\mathbf{e}']_\times \mathtt{P}'\mathtt{P}^+$, $\mathbf{e}' = \mathtt{P}'\mathbf{C}$, *and* $\mathbf{C}$ *is uniquely determined by equations* $\mathtt{P}\mathbf{C} = 0$ *and* $\det(\mathtt{P}^\top \mid \mathbf{C}) = \|\mathbf{C}\|^2$.

Note that (9.1) is invariant to the change of signs of $\mathbf{x}$, $\mathbf{x}'$ and $\mathtt{F}$ whereas (9.2) is not. Therefore (9.2) implies (9.1) (multiply (9.2) by $\mathbf{x}'^\top$ from the left), but not *vice versa*. Thus, the oriented epipolar constraint is stronger than the unoriented one.

9.3. RANSAC with Oriented Constraint

The standard seven-point algorithm [20] is used to hypothesize the fundamental matrix. The two-dimensional space of 3×3 matrices satisfying (9.1) for the seven sampled correspondences is found by QR factorization rather than SVD, as suggested in [40]. Each fundamental matrix is then tested whether it satisfies the oriented constraint (9.2). This is the only step in which the new algorithm differs form the standard one. The test can be performed very efficiently, requiring only 27 – 81 floating point operations (i.e. multiplications and additions). If the orientation test is passed, the support of the fundamental matrix is computed as the number of correspondences with Sampson's distance [20] below threshold.

9.4. Experiments

The RANSAC algorithm with the oriented constraints was tested on six standard image pairs, including wide-baseline stereo (experiments Juice, Shelf, Valbonne and the Great Wall) and narrow-baseline stereo (Leuven and Corridor).

Obtaining tentative correspondences. By a tentative correspondence we mean a pair of points $\mathbf{x} \leftrightarrow \mathbf{x}'$, where $\mathbf{x}$ is from the first image and $\mathbf{x}'$ is from second image. The set of tentative correspondences contains both inliers and outliers.

The tentative correspondences for the wide-baseline experiments were obtained automatically by matching normalized local affine frames [36]. Only mutually best candidates were selected as tentative correspondences.

In the narrow-baseline experiments, the Harris operator [18] was used to detect interest points. Point pairs with mutually maximal normalised cross-correlation of rectangular windows around

	models	rejected	passed [%]
Juice	1,221,932	760,354	37.77
Shelf	1,233,770	1,094,533	11.29
Valbonne	1,256,648	1,176,042	6.41
Great Wall	1,274,018	1,181,084	7.29
Leuven	1,194,238	336,515	71.82
Corridor	1,187,380	120,916	89.82

Table 9.2.: The number of fundamental matrices generated by RANSAC ('models') over 500,000 samples, the number of models rejected by the orientation constraint ('rejected') and the percentage of models that passed the test ('passed'). Note that regardless of the setting, there is always approximately 2.4 fundamental matrices per sample on average.

interest points were kept as tentative correspondences. The size of the used windows were 15 and 7 pixels for Leuven and Corridor, respectively. The proximity constraint, ensuring that the coordinates of corresponding points would not differ by more than 100 and 30 pixels respectively, was also used. The numbers of tentative correspondences and the numbers of inliers for each experiment are summarized in Tab. 9.1. The image pairs, with inliers and outliers superimposed, are depicted in Fig. 9.1.

