RANSAC for Dummies

With examples using the RANSAC toolbox for Matlab™ & Octave and more...

Marco Zuliani
marco.zuliani@gmail.com

vision.ece.ucsb.edu/~zuliani

November 19, 2011

To all the free thinkers,
who freely share their ideas.

Contents

1 Introduction **7**

2 Parameter Estimation In Presence of Outliers **8**

2.1 A Toy Example: Estimating 2D Lines 8

2.1.1 Maximum Likelihood Estimation 9

2.2 Outliers, Bias and Breakdown Point 11

2.2.1 Outliers 12

2.2.2 Bias 12

2.2.3 Breakdown Point 13

2.3 The Breakdown Point for a 2D Line Least Squares Estimator 13

3 RANdom Sample And Consensus **15**

3.1 Introduction 15

3.2 Preliminaries 16

3.3 RANSAC Overview 18

3.3.1 How many iterations? 19

3.3.2 Constructing the MSSs and Calculating q 20

3.3.3 Ranking the Consensus Set 23

3.4 Computational Complexity 25

3.4.1 Hypothesize Step 25

3.4.2 Test Step 25

3.4.3 Overall Complexity 25

3.5 Other RANSAC Flavors 25

4 RANSAC at Work **28**
4.1 The RANSAC Toolbox for Matlab™ & Octave . . . 28
4.1.1 `RANSAC.m` . . . 28
4.2 Some Examples Using the RANSAC Toolbox . . . 32
4.2.1 Estimating Lines . . . 32
4.2.2 Estimating Planes . . . 35
4.2.3 Estimating a Rotation Scaling and Translation . . . 36
4.2.4 Estimating Homographies . . . 40
4.3 Frequently Asked Questions . . . 44
4.3.1 What is the "right" value of σ? . . . 44
4.3.2 I want to estimate the parameters of my favourite model. What should I do? . . . 44
4.3.3 How do I use the toolbox for image registration purposes? . . . 44
4.3.4 Why the behaviour of RANSAC is not repeatable? . . . 45
4.3.5 What should I do if I find a bug in the toolbox? . . . 45
4.3.6 Are there any other RANSAC routines for Matlab? . . . 45

A Notation **46**

B Some Linear Algebra Facts **47**
B.1 The Singular Value Decomposition . . . 47
B.2 Relation Between the SVD Decomposition and the Eigen Decomposition 48
B.3 Fast Diagonalization of Symmetric 2×2 Matrices . . . 49
B.4 Least Square Problems Solved via SVD . . . 50
B.4.1 Solving $A\boldsymbol{\theta} = \boldsymbol{b}$ . . . 50
B.4.2 Solving $A\boldsymbol{\theta} = \mathbf{0}$ subject to $\|\boldsymbol{\theta}\| = 1$ . . . 51

C The Normalized Direct Linear Transform (nDLT) Algorithm **53**
C.1 Introduction . . . 53
C.2 Point Normalization . . . 54
C.3 A Numerical Example . . . 58
C.4 Concluding Remarks About the Normalized DLT Algorithm . . . 60

D Some Code from the RANSAC Toolbox **65**
D.1 Function Templates . . . 65
D.1.1 MSS Validation . . . 65

D.1.2 Parameter Estimation 66
D.1.3 Parameter Validation 68
D.1.4 Fitting Error 69
D.2 Source Code for the Examples 71
D.2.1 Line Estimation 71
D.2.2 Plane Estimation 74
D.2.3 RST Estimation 78
D.2.4 Homography Estimation 83

E GNU Free Documentation License **94**
1. APPLICABILITY AND DEFINITIONS 94
2. VERBATIM COPYING 95
3. COPYING IN QUANTITY 95
4. MODIFICATIONS 95
5. COMBINING DOCUMENTS 96
6. COLLECTIONS OF DOCUMENTS 97
7. AGGREGATION WITH INDEPENDENT WORKS 97
8. TRANSLATION 97
9. TERMINATION 97
10. FUTURE REVISIONS OF THIS LICENSE 97
ADDENDUM: How to use this License for your documents 97
References 98

1 Introduction

This tutorial and the toolbox for Matlab™ & Octave were mostly written during my spare time (with the loving disapproval of my wife), starting from some routines and some scattered notes that I reorganized and expanded after my Ph.D. years. Both the tutorial and the toolbox are supposed to provide a simple and quick way to start experimenting the RANSAC algorithm utilizing Matlab™ & Octave .

The notes may seem somewhat heterogeneous, but they collect some theoretical discussions and practical considerations that are all connected to the topic of robust estimation, more specifically utilizing the RANSAC algorithm.

Despite the fact that several users tested this package and sent me their invaluable feedback, it is possible (actually very probable) that these notes still contain typos or even plain mistakes. Similarly, the RANSAC toolbox may contain all sorts of bugs. This is why I really look forward to receive your comments: compatibly with my other commitments I will try to improve the quality of this little contribution in the fervent hope that somebody might find it useful.

I want to thank you here all the persons that have been intensively using the toolbox and provided me with precious suggestions, in particular Dong Li, Tamar Back, Frederico Lopes, Jayanth Nayak, David Portabella Clotet, Chris Volpe, Zhe Zang, Ali Kalihili, George Polchin.

Los Gatos, CA
November 2011

Marco Zuliani

2 Parameter Estimation In Presence of Outliers

This chapter introduces the problem of parameter estimation when the measurements are *contaminated by outliers*. To motivate the results that will be presented in the next chapters and to understand the power of RANSAC, we will study a simple problem: fitting a 2D line to a set of points on the plane. Despite its simplicity, this problem retains all the challenges that are encountered when the models used to explain the measurements are more complex.

2.1 A Toy Example: Estimating 2D Lines

Consider a set of N points $D = \{\boldsymbol{d}_1, \ldots, \boldsymbol{d}_N\} \subset \mathbb{R}^2$ and suppose we want to estimate the best line that fits such points.[1] For each point we wish to minimize *a monotonically increasing function* of the *absolute* value of the *signed* error:

$$e_{\mathcal{M}}(\boldsymbol{d}; \boldsymbol{\theta}) = \frac{\theta_1 x_1 + \theta_2 x_2 + \theta_3}{\sqrt{\theta_1^2 + \theta_2^2}} \qquad (2.1)$$

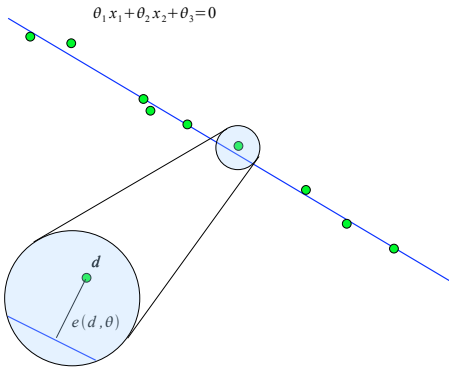


Figure 2.1: Line fitting example.

The sign of the error (2.1) accounts for the fact that the point lies either on the left or on the right semi-plane determined by the line. The parameter vector $\boldsymbol{\theta} \in \mathbb{R}^2$ describes the line according the implicit representation $\theta_1 x_1 + \theta_2 x_2 + \theta_3 = 0$ (this is the model $\mathcal{M}$ that we will

[1] Further details regarding the estimation of 2D lines can be found in Section 4.2.1.

use to fit the measurements). Note that the length of $\boldsymbol{\theta}$ is immaterial. This type of fitting is also known as *orthogonal regression*, since the distances of the sample points from the line are evaluated computing the orthogonal projection of the measurements on the line itself. Other type of regression can also be used, e.g. minimizing the distance of the projection of the measurements along the y axis (however such an approach produces an estimate of the parameters that is not invariant with respect a rotation of the coordinate system).

2.1.1 Maximum Likelihood Estimation

Imagine that the fitting error is modeled as a Gaussian random variable with zero mean and standard deviation σ_η, i.e. $e_{\mathcal{M}}(\boldsymbol{d};\boldsymbol{\theta}) \sim \mathcal{N}(0,\sigma_\eta)$. The maximum likelihood approach aims at finding the parameter vector that *maximizes the likelihood of the joint error distribution* defined as: $\mathcal{L}(\boldsymbol{\theta}) \stackrel{\text{def}}{=} p\left[e_{\mathcal{M}}(\boldsymbol{d}_1;\boldsymbol{\theta}),\ldots,e_{\mathcal{M}}(\boldsymbol{d}_N;\boldsymbol{\theta})\right]$. In the previous expression, p indicates the joint *probability distribution function* (pdf) of the errors. Intuitively, we are trying to find the parameter vector that maximizes the probability of observing the signed errors $e_{\mathcal{M}}(\boldsymbol{d}_i;\boldsymbol{\theta})$. Therefore we would like to calculate the estimate:

$$\hat{\boldsymbol{\theta}} = \operatorname*{argmax}_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta})$$

To simplify this maximization problem, we assume that the errors are *independent* (an assumption that should be made with some caution, especially in real life scenarios...) and we consider the log-likelihood $\mathcal{L}^*(\boldsymbol{\theta}) \stackrel{\text{def}}{=} \log \mathcal{L}(\boldsymbol{\theta})$. This trick allows us to simplify some calculations without affecting the final result, since the logarithm is a monotonically increasing function (and therefore the maximizer remains the same).

Under the previous assumptions we can write:

$$\mathcal{L}^*(\boldsymbol{\theta}) = \log \prod_{i=1}^{N} p[e_{\mathcal{M}}(\boldsymbol{d}_i;\boldsymbol{\theta})] = \sum_{i=1}^{N} \log p[e_{\mathcal{M}}(\boldsymbol{d}_i;\boldsymbol{\theta})] = \sum_{i=1}^{N} \left(\log \frac{1}{Z_G} - \frac{1}{2} \left(\frac{e_{\mathcal{M}}(\boldsymbol{d}_i;\boldsymbol{\theta})}{\sigma_\eta} \right)^2 \right)$$

where $Z_G = \sqrt{2\pi}\sigma_\eta$ is the normalization constant for the Gaussian distribution. Therefore the maximum likelihood estimate of the parameter vector is given by:

$$\hat{\boldsymbol{\theta}} = \operatorname*{argmax}_{\boldsymbol{\theta}} \sum_{i=1}^{N} \left(\log \frac{1}{Z_G} - \frac{1}{2} \left(\frac{e_{\mathcal{M}}(\boldsymbol{d}_i;\boldsymbol{\theta})}{\sigma_\eta} \right)^2 \right) = \operatorname*{argmin}_{\boldsymbol{\theta}} \sum_{i=1}^{N} \frac{1}{2} \left(\frac{e_{\mathcal{M}}(\boldsymbol{d}_i;\boldsymbol{\theta})}{\sigma_\eta} \right)^2 \quad (2.2)$$

The function to be minimized is convex[2] with respect to $\boldsymbol{\theta}$. and therefore the minimizer can be determined utilizing traditional iterative descent methods [Ber99, Lue03] (see Figure 2.2(a) for a numeric example and Section 4.2.1 for further details regarding the calculations). Note that (2.2) is nothing but the familiar *least square estimator*. This is a very well known results: for a more extensive treatment of this subject refer to [Men95].