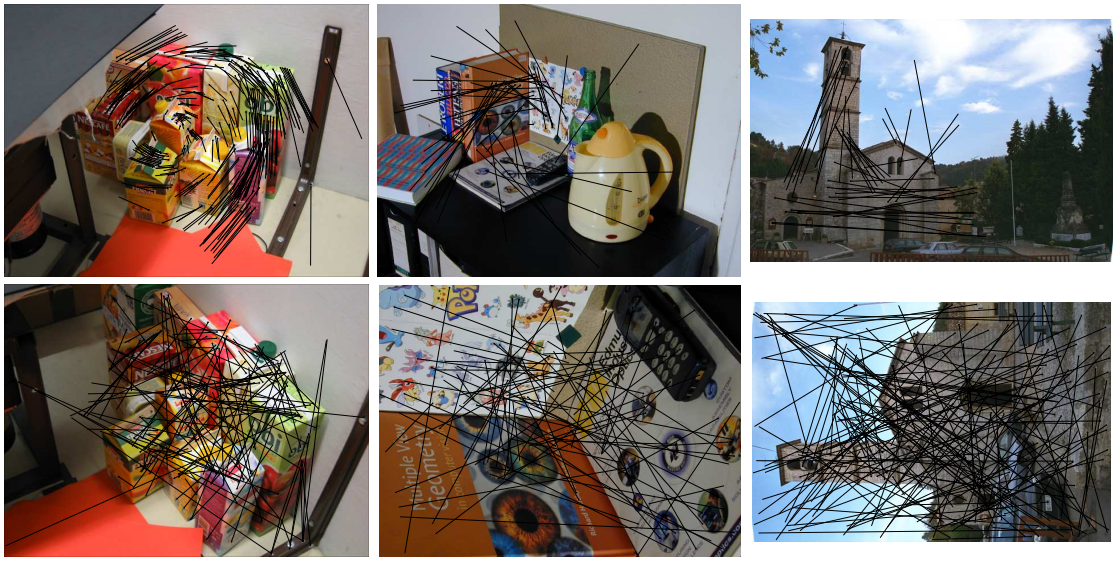
The fraction of rejected models. The number of hypotheses that can be rejected on the basis of the orientation constraint was measured. The results of the experiment are summarized in Tab. 9.2. A sample of seven correspondences was selected at random 500,000 times. The seven-point algorithm produces 1 to 3 fundamental matrices satisfying the unoriented epipolar constraint (9.1). The total number of the models is given in the 'models' column of Tab. 9.2. The number of models that are rejected by the orientation constraint (9.2) is shown in the 'rejected' column.

The fraction of rejected models varies widely. What affects the fraction of hypothesis that can be rejected solely based on the orientation constraint? This question goes far beyond the scope of this chapter. From the results of the experiments we observed that more models were rejected in the wide-baseline setting than in the narrow-baseline one. We believe this is due to different distribution of outliers which is caused by limited correspondence search window in the narrow-baseline case. The fraction of rejected models is proportional to the fraction of outliers among the tentative correspondences.

Running time. The time saved by using the oriented epipolar constraint was measured. The standard and oriented versions of RANSAC were executed 500 times and their running-time was recorded (on a PC with K7 2200+ processor). To ensure that both methods draw the same samples, the pseudo-random generator was initialized by the same seed.

9.5. Conclusions

In this chapter, RANSAC enhanced by the oriented epipolar constraint was experimentally evaluated. The application of the oriented constraint reduced the running time by 5% to 46%, com-



Juice Shelf Valbonne

Great Wall Leuven Corridor

Figure 9.1.: The experimental settings. Inliers and outliers are superimposed over the first and second images respectively.

	standard	oriented	speed-up [%]
Juice	1.4	0.9	35.08
Shelf	53.6	45.1	15.82
Valbonne	930.8	584.8	37.17
Great Wall	1109.0	599.3	45.96
Leuven	18.5	15.1	18.52
Corridor	1.2	1.2	5.77

Table 9.3.: Time (in ms) spent in standard and oriented versions of RANSAC and the relative speed-up (right column).

pared with standard RANSAC. The efficiency increase of RANSAC with the orientation constraint is achieved by reducing the number of verification steps. As a consequence, the more time-consuming the verification step is the higher relative speed-up is achieved. This applies not only to situations with large number of correspondences, but also to RANSAC-type algorithms that perform expensive verification procedures. An example of such an algorithm is MLESAC [57] which estimates the parameters of a mixture of inlier and outlier distributions. Since the evaluation of the orientation test takes virtually no time compared to the epipolar geometry computation, the epipolar geometry estimation via RANSAC should exploit the oriented epipolar constraint.