We want to emphasize that the assumption of (independent) *Gaussian* errors implies that the probability of finding a point that supports the model with a residual larger than $3\sigma_\eta$ is *less than 0.3%*. We may be interested in finding the expression of the maximum likelihood estimate when the pdf of the error *is not* Gaussian. In particular we will focus our attention on the Cauchy–Lorentz distribution:

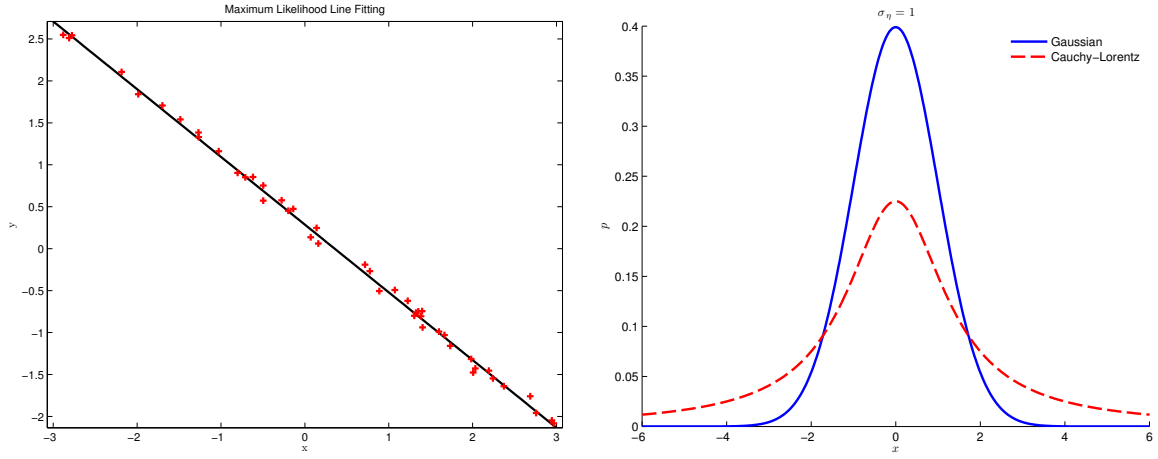
$$p\left[e_{\mathcal{M}}(\boldsymbol{d}_i;\boldsymbol{\theta})\right] = \frac{1}{Z_C}\frac{1}{1+\frac{1}{2}\left(\frac{e_{\mathcal{M}}(\boldsymbol{d}_i;\boldsymbol{\theta})}{\sigma_\eta}\right)^2}$$

where $Z_C = \sqrt{2}\pi\sigma_\eta$ is the normalization factor. The notation used in the previous formula *should not be misleading*: for this distribution the mean, the variance (or the higher moments) are *not* defined. The value for the so called *scale parameter* has been chosen to be consistent with the expression obtained in the Gaussian case. It is important to observe that the Cauchy–Lorentz distribution is characterized by *heavier tails* than the Gaussian distribution. Intuitively, this implies that the probability of finding a "large" error is higher if the distribution is Cauchy–Lorentz than if the distribution is Gaussian (see Figure 2.2(b)). If we derive the maximum likelihood estimate for this distribution we obtain:

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \sum_{i=1}^{N}\left(\log\frac{1}{Z_C} - \log\left(1+\frac{1}{2}\left(\frac{e_{\mathcal{M}}(\boldsymbol{d}_i;\boldsymbol{\theta})}{\sigma_\eta}\right)^2\right)\right) = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \sum_{i=1}^{N}\log\left(1+\frac{1}{2}\left(\frac{e_{\mathcal{M}}(\boldsymbol{d}_i;\boldsymbol{\theta})}{\sigma_\eta}\right)^2\right) \quad (2.3)$$

Also in this case the function to be minimized is convex with respect to $\boldsymbol{\theta}$ and therefore the minimizer can be computed utilizing traditional iterative descent methods.

[2] A function $f : \mathbb{R} \to \mathbb{R}$ is convex if its graph between x_1 and x_2 lies below any segment that connects $f(x_1)$ to $f(x_2)$. Formally speaking, if $\forall\lambda \in [0,1]$ we have that $f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2)$. This notion generalizes straightforwardly for vector functions (whose graph essentially looks like a cereal bowl).



(a) (b)

Figure 2.2: (a) Maximum likelihood line estimate assuming a Gaussian error distribution ($\sigma_\eta = 0.05$). (b) Comparison between the Gaussian and the Cauchy–Lorentz distribution for $\sigma_\eta = 1$. Note that the Cauchy–Lorentz distribution has "heavier tails" than the Gaussian distribution.

If we compare expression (2.2) with (2.3) we can readily see that they are *structurally very similar*. Indeed if we introduce the function ρ we can write:

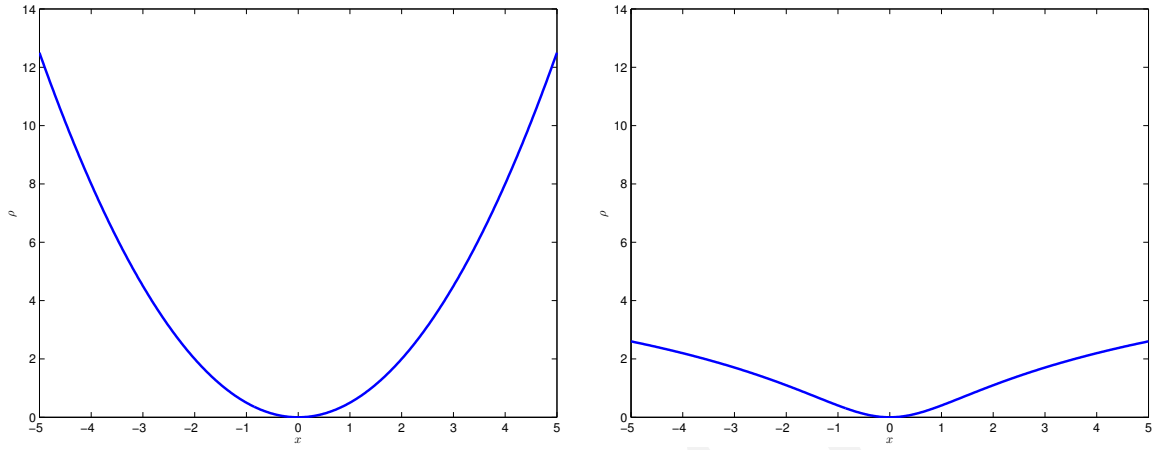
$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \sum_{i=1}^{N} \rho(e_{\mathcal{M}}(\boldsymbol{d}_i; \boldsymbol{\theta})) \tag{2.4}$$

Figure 2.3 compares the ρ function associated to the classical least square estimator (2.2) and to the maximum likelihood estimator based on the Cauchy–Lorentz distribution (2.3). From the plots it is clear that *large errors contribute less* to the sum of the residuals in the Cauchy–Lorentz case.

The family of estimators that can be written as in (2.4) are called *M-estimators* and often times they are considered robust alternatives to classical least square estimators. We will dig more into the issue of robustness in the next section.

2.2 Outliers, Bias and Breakdown Point

This section will introduce three important concepts that will outline the limitations of traditional least square estimation approaches and that will motivate the discussion of RANSAC in Chapter 3.



(a) (b)

Figure 2.3: The comparison of the ρ function for the (a) least square case and for the (b) Cauchy–Lorentz case. In both cases $\sigma_\eta = 1$.

2.2.1 Outliers

To the best of my knowledge, there does not exist a formal definition of *outlier*. My attempt to provide a definition which is general enough but still reasonably formal is the following:

> A datum is considered to be an outlier if it will not fit the "true" model instantiated by the "true"set of parameters within some error threshold that defines the maximum deviation attributable to the effects of noise.

Here we assume that there exists a model and a set of parameters (the "true" ones) that can exactly generate the observed measurements if they were observed in absence of noise. We also assume that we have a knowledge of the so called *noise scale*, i.e. we know what is the maximum perturbation of an observed valid measurement (i.e. produced by the true model, instantiated with the true parameters and measured in presence of noise).

2.2.2 Bias

To define the bias we follow quite closely the definition provided by Rousseeuw and Leroy [RL87]. Let $D_I \subset D$ be *a* set of inliers and let $D_{I/O}(m)$ be the previous set after m inliers have been *replaced* by outliers. As usual $\mathcal{M}$ indicates the model that we are considering.

Definition 1. *The bias associated to the model $\mathcal{M}$ and to sets D_I and $D_{I/O}(m)$ is defined as:*

$$\text{bias}_{\mathcal{M}}(m; D_I) \stackrel{\text{def}}{=} \sup_{D_{I/O}(m)} \text{dist}_{\mathcal{M}}\left(\boldsymbol{\theta}(D_I), \boldsymbol{\theta}(D_{I/O}(m))\right) \tag{2.5}$$

This quantity measures the maximum perturbation that can be caused to a parameter vector first estimated *only* using inliers, and then when m of such inliers are replaced by outliers. The function $\text{dist}_{\mathcal{M}}$ measures the distance between the parameter vectors. An appropriate choice for such function is crucial and often times the simple Euclidean norm is not suitable.[3]

2.2.3 Breakdown Point

Intuitively, the breakdown point of an estimator represents the minimum fraction of outliers that are sufficient to produce an arbitrary large bias:

Definition 2. *The breakdown point associated to the model $\mathcal{M}$ and to the set D_I is defined as:*

$$\text{BP}_{\mathcal{M}}(D_I) \stackrel{\text{def}}{=} \min\left\{\frac{m}{|D_I|} : \text{bias}_{\mathcal{M}}(m; D_I) = \infty\right\} \tag{2.6}$$

Quite often the bias does not depend on a particular set of points, but it is instead just related to the properties of the considered model and of the associated estimator.

2.3 The Breakdown Point for a 2D Line Least Squares Estimator

In the 2D line example considered in this chapter it is possible to show that the presence of one single outlier can arbitrarily bias the estimate of the parameters of the line. Intuitively we can arbitrarily modify the estimate of the line introducing an outlier and moving it around on the 2D plane (see Figure 2.4). Therefore the least square estimators breakdown point is 0%. Even if it is possible to build M-estimators such that $\text{BP}_{\mathcal{M}}(D_I) > 0$, this does not necessarily mean that they can be considered

[3] This constitutes the main difference between our definition of bias and Rousseeuw and Leroy's one [RL87]. If we consider the parameter vector that defines a 2D line, its length is immaterial and the only information that matters is its direction. Thus, taking the Euclidean distance (or any other p–norm based distance between two of such vectors) does not provide a meaningful measure of "how close" the two models are.