10 Conclusions

In the thesis, a robust estimator RANSAC was thoroughly analyzed. The thesis contributed to the state of the art in robust estimation of geometric models by several novel algorithms derived from RANSAC. These algorithms are addressing the stability, the sensitivity to the noise, and the computational complexity of the estimation process in two-view geometry. In particular,

- A known discrepancy (caused by noise on inliers) between theoretical prediction of the time required to find the solution and practically observed running times was traced to a tacit assumptions of RANSAC. The algorithm is modified to reach almost perfect agreement with theoretical predictions without any negative impact on the time complexity by applying local optimization to the-best-so-far samples. It was shown that the computational time is significantly reduced, when approximate models are hypothesized and coarse to fine strategy is used in the local optimization.

- An unified approach estimating model and its degenerate configuration (epipolar geometry and homography of a dominant plane) at the same time without a priori knowledge of the presence of the degenerate configuration (dominant plane) was derived.

- An algorithm, called PROSAC, exploiting (possibly noisy) match quality to modify the sampling strategy was introduced. The quality of a match is an often freely available quantity in the matching problem. The approach increased the efficiency of the proposed algorithm while keeping the same robustness as RANSAC in the worst-case situation (when the match quality is unrelated to whether a correspondence is a mismatch or not).

- The concept of randomized cost function evaluation in RANSAC was introduced. Superiority of randomized RANSAC (R-RANSAC) over the RANSAC with deterministic evaluation was shown.

- It was shown that using oriented geometric constraints that arise from a realistic model of physical camera devices saves non-negligible fraction of computational time, without any negative side effect.

Most of the algorithms in the thesis were motivated by (and presented on) estimation of a multi-view geometry. The algorithms are, however, general robust estimation techniques and can be easily used in other application areas as well. Moreover, all algorithms proposed in the thesis can, and are recommended, to use together, as no initial assumptions of any two of the algorithms are mutually exclusive.

Bibliography

[1] A. Baumberg. Reliable feature matching across widely separated views. In *CVPR00*, pages I:774–781, 2000.

[2] M. Brown and D. Lowe. Recognising panoramas. In *Proc. ICCV03*, volume I, pages 1218–1225, October 2003.

[3] O. Chum, J. Matas, and J. Kittler. Locally optimized RANSAC. In *Proc. DAGM*. Springer-Verlag, 2003.

[4] O. Chum, T. Pajdla, and P. Sturm. The geometric error for homographies. *Computer Vision and Image Understanding*, To appear 2004.

[5] Ondřej Chum and Jiří Matas. Randomized RANSAC with T(d,d) test. In *Proc. of BMVC*, volume 2, pages 448–457. BMVA, 2002.

[6] Ondřej Chum and Jiří Matas. Matching with PROSAC - progressive sampling concensus. In *Proc. of the CVPR*, pages –, June 2005.

[7] Ondřej Chum, Jiří Matas, and Štěpán Obdržálek. Enhancing RANSAC by generalized model optimization. In *Proc. of the ACCV*, volume 2, pages 812–817, January 2004.

[8] Ondřej Chum, Tomáš Werner, and Jiří Matas. Epipolar geometry estimation via RANSAC benefits from the oriented epipolar constraint. In *Proc. of ICPR 2004*, volume 1, pages 112–115. IEEE CS, August 2004.

[9] Ondřej Chum, Tomáš Werner, and Jiří Matas. Epipolar geometry estimation unaffected by the dominant plane. In *Proc. of the CVPR*, June 2005.

[10] Ondřej Chum, Tomáš Werner, and Tomáš Pajdla. Joint orientation of epipoles. In Richard Harvey, editor, *BMVC 2003: Proceedings of the 14th British Machine Vision Conference*, volume 1, pages 73–82, London, UK, September 2003. BMVA, BMVA.

[11] J. Clarke, S. Carlsson, and A. Zisserman. Detecting and tracking linear features efficiently. In *Proc. 7th BMVC*, pages 415–424, 1996.

[12] Hugo Cornelius, Radim Šára, Daniel Martinec, Tomáš Pajdla, Ondřej Chum, and Jiří Matas. Towards complete free-form reconstruction of complex 3D scenes from an unordered set of uncalibrated images. In D. Comaniciu, R. Mester, and K. Kanatani, editors, *Proc ECCV Workshop Statistical Methods in Video Processing*, volume LNCS 3247, pages 1–12, Heidelberg, Germany, May 2004. Springer-Verlag.

[13] H. S. M. Coxeter. *The Real Projective Plane : with an Appendix for Mathematica by George Beck*. Springer-Verlag, New York, USA, 3rd edition, 1992.

[14] A.P. Dempster, N.M Laird, and D.B Rubin. Maximum likelihood from incoplete data via the EM algorithm. *Journal of the Royal Statictical Society*, 39:185–197, 1977.

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

[16] A. W. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In *Proc. of CVPR*, volume 1, pages 125–132, 2001.

[17] J. Flusser and T. Suk. Pattern recognition by affine moment invariants. *Journal of the Pattern Recognition*, 26(1):167–174, 1993.

[18] C. J. Harris and M. Stephens. A combined corner and edge detector. In *Proc. of Alvey Vision Conference*, pages 147–151, 1988.

[19] R. Hartley. In defence of the 8-point algorithm. In *ICCV95*, pages 1064–1070, 1995.

[20] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge, UK, 2000.

[21] R.I. Hartley and P. Sturm. Triangulation. In *ARPA94*, pages II:957–966, 1994.

[22] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*. Cambridge University, Cambridge, 2nd edition, 2003.

[23] P.V.C. Hough. Method and means for recognizing complex patterns. U.S. Patent 3069654, 1962.

[24] P. J. Huber. *Robust Statistics*. John Willey and Sons, New York, 1nd edition, 1981.

[25] J. Illingworth and J. Kittler. A survey of the Hough transform. *Computer Vision, Graphics and Image Processing*, 44:87–116, 1998.

[26] Michal Irani and P. Anandan. Parallax geometry of pairs of points for 3d scene analysis. In *ECCV (1)*, pages 17–30, 1996.

[27] Heikki Kälviäinen, P. Hirvonen, Lei Xu, and E. Oja. Probabilistic and non-probabilistic hough transforms: Overview and comparisons. *IVC*, 13(4):239–252, May 1995.

[28] S. Laveau and O.D. Faugeras. Oriented projective geometry for computer vision. In *Proc. of ECCV*, volume I, pages 147–156, 1996.

[29] Stéphane Laveau and Olivier Faugeras. 3-D scene representation as a collection of images. In *ICPR '94*, pages 689–691, October 9–13 1994.

[30] Peter M. Lee. Sequential probability ratio test. University of York. www.york.ac.uk/depts/maths/teaching/pml/ais/sprt.ps.

[31] A. Leonardis, H. Bischof, and R. Ebensberger. Robust recognition using eigenimages. Technical Report TR-47, PRIP, TU Wien, 1997.

[32] D.G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[33] Jiří Matas, Petr Bílek, and Ondřej Chum. Rotational invariants for wide-baseline stereo. In *Proc. of CVWW'02*, pages 296–305, February 2002.

[34] Jiří Matas and Ondřej Chum. Randomized RANSAC with $T_{d,d}$ test. *Image and Vision Computing*, 22(10):837–842, September 2004.

[35] Jiří Matas, Ondřej Chum, Martin Urban, and Tomáš Pajdla. Robust wide-baseline stereo from maximally stable extremal regions. *Image and Vision Computing*, 22(10):761–767, Sep 2004.

[36] Jiří Matas, Štěpán Obdržálek, and Ondřej Chum. Local affine frames for wide-baseline stereo. In *Proc. ICPR*, volume 4, pages 363–366. IEEE CS, Aug 2002.

[37] P. McLauchlan and A. Jaenicke. Image mosaicing using sequential bundle adjustment. In *Proc. BMVC*, pages 616– 62, 2000.

[38] K. Mikolajczyk and C. Schmid. An affine invariant interest point detector. In *Proc. ECCV*, volume 1, pages 128–142, 2002.

[39] D.R. Myatt, P.H.S. Torr, S.J. Nasuto, J.M. Bishop, and R. Craddock. Napsac: High noise, high dimensional robust estimation - it's in the bag. In *BMVC02*, volume 2, pages 458–467, 2002.

[40] D. Nister. An efficient solution to the five-point relative pose problem. In *Proc. of CVPR*, volume II, pages 195–202, 2003.