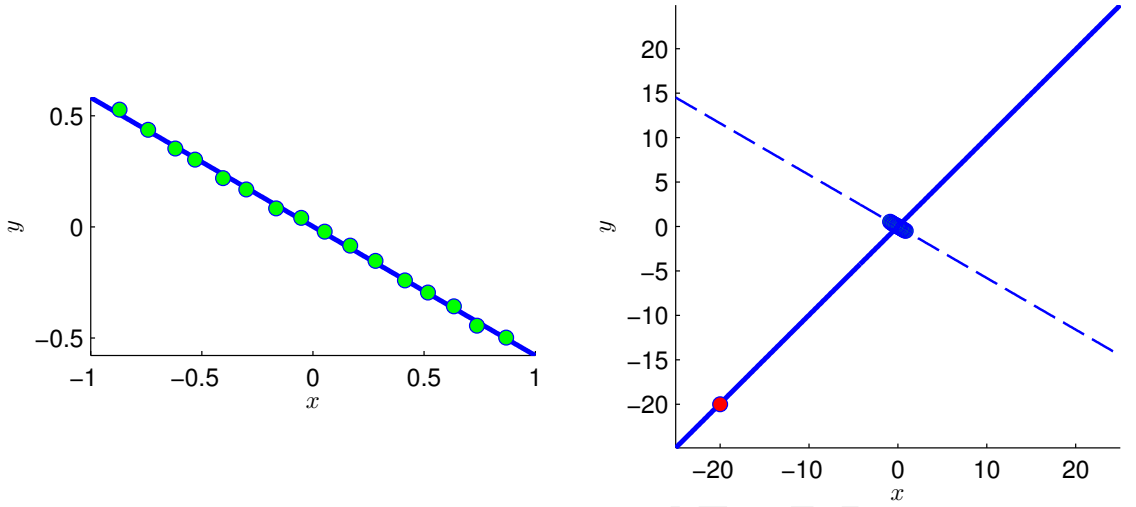


Figure 2.4: These figures describe the effect of one outlier (red dot) on the least square estimation of the parameters of a 2D line.

robust for practical purposes. Certainly we can obtain some improvements if we have more information regarding the error statistics, but in practical scenarios we seek for robustness exactly when this type of information is lacking: robustness in presence of non–modeled distributions! Moreover, even if the bias stays bounded, the outliers will still affect the estimate, which will differ from the "true value" of the parameter vector.

Despite the very intuitive and qualitative character of the previous considerations and the simplicity of the 2D line estimation example, we believe that it is now clear why the estimation of the parameters of a model in presence of outliers constitutes a very tricky and challenging problem. RANSAC is an effective algorithm to cope with these types of problems.

3 RANdom Sample And Consensus

This chapter is structured as follows. First we will briefly introduce the RANSAC algorithm in Section 3.1. In Section 3.2 we will present the notation and provide some definitions that will facilitate the discussion of the technical details of the algorithm (Section 3.3 and 3.4). Finally we will present related works in the literature in Section 3.5.

3.1 Introduction

The RANSAC algorithm (*RANdom Sample And Consensus*) was first introduced by Fischler and Bolles [FB81] in 1981, as a method to estimate the parameters of a certain model[1] starting from a set of data contaminated by large amounts of *outliers*. In this tutorial, following the definition given in Section 2.2, a datum is considered to be an outlier if it will not fit the "true" model instantiated by the "true"' set of parameters within some error threshold that defines the maximum deviation attributable to the effect of noise. The percentage of outliers which can be handled by RANSAC can be larger than 50% of the entire data set. Such a percentage, know also as the *breakdown point*, is commonly assumed to be the practical limit for many other commonly used techniques for parameter estimation (such as all the least squares flavors or robust techniques like the least median of squares [Hub81, RL87, Men95, Zha97, Ste99]). We want to mention here a robust estimator proposed by Stewart called MINPRAN [Ste95], capable of estimating the parameters

[1] Fischler and Bolles used RANSAC to solve the Location Determination Problem (LDP), where the goal is to determine the points in the space that project onto an image into a set of landmarks with known locations.

of a model using datasets containing more than 50% of outliers.

Despite many modifications, the RANSAC algorithm is essentially composed of two steps that are repeated in an iterative fashion (*hypothesize–and–test framework*):

- **Hypothesize**. First *minimal sample sets* (MSSs) are randomly selected from the input dataset and the model parameters are computed using *only* the elements of the MSS. The cardinality of the MSS is the smallest sufficient to determine the model parameters[2] (as opposed to other approaches, such as *least squares*, where the parameters are estimated using *all* the data available, possibly with appropriate weights).

- **Test**. In the second step RANSAC checks which elements of the entire dataset are consistent with the model instantiated with the parameters estimated in the first step. The set of such elements is called *consensus set* (CS).

RANSAC terminates when the probability of finding a better ranked CS drops below a certain threshold. In the original formulation the ranking of the CS was its cardinality (i.e. CSs that contain more elements are ranked better than CSs that contain fewer elements).

3.2 Preliminaries

To facilitate the discussion that follows, it is convenient to introduce a suitable formalism to describe the steps for the estimation of the model parameters and for the construction of the CS. As usual we will denote vectors with boldface letters and the superscript $^{(h)}$ will indicate the h^{th} iteration. The symbol $\hat{x}$ indicates the estimated value of the quantity x. The input dataset, which is composed of N elements, is indicated by $D = \{\boldsymbol{d}_1, \ldots, \boldsymbol{d}_N\}$ and we will indicate a MSS with the letter s. Let $\boldsymbol{\theta}(\{\boldsymbol{d}_1, \ldots, \boldsymbol{d}_h\})$ be the parameter vector estimated using the set of data $\{\boldsymbol{d}_1, \ldots, \boldsymbol{d}_h\}$, where $h \geq k$ and k is the cardinality of the MSS. The *model space* $\mathcal{M}$ is defined as:

$$\mathcal{M}(\boldsymbol{\theta}) \stackrel{\text{def}}{=} \left\{\boldsymbol{d} \in \mathbb{R}^d : f_{\mathcal{M}}(\boldsymbol{d}; \boldsymbol{\theta}) = 0\right\}$$

where $\boldsymbol{\theta}$ is a parameter vector and $f_{\mathcal{M}}$ is a smooth function whose zero level set contains all the points that fit the model $\mathcal{M}$ instantiated with the parameter vector

[2]Suppose we want to estimate a line: in this case the cardinality of the MSS is 2, since at least two distinct points are needed to uniquely define a line.

$\boldsymbol{\theta}$. We define the error associated with the datum $\boldsymbol{d}$ with respect to the model space as the distance from $\boldsymbol{d}$ to $\mathcal{M}(\boldsymbol{\theta})$:

$$e_{\mathcal{M}}(\boldsymbol{d}, \boldsymbol{\theta}) \stackrel{\text{def}}{=} \min_{\boldsymbol{d}' \in \mathcal{M}(\boldsymbol{\theta})} \text{dist}(\boldsymbol{d}, \boldsymbol{d}')$$

where $\text{dist}(\cdot, \cdot)$ is an appropriate distance function. Using this error metric, we define the CS as:

$$S(\boldsymbol{\theta}) \stackrel{\text{def}}{=} \{\boldsymbol{d} \in D : e_{\mathcal{M}}(\boldsymbol{d}; \boldsymbol{\theta}) \leq \delta\} \tag{3.1}$$

where δ is a threshold that can either be inferred from the nature of the problem or, under certain hypothesis, estimated automatically [WS04] (see Figure 3.1 for a pictorial representation of the previous definitions). In the former case, if we want

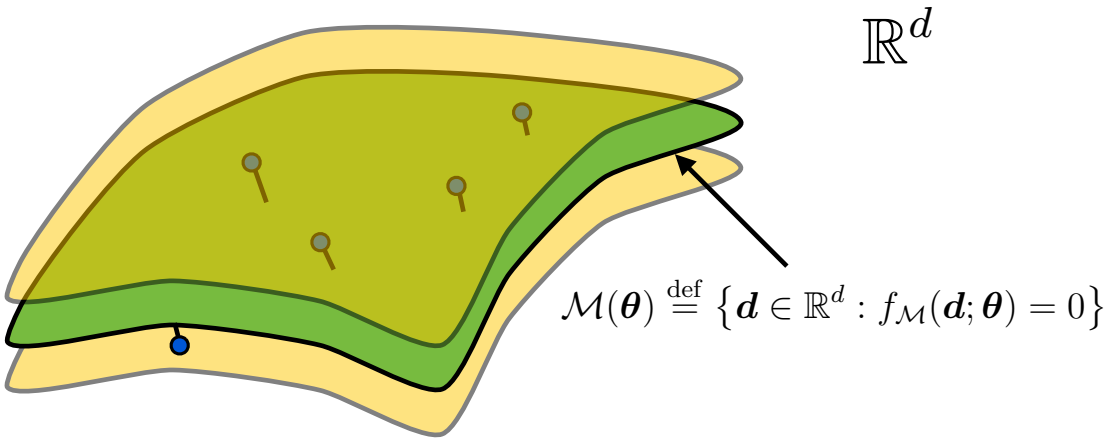


Figure 3.1: This figure pictorially displays the model space $\mathcal{M}$ as a green surface (the locus for which $f_{\mathcal{M}}(\boldsymbol{d}; \boldsymbol{\theta}) = 0$). The yellow surfaces represent the boundaries for a datum to be considered an inlier (imagine that the distance function is the Euclidean distance, hence the smallest distance between any two points on the yellow surface and green surface is δ). Note that the structure of the green surface is both defined by the model $\mathcal{M}$ and by the parameter vector $\boldsymbol{\theta}$. The inliers, represented as blue dots, lie in between the two yellow "crusts".

to relate the value of δ to the statistics of the noise that affects the data and the distance function is the Euclidean norm, we can write:

$$e_{\mathcal{M}}(\boldsymbol{d}, \boldsymbol{\theta}) = \min_{\boldsymbol{d}' \in \mathcal{M}(\boldsymbol{\theta})} \sqrt{\sum_{i=1}^{n} (d_i - d'_i)^2} = \sqrt{\sum_{i=1}^{n} (d_i - d^*_i)^2}$$

where $\boldsymbol{d}^*$ is the orthogonal projection of $\boldsymbol{d}$ onto the model space $\mathcal{M}(\boldsymbol{\theta})$. Now suppose that the datum $\boldsymbol{d}$ is affected by Gaussian noise $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma_\eta I)$ so that $\boldsymbol{\eta} = \boldsymbol{d} - \boldsymbol{d}^*$. Our goal is to calculate the value of δ that bounds, with a given probability P_{inlier}, the error generated by a true inlier contaminated with Gaussian noise. More formally we want to find the value δ such that:

$$P[e_{\mathcal{M}}(\boldsymbol{d}, \boldsymbol{\theta}) \leq \delta] = P_{inlier} \tag{3.2}$$

Following [HZ03], p. 118, we can write the following chain of equations:

$$P[e_{\mathcal{M}}(\boldsymbol{d}, \boldsymbol{\theta}) \leq \delta] = P\left[\sum_{i=1}^{n} \eta_i^2 \leq \delta^2\right] = P\left[\sum_{i=1}^{n} \left(\frac{\eta_i}{\sigma_\eta}\right)^2 \leq \frac{\delta^2}{\sigma_\eta^2}\right]$$

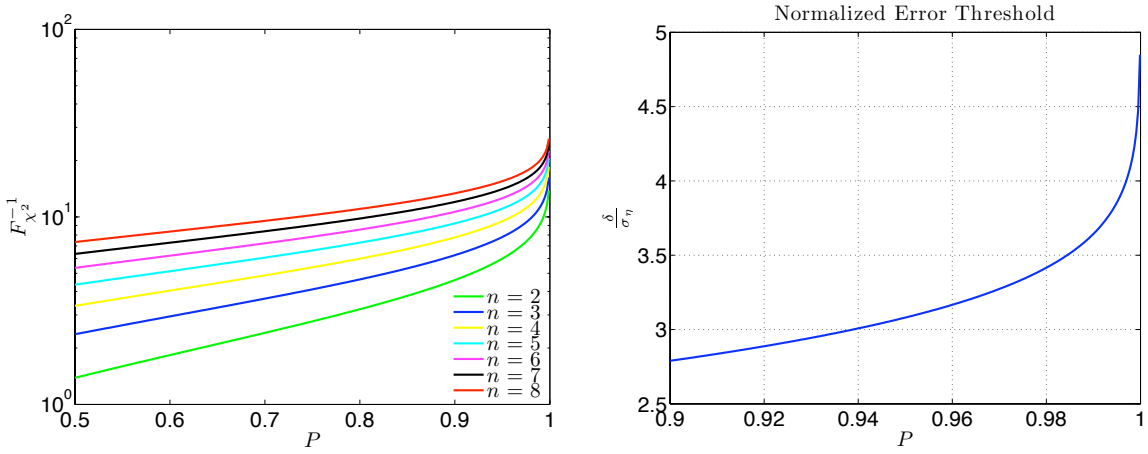
and since $\eta_i/\sigma_\eta \sim \mathcal{N}(0,1)$, the random variable $\sum_{i=1}^{n} \left(\frac{\eta_i}{\sigma_\eta}\right)^2$ has a χ_n^2 distribution. Hence:

$$\delta = \sigma_\eta \sqrt{F_{\chi_n^2}^{-1}(P_{inlier})} \tag{3.3}$$

where $F_{\chi_n^2}^{-1}$ is the inverse cumulative distribution function associated with a χ_n^2 random variable. Figure 3.2(a) displays the function $F_{\chi_n^2}^{-1}$ for different values of n. Note that when P_{inlier} tends to one (i.e. we want to pick an error threshold such that all the inliers will be considered) the value of $F_{\chi_n^2}^{-1}$ diverges to infinity. Values of P_{inlier} close to one will return a large threshold with the risk of including some outliers as well. On the other hand, too small values of P_{inlier} will generate a value for δ which is too tight, and possibly some inliers will be discarded.

3.3 RANSAC Overview

A pictorial representation of the RANSAC fundamental iteration together with the notation just introduced is shown in Figure 3.3. As mentioned before, the RANSAC algorithm is composed of two steps that are repeated in an iterative fashion (hypothesize-and-test framework). First a MSS $s^{(h)}$ is selected from the input dataset and the model parameters $\boldsymbol{\theta}^{(h)}$ are computed using *only* the elements of the selected MSS. Then, in the second step, RANSAC checks which elements in the dataset D are consistent with the model instantiated with the estimated parameters and, if it is the case, it updates the current best CS S^* (which, in the original Fischler and Bolles



(a) (b)

Figure 3.2: Figure (a) shows the function $F^{-1}_{\chi^2_n}$ for different values of n. Note that the vertical axis values are logarithmically spaced). Figure (b) displays the normalized error threshold for the symmetric transfer error of planar correspondences ($n = 2 + 2 = 4$).

formulation, is the CS with the largest cardinality). The algorithm terminates when the probability of finding a better CS drops below a certain threshold. In the next paragraphs we will discuss how to estimate the number of iterations that RANSAC is supposed to perform and other approaches to rank the CSs.

3.3.1 How many iterations?

Let q be the probability of sampling from the dataset D a MSS s that produces an accurate estimate of the model parameters. Consequently, the probability of picking a MSS containing at least one outlier (i.e. a MSS that produces a biased estimate of the true model parameter vector) is $1 - q$. If we construct h different MSSs, then the probability that *all of them are contaminated by outliers* is $(1 - q)^h$ (this quantity tends to zero for h going to infinity: sooner or later we will pick something good!). We would like to pick h (i.e. the number of iterations) large enough so that the probability $(1 - q)^h$ is smaller or equal than a certain probability threshold ε (often called *alarm rate*), i.e. $(1 - q)^h \leq \varepsilon$. The previous relation can be inverted so that we can write:

$$h \leq \left\lceil \frac{\log \varepsilon}{\log (1 - q)} \right\rceil$$

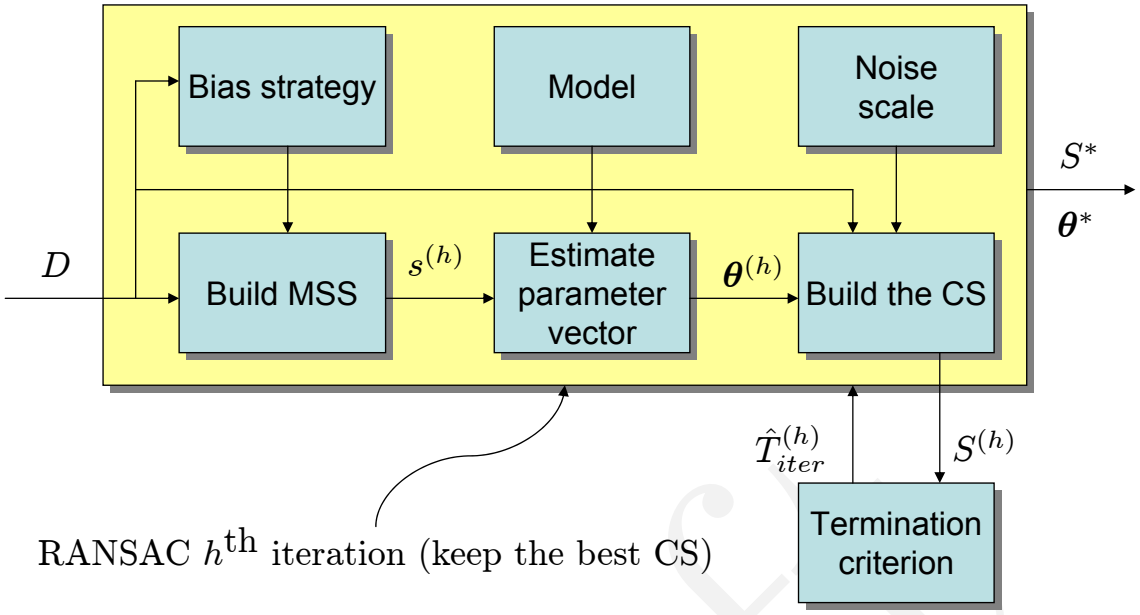


Figure 3.3: Pictorial representation of the fundamental RANSAC iteration.

where $\lceil x \rceil$ denotes the smallest integer larger than x. Therefore we can set the threshold for the iterations to:

$$\hat{T}_{iter} = \left\lceil \frac{\log \varepsilon}{\log (1-q)} \right\rceil \tag{3.4}$$

3.3.2 Constructing the MSSs and Calculating q

If we imagine that the inliers inside the dataset D are noise free, then any MSS entirely composed of inliers will generate the "true" value of the parameter vector.[3] If all the elements in the dataset have *the same probability of being selected*, then the probability of obtaining a MSS composed only of inliers is:

$$q = \frac{\binom{N_I}{k}}{\binom{N}{k}} = \frac{N_I!(N-k)!}{N!(N_I-k)!} = \prod_{i=0}^{k-1} \frac{N_I - i}{N - i} \tag{3.5}$$

[3]This statement is not completely true. In real life applications it is not possible to disregard numerical approximations or to pick the noise threshold δ to be an arbitrarily small positive number. Therefore there exist configurations of inliers such that the estimate of the parameter vector can still be biased (see [TM05] for further insights regarding this issue).

where N_I is the total number of inliers. Note that if $N, N_I \gg k$, then q is approximately equivalent to the probability of picking for k times an inlier from the dataset (with re-insertion). In fact:

$$q = \prod_{i=0}^{k-1} \frac{N_I - i}{N - i} \approx \left(\frac{N_I}{N}\right)^k \tag{3.6}$$

Unfortunately, to compute q we should know N_I, which is generally not known a priori. However it is easy to verify that for any $\hat{N}_I \leq N_I$ we have $q(\hat{N}_I) \leq q(N_I)$ and consequently $(1 - q(N_I))^h \geq \left(1 - q(\hat{N}_I)\right)^h$ (where we made explicit the dependency of q on the number of inliers). Therefore we can estimate the maximum number of iterations using the cardinality of the *largest set of inliers found so far* (call this $\hat{N}_I$), which can be regarded as a conservative estimate of N_I. Hence, the iteration threshold can be fixed to:

$$\hat{T}_{iter} = \left\lceil \frac{\log \varepsilon}{\log\left(1 - q(\hat{N}_I)\right)} \right\rceil \tag{3.7}$$

Note however that some researchers (for example Tordoff et al. [TM05]) consider this threshold on the number of the iterations to be *over-optimistic*, since in presence of noisy data it is not enough to generate a MSS composed only of inliers to obtain a reliable estimate of the parameters of the model. This observation motivates some of the considerations that will be outlined in the next section.