[41] D. Nister. Preemptive RANSAC for live structure and motion estimation. In *Proc. ICCV03*, volume I, pages 199–206, October 2003.

[42] Štěpán Obdržálek and Jiří Matas. Image retrieval using local compact DCT-based representation. In *Proc. of DAGM'03*, number 2781 in LNCS, pages 490–497. Springer-Verlag, 9 2003.

[43] Athanasios Papoulis and S. Unnikrishna Pillai. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill Series in Electrical and Computer Engineering. McGraw-Hill, Boston, USA, 4th edition, 2002.

[44] Marc Pollefeys, Frank Verbiest, and Luc van Gool. Surviving dominant planes in uncalibrated structure and motion recovery. In *Proc. ECCV*, volume 1, pages 837–851. Springer-Verlag, 2002.

[45] P. Pritchett and A. Zisserman. Wide baseline stereo matching. In *Proc. ICCV*, pages 754–760, 1998.

[46] Peter J. Rousseeuw and Annick M. Leroy. *Robust Regression and Outlier Detection*. Wiley, 1987.

[47] F. Schaffalitzky and A. Zisserman. Viewpoint invariant texture matching and wide baseline stereo. In *Proc. 8th ICCV, Vancouver, Canada*, July 2001.

[48] M.I. Schlesinger. A connection between learning and self-learning in the pattern recognition (in Russian). *Kibernetika*, 2:81–88, 1968.

[49] C. Schmid and R. Mohr. Local grayvalue invariants for image retrieval. *PAMI*, 19(5):530–535, May 1997.

[50] J. Sivic and A. Zisserman. Video Google: A text retrieval approach to object matching in videos. In *Proceedings of the International Conference on Computer Vision*, pages 1470 – 1477, October 2003.

[51] Charles V. Stewart. MINPRAN: a new robust estimator for computer vision. *PAMI*, 17(10):925–938, October 1995.

[52] Jorge Stolfi. *Oriented Projective Geometry: A Framework for Geometric Computations*. Academic Press, Inc., 1250 Sixth Avenue, San Diego, CA 92101, 1991.

[53] B. Tordoff and D.W. Murray. Guided sampling and consensus for motion estimation. In *Proc. 7th ECCV*, volume 1, pages 82–96. Springer-Verlag, 2002.

[54] P. H. S. Torr. *Outlier Detection and Motion Segmentation*. PhD thesis, Dept. of Engineering Science, University of Oxford, 1995.

[55] P. H. S. Torr. An assessment of information criteria for motion model selection. In *Proc. CVPR*, pages 47–53. IEEE, 1997.

[56] P. H. S. Torr, A. W. Fitzgibbon, and A. Zisserman. The problem of degeneracy in structure and motion recovery from uncalibrated image sequences. *International Journal of Computer Vision*, 32(1):27–44, aug 1999.

[57] P. H. S. Torr and A. Zisserman. MLESAC: A new robust estimator with application to estimating image geometry. *CVIU*, 78:138–156, 2000.

[58] P.H.S. Torr, A. Zisserman, and S.J. Maybank. Robust detection of degenerate configurations while estimating the fundamental matrix. *CVIU*, 71(3):312–333, Sep. 1998.

[59] Bill Triggs, Philip McLauchlan, Richard Hartley, and Andrew Fitzgibbon. A comprehensive survey of bundle adjustment in computer vision. In *Proc. Vision Algorithms: Theory and Practice. International Workshop on Vision Algorithms*, number 1883 in LNCS, pages 298–372. Springer Verlag, 1999.

[60] Tinne Tuytelaars and Luc Van Gool. Wide baseline stereo matching based on local, affinely invariant regions. In *Proc. 11th BMVC*, 2000.

[61] Abraham Wald. *Sequential analysis*. Dover, New York, 1947.

[62] Tomáš Werner and Tomáš Pajdla. Oriented matching constraints. In *Proc. of BMVC*, pages 441–450. BMVA, September 2001.

[63] Lei Xu, E. Oja, and P. Kultanen. A new curve detection method: Randomized hough transform (rht). *Pattern Recognition Letters*, 11:331–338, 1990.

[64] Zhengyou Zhang, Rachid Deriche, Olivier Faugeras, and Quang-Tuan Luong. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry,. *Artificial Intelligence, December 1995*, 78:87–119, 1995.

A The Number of Points Epipolar Geometry Is Estimated From

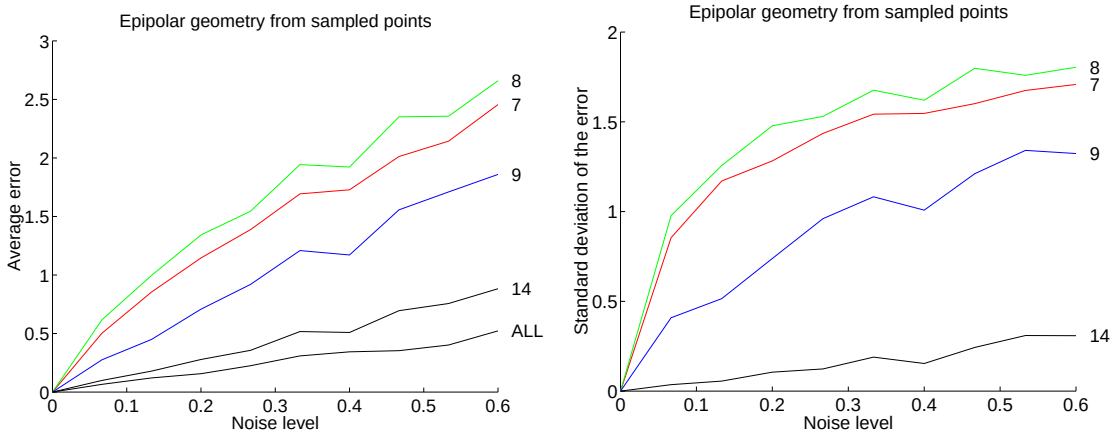


Figure A.1.: The average error (left) and the standard deviation of the error for samples of 7,8,9, 14 and all 100 points respectively with respect to the noise level.

An experiment testing the dependence of the quality of the epipolar geometry on the number of points (and the algorithm) used to estimate a fundamental matrix is presented. It has been mentioned in [54], that the fundamental matrix estimated from a seven point sample is more precise than the one estimated form eight points using a linear algorithm [19]. The difference stems in the difference in the seven-point and the eight-point algorithm. The seven point algorithm calculates directly a rank defficient 3-by-3 matrix $\mathtt{F}$ that satisfies the epipolar constraints $\mathbf{x}_i'^{\top}\mathtt{F}\mathbf{x}_i = 0$ for seven given correspondences $\{\mathbf{x}_i \leftrightarrow \mathbf{x}\}_{i=1}^{7}$. The eight-point algorithm finds in general regular matrix $\mathtt{G}$ (which is not a fundamental matrix) that satisfies $\mathbf{x}_i'^{\top}\mathtt{G}\mathbf{x}_i = 0$ for the eight given correspondences. To obtained a (rank deficient) fundamental matrix, the matrix $\mathtt{G}$ is projected into the space of singular matrices.

The following experiment shows that the observation holds only for eight point samples and that by increasing the number of points gives more stable results than those obtained when the fundamental matrix is computed from seven points only.

Experiment: This experiment shows, how the quality of a hypothesis depends on the number of correspondences used to calculate the fundamental matrix. For seven points, the seven point algorithm was used [54] and for eight and more points the linear algorithm [19] was used. The course of experiment was as follows. Noise of different levels was added to the noise-free image points correspondences divided into two sets of hundred correspondences. Samples of different

sizes were drawn from the first set and the average error over the second was computed. This was repeated 1000 times for each noise level. Results are displayed in Figure A.1.

This experiment suggests, that the more points are used to estimate the model (in this case fundamental matrix) the more precise solution is obtained (with the exception of eight points). The experiment also shows that the minimal sample gives hypotheses of rather poor quality. One can use different cost functions that are more complicated than simply the number of inliers, but evaluating this function only at parameters arising from the minimal sample will get results at best equal to the proposed method of local optimization.