Remark 1. *Sometimes there exists some a priori information regarding the probability that a datum is an inlier or an outlier. This information can be used to guide the sampling procedure in a more principled manner (see also [VL01, KK04, CM05, TM05]). In the biased sampling case the probability of picking the element $\boldsymbol{d}_j$ at the l^{th} draw will be denoted as:*

$$P(\boldsymbol{d}_j | D \setminus \{\boldsymbol{d}_{i^{(1)}}, \ldots, \boldsymbol{d}_{i^{(l-1)}}\}) \tag{3.8}$$

where $i^{(l)}$ denotes the index of the element in the dataset D obtained at the l^{th} draw. The probability (3.8) is determined by the bias $\omega(\boldsymbol{d}) \in \mathbb{R}_+$*, which incorporates the a priori information regarding the likelihood of an element to be an inlier. In an ideal case, the weight of any inlier is* larger *than the weight of any outlier, but in real life*

scenarios this hypothesis does not necessarily hold true (otherwise the detection of the inliers versus the outlier would be greatly *simplified: we would just have to identify the correct threshold for the weights). Clearly, the computation of the probability of obtaining a MSS composed only of inliers is complicated by the fact that the order in which the elements are sampled* matters*: the element picked at the l^{th} draw modifies the probability of picking a certain element at the $(l+1)^{th}$ draw. Thus the probability of sampling k inliers in the order specified by the indices $i_1, \ldots, i_k$ is:*

$$\begin{aligned} P(\boldsymbol{d}_{i_1}, \ldots, \boldsymbol{d}_{i_k}) &= \\ (1^{th} \text{ draw}) \quad & P(\boldsymbol{d}_{i_1} \in D_I)\cdot \\ (2^{nd} \text{ draw}) \quad & P(\boldsymbol{d}_{i_2} \in D_I \setminus \{\boldsymbol{d}_{i_1}\})\cdot \\ & \vdots \\ (k^{th} \text{ draw}) \quad & P(\boldsymbol{d}_{i_k} \in D_I \setminus \{\boldsymbol{d}_{i_1}, \ldots, \boldsymbol{d}_{i_{k-1}}\}) \end{aligned} \tag{3.9}$$

The previous expression confirms what was anticipated earlier, i.e. that the computation of q is complicated by the fact that the order of selection of the elements matters. To compute q we need to enumerate all the possible permutations of k inliers (recall that k is the cardinality of a MSS), which turn out to be:

$$\text{Number of permutations of the inliers} = \frac{N_I!}{(N_I - k)!} = \prod_{i=0}^{k-1}(N_I - i)$$

If $\mathcal{I}$ is the set of all the permutations of k indices of inliers then we can write:

$$q = \sum_{\{i_1,\ldots,i_k\}\in\mathcal{I}} P(\boldsymbol{d}_{i_1}, \ldots, \boldsymbol{d}_{i_k}) \tag{3.10}$$

where the elements of the sums can be expanded using (3.10). This expression can be used to design a function that updates the probability every time the number of inliers is updated. However this is not a trivial task, since the number of the terms of the summation can be extremely large[4].

[4]If the number of inliers is $N_I = 100$ and the cardinality of the MMS is $k = 4$ the number of terms in the summation (3.10) is almost four millions.

3.3.3 Ranking the Consensus Set

In the original formulation of RANSAC, the ranking r of a consensus set was nothing but its cardinality:

$$r(CS) \stackrel{\text{def}}{=} |CS|$$

In other words CSs that are larger are ranked higher. Thus RANSAC can be seen as an *optimization* algorithm [TZ00] that *minimizes* the cost function:

$$C_{\mathcal{M}}(D;\boldsymbol{\theta}) = \sum_{i=1}^{N} \rho(\boldsymbol{d}_i, \mathcal{M}(\boldsymbol{\theta})) \tag{3.11}$$

where:

$$\rho(\boldsymbol{d}, \mathcal{M}(\boldsymbol{\theta})) = \begin{cases} 0 & |e_{\mathcal{M}}(\boldsymbol{d}, \boldsymbol{\theta})| \leq \delta \\ 1 & \text{otherwise} \end{cases} \tag{3.12}$$

This observation leads immediately to an approach based on *M-estimators* [Hub81, RL87, Zha97], the idea being to substitute the function ρ with a more sensible one. The first function that we consider is the following:

$$\rho(\boldsymbol{d}, \mathcal{M}(\boldsymbol{\theta})) = \begin{cases} e_{\mathcal{M}}(\boldsymbol{d}, \boldsymbol{\theta}) & e_{\mathcal{M}}(\boldsymbol{d}, \boldsymbol{\theta}) \leq \delta \\ \delta & \text{otherwise} \end{cases}$$

Using this re-descending M-estimator, the inliers are scored according to their *fitness to the model*, while the outilers are given a constant weight. Torr et al. refer to this modification of the original RANSAC algorithm with the name MSAC, i.e. M-estimator SAmple and Consensus. We agree with their claim:

> The implementation of this new method yields a modest to hefty benefit to all robust estimations with absolutely no additional computational burden. Once this is understood there is no reason to use RANSAC in preference to this method.

A further improvement can be obtained modifying RANSAC in order to maximize the likelihood of the solution. This is the approach implemented by MLESAC [TZ00], a variation of RANSAC that evaluates the *likelihood* of the hypothesis by representing the error distribution as a *mixture model*. More precisely, the probability distribution of the error for entire dataset (comprising both the iniers and outliers) can be modeled

as the mixture of two distributions (one taking into account the inliers, tho other the outliers) so that the likelihood can be expressed as:

$$p[e(D,\mathcal{M}(\boldsymbol{\theta}))|\boldsymbol{\theta}] = \prod_{i=1}^{N} \left(\gamma p\left[e(\boldsymbol{d}_i,\mathcal{M}(\boldsymbol{\theta})|\text{the } i^{th} \text{ element is an inlier}\right] + (1-\gamma)p\left[e(\boldsymbol{d}_i,\mathcal{M}(\boldsymbol{\theta})|\text{the } i^{th} \text{ element is an outlier}\right]\right)$$

A common practice is to maximize the log-likelihood, which is given by:

$$\mathcal{L}^*[e(D,\mathcal{M}(\boldsymbol{\theta}))|\boldsymbol{\theta}] = \sum_{i=1}^{N} \log\left(\gamma p\left[e(\boldsymbol{d}_i,\mathcal{M}(\boldsymbol{\theta}))|\text{the } i^{th} \text{ element is an inlier}\right] + (1-\gamma)p\left[e(\boldsymbol{d}_i,\mathcal{M}(\boldsymbol{\theta}))|\text{the } i^{th} \text{ element is an outlier}\right]\right)$$

Often times the error distributions for the inliers is modeled with a Gaussian distribution:

$$p\left[e(\boldsymbol{d}_i,\mathcal{M}(\boldsymbol{\theta}))|\text{the } i^{th} \text{ element is an inlier}\right] = \frac{1}{Z}\exp\left(-\frac{e(\boldsymbol{d}_i,\mathcal{M}(\boldsymbol{\theta})^2}{2\sigma_\eta^2}\right)$$

where Z is the appropriate normalization constant and σ_η indicates the noise standard deviation (see also equation (3.3)). The error statistics for the outliers is described by a uniform distribution:

$$p\left[e(\boldsymbol{d}_i,\mathcal{M}(\boldsymbol{\theta}))|\text{the } i^{th} \text{ element is an outlier}\right] = \begin{cases} \frac{1}{2e_{max}} & |e(\boldsymbol{d}_i,\mathcal{M}(\boldsymbol{\theta}))| \leq e_{max} \\ 0 & \text{otherwise} \end{cases}$$

where e_{max} represents the largest error which can be possibly induced by the presence of outliers (an estimate of such quantity can be obtained from the context the data are drawn from). Note that in this case we need to estimate two quantities: the parameter $\boldsymbol{\theta}$ that maximizes the likelihood and the mixture coefficient γ. This is traditionally done using the *expectation maximization approach*.[5]

[5] A good practical reference with Matlab™ & Octave code to perform this task can be found in [Nab03]

3.4 Computational Complexity

In this section we will briefly discuss the computational complexity associated to RANSAC.

3.4.1 Hypothesize Step

At each iteration we need to compute the parameters of the model starting from the MSS (whose cardinality is k). Let's define the cost of this operation to be $C_{estimate}(k)$.

3.4.2 Test Step

Once the parameters have been estimated we need to evaluate how many data fit the model. If the cost associated to compute the fitting of one single element is $C_{fitting}$, then the overall cost is $NC_{fitting}$. Note that the approach by Chum et al. [CM02] aims at reducing the complexity of this stage.

3.4.3 Overall Complexity

Putting together the cost of the hypothesize and test steps, the overall complexity of the algorithm in the worst case scenario is:

$$\text{Complexity} = O\left(T_{iter}(C_{estimate}(k) + NC_{fitting})\right) \tag{3.13}$$

3.5 Other RANSAC Flavors

Since 1981 RANSAC has become a fundamental tool in the computer vision and image processing community. In 2006, for the 25^{th} anniversary of the algorithm, a workshop was organized at the International Conference on Computer Vision and Pattern Recognition (CVPR) to summarize the most recent contributions and variations to the original algorithm, mostly meant to improve the speed of the algorithm, the robustness and accuracy of the estimated solution and to decrease the dependency from user defined constants. We regard with amusement the effort of the researchers in finding colorful variations to the name of the original algorithm. Some of these approaches (with their relative names) will be described in the following paragraphs.

As pointed out by Torr et al. , RANSAC can be sensitive to the choice of the correct noise threshold that defines which data points fit a model instantiated with a certain set of parameters. If such threshold is too large, then all the hypotheses tend to be ranked equally (good). On the other hand, when the noise threshold is too small, the estimated parameters tend to be unstable (i.e. by simply adding or removing a datum to the set of inliers, the estimate of the parameters may fluctuate). To partially compensate for this undesirable effect, Torr et al. proposed two modification of RANSAC called MSAC (M-estimator SAmple and Consensus) and MLESAC (Maximum Likelihood Estimation SAmple and Consensus) [TZ00]. The main idea is to evaluate the quality of the consensus set (i.e. the data that fit a model and a certain set of parameters) calculating its likelihood (whereas in the original formulation by Fischler and Bolles the rank was the cardinality of such set). More details have been provided in Section 3.3.3. An extension to MLESAC which keeps into account the prior probabilities associated to the input dataset is proposed by Tordoff et al. in [TM05]. The resulting algorithm is dubbed Guided-MLESAC. Along similar lines, Chum proposed to guide the sampling procedure if some a priori information regarding the input data is known, i.e. whether a datum is likely to be an inlier or an outlier. The proposed approach is called PROSAC, PROgressive SAmple Consensus [CM05].

Chum et al. also proposed a randomized version of RANSAC called R-RANSAC [CM02] to reduce the computational burden to identify a good CS. The basic idea is to initially evaluate the goodness of the currently instantiated model using only a reduced set of points instead of the entire dataset. A sound strategy will tell with high confidence when it is the case to evaluate the fitting of the entire dataset or when the model can be readily discarded. It is reasonable to think that the impact of this approach is more relevant in cases where the percentage of inliers is large. The type of strategy proposed by Chum et al. is called *preemption scheme*. In [Nis03], Nistér proposed a paradigm called Preemptive RANSAC that allows real time robust estimation of the structure of a scene and of the motion of the camera. The core idea of the approach consists in generating a fixed number of hypothesis so that the comparison happens with respect to the quality of the generated hypothesis rather than against some absolute quality metric.

Other researchers tried to cope with difficult situations where the noise scale is not know and/or multiple model instances are present. The first problem has been tackled in the the work by Wang and Suter [WS04]. Some solutions for the second

problem where instead proposed in [ZKM05] and more recently in [TF08]. More specifically, Toldo et al. represent each datum with the characteristic function of the set of random models that fit the point. Then multiple models are revealed as clusters which group the points supporting the same model. The clustering algorithm, called J-linkage, does not require prior specification of the number of models, nor it necessitate manual parameters tuning.

RANSAC has also been tailored for recursive state estimation applications, where the input measurements are corrupted by outliers and Kalman filter approaches, which rely on a Gaussian distribution of the measurement error, are doomed to fail. Such an approach, dubbed KALMANSAC, is presented in [VJFS05].

4 RANSAC at Work

In this chapter we will describe some simple applications of RANSAC. To facilitate the discussion we will utilize the RANSAC Toolbox for Matlab™ & Octave .

4.1 The RANSAC Toolbox for Matlab™ & Octave

In this section we briefly introduce the RANSAC Toolbox for Matlab™ & Octave . This toolbox is highly customizable and it is designed to be a flexible research/didactic resource.

To install the package first download the `.zip` archive from the Mathworks file exchange website. After unzipping the archive run the script `SetPathLocal.m` to set the local paths and some support global variables. To check if updates are available run the script `RANSAC_update.m`. The sub-directory `Examples` contains some example scripts to estimate the parameters of lines and planes as well as rotation, scaling and translations (RST) and homographic transformations. Each example will be described in some more detail in the next sections.

4.1.1 `RANSAC.m`

This is the driver function that implements the RANSAC algorithm. Here we will describe in detail the options that are accessible to the user and that can be listed issuing the usual command `help RANSAC`.

Input Parameters

- `X`: this matrix collects the input dataset. Its dimensions are $d \times N$ and the i^{th} column contains the datum $\boldsymbol{d}_i$.

- `options`: this structure collects the options for the algorithm. If an option is not specified the function will either use a default value or issue an error message (if a default value for the desired option does not exist).

  - `sigma`: scalar value of the noise standard deviation under the assumption that the components of the input are corrupted by Gaussian noise with covariance matrix $\sigma_\eta I$. See the discussion in Section 3.2 and Equation (3.3).

  - `P_inlier`: the probability that a point whose fitting error is less or equal than δ is actually an inlier (see Equation (3.2)). Default value $P_{inlier} = 0.99$.

  - `T_noise_squared`: when this value is provided, it will force the error threshold to be equal to δ^2. If this value is provided, the calculation of the threshold using the value of σ_η (see Equation (3.3)) is *overridden*.

  - `epsilon`: False alarm rate, i.e. the probability that the algorithm throughout all the iterations will *never* sample a MSS containing *only* inliers (see Section 3.3.1). Default value $\varepsilon = 0.001$.

  - `Ps`: Probability mass distribution to bias the sampling: `Ps(i)` $= P(\boldsymbol{d}_i$ is sampled). Note that $P_s \in \mathbb{R}_+^N$ must be normalized so that $\sum_{i=1}^N P(\boldsymbol{d}_i \text{ is sampled}) = 1$. The default value is a uniform distribution, i.e. $P(\boldsymbol{d}_i \text{ is sampled}) = \frac{1}{N}$.

  - `ind_tabu`: logical array that identifies the elements that *should not* be selected to construct the MSSs during the sampling (default is empty).

  - `validateMSS_fun`: function that validates the MSS sampled from the dataset (see 4.2.4). A typical template for this function is found in Appendix D.1.1.

  - `est_fun`: function that returns the estimate of the parameter vector starting from a set of data. A typical template for this function is found in Appendix D.1.2.

- `validateTheta_fun`: function that validates the parameter vector estimated from the MSS. A typical template for this function is found in Appendix D.1.3.
- `man_fun`: function that returns the fitting error of the data. A typical template for this function is found in Appendix D.1.4.
- `mode`: A string that specifies the algorithm flavour, specifically the raking of the CS, as described in see Section 3.3.3 (default = `MSAC`):
  - `'RANSAC'`: original Fischler and Bolles formulation [FB81], the ranking of the CS is its cardinality (see (3.11))
  - `'MSAC'`: Torr and Zisserman MSAC formulation [TZ00], the ranking of the CS is based on a M-estimator (see (3.12))
  - `'MLESAC'`: Torr and Zisserman MLESAC formulation [TZ00], the ranking of the CS is based on a M-estimator (see Section (3.3.3))
- `max_iters`: maximum number of iterations allowed (overrides the threshold (3.4), default = ∞)
- `min_iters`: minimum number of iterations required (overrides the threshold (3.4), default = 0)
- `max_no_updates`: maximum number of iterations with no updates (default = 0)
- `reestimate`: true to resestimate the parameter vector using all the detected inliers (default = false)
- `fix_seed`: set to true to fix the seed of the random number generator so that the results on the same dataset are *repeatable* (default = false)
- `verbose`: true for verbose output (default = true)
- `notify_iters`: if the verbose output is enabled this parameter specifies the maximim number of iterations that will occur before a progress messages is displayed (default is empty).

Output Parameters

- `results`: a structure containing the results of the estimation. Its fields are the following:

  - `Theta`: The vector of the estimates of the parameters of the model (see Section 3.2).
  - `E`: A $1 \times N$ vector containing the fitting error for each data point calculated using the function `man_fun`.
  - `CS`: A $1 \times N$ logical vector whose entries are true for the data points that have been labelled as inliers.
  - `J`: Overall ranking of the solution (see Section 3.3.3).
  - `iter`: Number of iterations performed by RANSAC before converging.

- `results`: this structure collects the options for the algorithm, after the default values have been set.

4.2 Some Examples Using the RANSAC Toolbox

4.2.1 Estimating Lines

In this section we will describe how to estimate a line that fits a set of 2D points in presence of outliers utilizing RANSAC. We will approach the problem in two different ways. The first one, called *algebraic regression*, aims at minimizing an algebraic constraint, whereas the second approach, called *orthogonal regression* provides the maximum likelihood estimate of the parameters (see the introductory Section 2.1.1).

Parameter Estimation (Algebraic Regression)

The implicit model of a line is $\theta_1 x_1 + \theta x_2 + \theta_3 = 0$. The scaling of the parameter vector $\boldsymbol{\theta}$ is immaterial, therefore we just need two equations (in other words two points will suffice to estimate a line, i.e. the cardinality of the MSS is 2). In case we are dealing with N points the following equation must hold:

$$\sum_{j=1}^{N} \left(\theta_1 \boldsymbol{x}_1^{(j)} + \theta_2 \boldsymbol{x}_2^{(j)} + \theta_3 \right)^2 = 0$$

This implies that:

$$\theta_3 = -\frac{1}{N} \sum_{j=1}^{N} \left(\theta_1 \boldsymbol{x}_1^{(j)} + \theta_2 \boldsymbol{x}_2^{(j)} \right) = -\theta_1 \bar{x} - \theta_2 \bar{y} \tag{4.1}$$

where the barred quantities indicate the sample means. If we define:

$$\Delta^{(j)} \stackrel{\text{def}}{=} \theta_1 \boldsymbol{x}_1^{(j)} + \theta_2 \boldsymbol{x}_2^{(j)} + \theta_3 = \theta_1 \tilde{\boldsymbol{x}}_1^{(j)} + \theta_2 \tilde{\boldsymbol{x}}_2^{(j)}$$

(where the tilde denotes the corresponding quantities after the mean value has been removed) in order to estimate the line of interest we can minimize the cost function:

$$J = \sum_{j=1}^{N} \left(\Delta^{(j)} \right)^2 \tag{4.2}$$

To avoid the trivial solution given by $\theta_1 = \theta_2 = 0$ we can perform the minimization subject to the constraint $\theta_1^2 + \theta_2^2 = 1$. If we construct the following matrix:

$$A \stackrel{\text{def}}{=} \begin{bmatrix} \tilde{\boldsymbol{x}}_1^{(1)} & \tilde{\boldsymbol{x}}_2^{(1)} \\ \vdots & \vdots \\ \tilde{\boldsymbol{x}}_1^{(N)} & \tilde{\boldsymbol{x}}_2^{(N)} \end{bmatrix} \tag{4.3}$$

the minimization problem can be rewritten in matrix form as:

$$\boldsymbol{\psi}^* = \underset{\substack{\theta_1,\theta_2\in\mathbb{R} \\ \theta_1^2+\theta_2^2=1}}{\operatorname{argmin}} J = \underset{\substack{\boldsymbol{\psi}\in\mathbb{R}^2 \\ \|\boldsymbol{\psi}\|=1}}{\operatorname{argmin}} \|A\boldsymbol{\psi}\|^2 \tag{4.4}$$

where:

$$\boldsymbol{\psi} \stackrel{\text{def}}{=} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}$$

This type of problems can be readily solved using the SVD decomposition of A (see Appendix B.4.2). Note that once we estimate $\boldsymbol{\psi}$ we can immediately recover θ_3 from (4.1). If we are willing to sacrifice some accuracy in the estimation of the parameters, we can solve the previous problem in closed form. The SVD decomposition of the matrix A is related to the eigen decomposition of $A^T A$ as stated in Lemma 1. Therefore the solution of the problem is obtained calculating the eigenvector corresponding to the smallest eigenvalue. A quick way to compute these quantities is described in Appendix B.3. The source code that implements the line parameter estimation is listed in Appendix D.2.1.

Remark 2. *The least square problem to estimate the parameters of the line can be set up in a different manner. If we define the matrix A to be:*

$$A \stackrel{\text{def}}{=} \begin{bmatrix} \boldsymbol{x}_1^{(1)} & \boldsymbol{x}_2^{(1)} & 1 \\ \vdots & \vdots & \\ \boldsymbol{x}_1^{(N)} & \boldsymbol{x}_2^{(N)} & 1 \end{bmatrix}$$

then we need to solve the new constrained minimization problem:

$$\boldsymbol{\theta}^* = \underset{\substack{\boldsymbol{\theta}\in\mathbb{R}^3 \\ \|\boldsymbol{\theta}\|=1}}{\operatorname{argmin}} \|A\boldsymbol{\theta}\|^2$$

This formulation is thoretically equivalent to (4.4), however practically it is not numerically sound. Without digging too much into the analytical details, the quantities that appear in A are not well scaled, more specifically the presence of ones in the last column (a quantity known with no error uncertainty) is likely to have a different order of magnitude than the remaining components of the matrix (the measurements affected by noise).

Parameter Estimation (Orthogonal Regression)

This method returns the maximum likelihood estimate of the parameters of the line. As explained in Section 2.1.1, the final goal is to solve the non linear optimization problem:

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}}\, J(\boldsymbol{\theta}) = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \sum_{i=1}^{N} \rho\left(e_{\mathcal{M}}(\boldsymbol{x}_i;\boldsymbol{\theta})\right)$$

The solution can be obtained via an iterative descent method [Lue03, Ber99]. Most of these methods require the computation of the gradient. One possibility is to evaluate the gradient numerically, but the results are usually more accurate if the expression of the gradient is computed in closed form. Because of the chain rule we can write:

$$\frac{\partial J}{\partial \theta_j}(\boldsymbol{\theta}) = \sum_{i=1}^{N} \frac{\partial \rho}{\partial e_{\mathcal{M}}}\left(e_{\mathcal{M}}(\boldsymbol{x}_i;\boldsymbol{\theta})\right) \frac{\partial e_{\mathcal{M}}}{\partial \theta_j}(\boldsymbol{x}_i;\boldsymbol{\theta})$$

Since $e_{\mathcal{M}}(\boldsymbol{x};\boldsymbol{\theta}) = \frac{\theta_1 x_1 + \theta_2 x_2 + \theta_3}{\sqrt{\theta_1^2+\theta_1^2}}$, after some simple algebra, we have:

$$\begin{aligned}
\frac{\partial e_{\mathcal{M}}}{\partial \theta_1}(\boldsymbol{x};\boldsymbol{\theta}) &= \frac{x_1}{\sqrt{\theta_1^2+\theta_2^2}} - \frac{\theta_1}{\theta_1^2+\theta_2^2} e_{\mathcal{M}}(\boldsymbol{x};\boldsymbol{\theta}) \\
\frac{\partial e_{\mathcal{M}}}{\partial \theta_2}(\boldsymbol{x};\boldsymbol{\theta}) &= \frac{x_2}{\sqrt{\theta_1^2+\theta_2^2}} - \frac{\theta_2}{\theta_1^2+\theta_2^2} e_{\mathcal{M}}(\boldsymbol{x};\boldsymbol{\theta}) \\
\frac{\partial e_{\mathcal{M}}}{\partial \theta_3}(\boldsymbol{x};\boldsymbol{\theta}) &= \frac{1}{\sqrt{\theta_1^2+\theta_2^2}}
\end{aligned}$$

The values of the gradient of estimator ρ associated to the Gaussian error distribution is:

$$\frac{\partial \rho}{\partial e_{\mathcal{M}}}\left(e_{\mathcal{M}}(\boldsymbol{x}_i;\boldsymbol{\theta})\right) = \frac{e_{\mathcal{M}}(\boldsymbol{x}_i;\boldsymbol{\theta})}{\sigma_\eta^2}$$

whereas for the Cauchy–Lorentz distribution is:

$$\frac{\partial \rho}{\partial e_{\mathcal{M}}}\left(e_{\mathcal{M}}(\boldsymbol{x}_i;\boldsymbol{\theta})\right) = \frac{\frac{e_{\mathcal{M}}(\boldsymbol{x}_i;\boldsymbol{\theta})}{\sigma_\eta}}{1+\frac{1}{2}\left(\frac{e_{\mathcal{M}}(\boldsymbol{x}_i;\boldsymbol{\theta})}{\sigma_\eta}\right)^2}$$

Error Estimation

The fitting error is defined as the squared signed distance between the point $\boldsymbol{x}$ and the line $\mathcal{M}(\boldsymbol{\theta})$:

$$e_{\mathcal{M}}(\boldsymbol{x};\boldsymbol{\theta})^2 = \frac{(\theta_1 x_1 + \theta_2 x_2 + \theta_3)^2}{\theta_1^2 + \theta_2^2}$$

If we assume that the point $\boldsymbol{x}$ is affected by Gaussian noise, then $e_{\mathcal{M}}(\boldsymbol{x};\boldsymbol{\theta})^2$ is χ^2_2 distributed. The function that computes the fitting error of the data is listed in Appendix D.2.1.

4.2.2 Estimating Planes

We will describe how it is possible to utilize RANSAC to identify points in $\mathbb{R}^3$ that belong to a plane in the space (i.e. to an affine linear manifold in $\mathbb{R}^3$) and simultaneously estimate the parameters of such manifold.

Parameter Estimation

The implicit model of the plane containing the points $\boldsymbol{x}^{(1)},\ldots,\boldsymbol{x}^{(N_I)}$ can be described by a system of linear equations:

$$\begin{cases} \theta_1 x_1^{(1)} + \theta_2 x_2^{(1)} + \theta_3 x_3^{(1)} + \theta_4 &= 0 \\ \vdots & \vdots \quad \vdots \\ \theta_1 x_1^{(N)} + \theta_2 x_2^{(N)} + \theta_3 x_3^{(N)} + \theta_4 &= 0 \end{cases}$$

It is convenient to group these equations in matrix form, so that we can write:

$$\begin{bmatrix} \left(\boldsymbol{x}^{(1)}\right)^T & 1 \\ \vdots & \vdots \\ \left(\boldsymbol{x}^{(N)}\right)^T & 1 \end{bmatrix} \boldsymbol{\theta} = A\boldsymbol{\theta} = 0 \tag{4.5}$$

Therefore the parameter vector that instantiates the plane containing the points $\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(N_I)}$ is given by:

$$\boldsymbol{\theta}^* = \underset{\substack{\boldsymbol{\theta}\in\mathbb{R}^4 \\ \|\boldsymbol{\theta}\|=1}}{\operatorname{argmin}} \|A\boldsymbol{\theta}\|^2 \tag{4.6}$$

Note that the model has only *three* degrees of freedom, since the length of the parameter vector is immaterial. The solution of (4.6) can be readily obtained using the SVD decomposition of A (see Appendix B.4.2). Since a plane is uniquely defined at least by three points, the cardinality of the MSS is $k = 3$. Therefore, using the notation introduced in Section 3.2, the model manifold (i.e. an affine space) can be expressed as:

$$\mathcal{M}(\boldsymbol{\theta}) \stackrel{\text{def}}{=} \left\{ \boldsymbol{x} \in \mathbb{R}^3 : \begin{bmatrix} \boldsymbol{x}^T & 1 \end{bmatrix} \boldsymbol{\theta} = 0 \right\}$$

The Matlab™ & Octave function that implements the parameter estimation routine is listed in Appendix D.2.2.

Error Estimation

The fitting error is defined as the distance between the point $\boldsymbol{x}$ and the manifold $\mathcal{M}(\boldsymbol{\theta})$. For each data point $\boldsymbol{x}$ the unique solution of $\operatorname{argmin}_{\boldsymbol{x}'\in\mathcal{M}(\boldsymbol{\theta})} \|\boldsymbol{x} - \boldsymbol{x}'\|^2$ can be obtained using the method of the Lagrange multipliers. The squared distance between $\boldsymbol{x}$ and its orthogonal projection onto $\mathcal{M}(\boldsymbol{\theta})$ (i.e. the plane) is given by the well know Line expression:

$$e(\boldsymbol{x}, \mathcal{M}(\boldsymbol{\theta}))^2 = \frac{\left(\begin{bmatrix} \boldsymbol{x}^T & 1 \end{bmatrix} \boldsymbol{\theta} \right)^2}{\|\boldsymbol{\theta}_{1:3}\|^2}$$

If we assume that the point $\boldsymbol{x}$ is affected by Gaussian noise, then $e(\boldsymbol{x}, \mathcal{M}(\boldsymbol{\theta}))^2$ is χ^2_3 distributed. The function that computes the fitting error of the data is listed in Appendix D.2.2.

4.2.3 Estimating a Rotation Scaling and Translation

We will describe how it is possible to utilize RANSAC to identify points in $\mathbb{R}^2$ that are related via a rotation, a scaling and a translation (RST) and simultaneously estimate the parameters of such transformation. The functional form of the RST

transformation (in Euclidean coordinates) is given by:

$$\boldsymbol{T}_{\boldsymbol{\theta}}(\boldsymbol{y}) = s \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \boldsymbol{y} + \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} \tag{4.7}$$

The parameter vector is $\boldsymbol{\theta} = \begin{bmatrix} \theta_1 & \theta_2 & \theta_3 & \theta_4 \end{bmatrix}^T = \begin{bmatrix} s\cos\phi & s\sin\phi & t_1 & t_2 \end{bmatrix}^T$.

Parameter Estimation

The parameters of an RST transformation can be estimated in a least square sense after reordering (4.7) as:

$$\underbrace{\boldsymbol{T}_{\boldsymbol{\theta}}(\boldsymbol{y})}_{\boldsymbol{b}} = \underbrace{\begin{bmatrix} y_1 & -y_2 & 1 & 0 \\ y_2 & y_1 & 0 & 1 \end{bmatrix}}_{A} \boldsymbol{\theta} = \begin{bmatrix} S & I \end{bmatrix} \boldsymbol{\theta} \tag{4.8}$$

where S is skew symmetric and I is the identity. Each point correspondence contributes for two equations and since the total number of parameters is four we need at least two point correspondences. Thus the cardinality of the MSS is $k = 2$. To estimate the parameter models starting from the MSS we need to solve a system in four equations and four unknowns. In principle we could use the same least square approach[1] used to solve the overdetermined linear system obtained from more than two corresponding point pairs.

However the parameter estimation using solely the elements in the MSS should be *as fast as possible* since it needs to be performed in every iteration of RANSAC, as shown by equation (3.13) (whereas the estimation starting from the CS is in general performed only at the end). We will demonstrate how a more careful analysis of the previous equations can lead to a relevant improvement in the performance. Grouping the equations (4.8) obtained using two point correspondences we can write:

$$\begin{bmatrix} \boldsymbol{T}_{\boldsymbol{\theta}}(\boldsymbol{y}^{(1)}) \\ \boldsymbol{T}_{\boldsymbol{\theta}}(\boldsymbol{y}^{(2)}) \end{bmatrix} = \begin{bmatrix} S^{(1)} & I \\ S^{(2)} & I \end{bmatrix} \boldsymbol{\theta}$$

If we define the matrix $M \stackrel{\text{def}}{=} S^{(1)} - S^{(2)}$ (which is nothing but the Schur complement

[1] Least squares can be solved efficiently via Gaussian elimination. This is implemented in Matlab™ & Octave using the compact notation `x=A\b;`. See also Appendix B.4.1.

of the bottom left identity), we can decompose the previous equation as:

$$\begin{bmatrix} \boldsymbol{T}_{\boldsymbol{\theta}}(\boldsymbol{y}^{(1)}) \\ \boldsymbol{T}_{\boldsymbol{\theta}}(\boldsymbol{y}^{(2)}) \end{bmatrix} = \begin{bmatrix} I & I \\ 0 & I \end{bmatrix} \begin{bmatrix} M & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ S^{(2)} & I \end{bmatrix} \boldsymbol{\theta} = \begin{bmatrix} M & I \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ S^{(2)} & I \end{bmatrix} \boldsymbol{\theta}$$

Bringing the matrices that multiply the parameter vector on the left hand side of the previous equality we get:

$$\begin{bmatrix} I & 0 \\ -S^{(2)} & I \end{bmatrix} \begin{bmatrix} M^{-1} & -M^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} \boldsymbol{T}_{\boldsymbol{\theta}}(\boldsymbol{y}^{(1)}) \\ \boldsymbol{T}_{\boldsymbol{\theta}}(\boldsymbol{y}^{(2)}) \end{bmatrix} = \\ = \begin{bmatrix} I & 0 \\ -S^{(2)} & I \end{bmatrix} \begin{bmatrix} M^{-1} \left(\boldsymbol{T}_{\boldsymbol{\theta}}(\boldsymbol{y}^{(1)}) - \boldsymbol{T}_{\boldsymbol{\theta}}(\boldsymbol{y}^{(2)}) \right) \\ \boldsymbol{T}_{\boldsymbol{\theta}}(\boldsymbol{y}^{(2)}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\theta}_{1:2} \\ \boldsymbol{\theta}_{3:4} \end{bmatrix}$$

Thus we can recover the parameter vector via simple algebraic operations and back substitutions, first computing $\boldsymbol{\theta}_{1:2} = M^{-1} \left(\boldsymbol{T}_{\boldsymbol{\theta}}(\boldsymbol{y}^{(1)}) - \boldsymbol{T}_{\boldsymbol{\theta}}(\boldsymbol{y}^{(2)}) \right)$ and then $\boldsymbol{\theta}_{3:4} = -S^{(2)}\boldsymbol{\theta}_{1:2} + \boldsymbol{T}_{\boldsymbol{\theta}}(\boldsymbol{y}^{(2)})$. Note that $M^{-1} = \frac{1}{|M|} M^T$ and $|M| = M_{1,1}^2 + M_{1,2}^2$. The total count of basic algebraic operations is listed in Table 4.1. The parameter vector can be estimated performing 12 multiplications and 11 additions. The *time unit* column shows at which stage the operation can be performed (assuming that the operations for each single task can be executed in parallel). The overall estimation will require just 6 clock cycles (on modern CPUs capable of streamed operations). The Matlab™ & Octave function that computes the parameter vector (`estimate_RST.m`) is listed in Appendix D.2.3.

Table 4.1: Counting operations for the RST parameter estimation using the MSS

Task	Multiplications	Additions	Time Unit
M	0	2	1
$\lvert M \rvert$	2	1	2
M^{-1}	2	0	3
$\boldsymbol{T}_{\boldsymbol{\theta}}(\boldsymbol{y}^{(1)}) - \boldsymbol{T}_{\boldsymbol{\theta}}(\boldsymbol{y}^{(2)})$	0	2	1
$\boldsymbol{\theta}_{1:2}$	4	2	4
$-S^{(2)}\boldsymbol{\theta}_{1:2}$	4	2	5
$\boldsymbol{\theta}_{3:4}$	0	2	6
	12	11	

Note that the normalization procedure described in Section C.2[2] allows us to greatly simplify the estimation of the parameters also when the number of points is larger than two (i.e. the cardinality of the MSS). Let's first recall the following definition:

$$d_{av} \stackrel{\text{def}}{=} \frac{1}{N}\sum_{i=1}^{N} \|\boldsymbol{y}^{(i)} - \boldsymbol{y}_c\|$$

and let's set $s \stackrel{\text{def}}{=} \frac{\sqrt{2}}{d_{av}}$ so that $\bar{\boldsymbol{y}}^{(i)} \stackrel{\text{def}}{=} s\left(\boldsymbol{y}^{(i)} - \boldsymbol{y}_c\right)$. Mutatis mutandis, the normal equations originated from (4.8) are:

$$\left(\bar{A}^T\bar{A}\right)\boldsymbol{\theta} = \bar{A}^T\boldsymbol{b}$$

where the bar indicates that the coordinates have been normalized. To simplify the notation for the next formula let's define $\boldsymbol{p}^{(i)} = \bar{\boldsymbol{y}}^{(i)}$ and $\boldsymbol{q}^{(i)} = \overline{\boldsymbol{T}_{\boldsymbol{\theta}}(\boldsymbol{y}^{(i)})}$. Because of the normalization, simple algebra is enough to reveal the diagonal structure of the normal matrix:

$$\bar{A}^T\bar{A} = \begin{bmatrix} a & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & N & 0 \\ 0 & 0 & 0 & N \end{bmatrix}$$

where $a = \sum_{i=1}^{N}\left((p_1^{(i)})^2 + (p_2^{(i)})^2\right)$. Hence:

$$\begin{aligned} \theta_1 &= \frac{1}{a}\sum_{i=1}^{N} p_1^{(i)}q_1^{(i)} + p_2^{(i)}q_2^{(i)} \\ \theta_2 &= \frac{1}{a}\sum_{i=1}^{N} -p_2^{(i)}q_1^{(i)} + p_1^{(i)}q_2^{(i)} \\ \theta_3 &= 0 \\ \theta_4 &= 0 \end{aligned}$$

Note that $\theta_3 = \theta_4 = 0$, since they express the translation in the normalized coordinate frame. The last step consist in removing the effect of the normalization from the parameters (refer to equation (C.3) for more details). The procedure just illustrated is clearly more efficient (in terms of computational complexity) than solving an

[2]The discussion is related to the estimation of homographies, but the procedure is valid in more general contexts.

overdetermined linear system (either via QR factorization or SVD decomposition).

4.2.4 Estimating Homographies

In this section we will study in more detail the design of a RANSAC algorithm to estimate the parameters of an homography given a set of point correspondences (of course, possibly contaminated by outliers). Before continuing we recall that an homography is a linear transformation in the projective space that relates two views of a planar scene obtained via an ideal pin-hole camera. The functional form of the homographic transformation (in Euclidean coordinates) is given by the *nonlinear* relation:

$$\boldsymbol{T}_{\boldsymbol{\theta}}(\boldsymbol{y}) = \begin{bmatrix} \frac{\theta_1 y_1 + \theta_4 y_2 + \theta_7}{\theta_3 y_1 + \theta_6 y_2 + \theta_9} \\ \frac{\theta_2 y_1 + \theta_5 y_2 + \theta_8}{\theta_3 y_1 + \theta_6 y_2 + \theta_9} \end{bmatrix}$$

We also recall that an homography is parameterized by nine numbers but it only has *eight* degrees of freedom, since scaling in the projective space is immaterial.

Parameter Estimation

The parameters of the homography can be estimated via the normalized DLT algorithm, which is discussed in some more detail in Appendix C. The Matlab™ & Octave function that implements the parameter estimation routine is listed in Appendix D.2.4.

The Sideness Constraint

As we saw in Section 3.4 the complexity of RANSAC depends also on the effort needed to estimate the parameters starting from the MSS (see Equation(3.13)). It should be noticed that certain configurations of the points that compose the MSS are not "physically" meaningful even though they would yield mathematically acceptable estimates of the parameters. Sometimes detecting this pathological configurations is computationally cheap if compared to the effort of estimating the parameters from the MSS. In this section we will present a constraint that must be satisfied in order to produce meaningful homographies[3].

[3]Note that the only requirement for an homography is to be non-singular. Here we refer to the fact that an homography should represent a valid transformation between images.

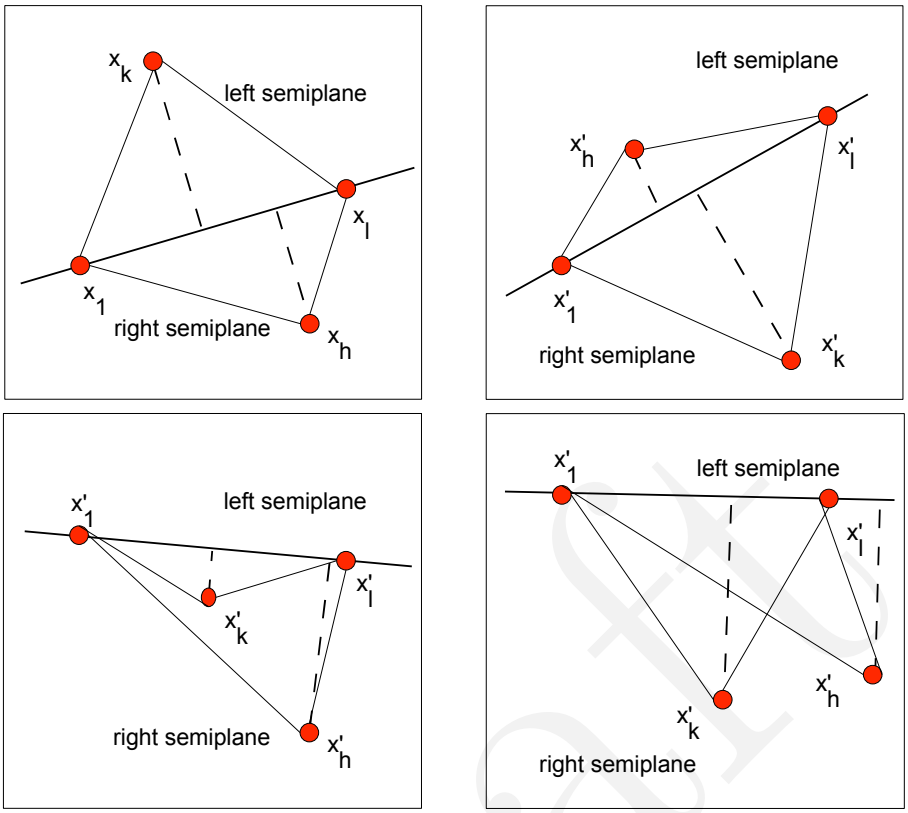


Figure 4.1: The top left figure displays the arrangements of the points composing the MSS in the first image and the line used to compute the sideness constraint. The remaining figures show possible configurations of the corresponding points where the sidness constraint is violated.

Consider a MSS formed by the following 4 point correspondences:

$$MSS = \left\{ \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}'_1 \end{bmatrix}, \ldots, \begin{bmatrix} \boldsymbol{x}_4 \\ \boldsymbol{x}'_4 \end{bmatrix} \right\}$$

Consider the line passing through the points $\boldsymbol{x}_1$ and $\boldsymbol{x}_l$ such that the points $\boldsymbol{x}_h$ and $\boldsymbol{x}_k$ lie in different semi-planes (as long as there do not exist collinear triplets it is always possible to find an arrangement of the distinct indices l, h and k in $\{2, 3, 4\}$ such that the previous condition is satisfied). The fundamental observation is that the points $\boldsymbol{x}'_h$ and $\boldsymbol{x}'_k$ must lie with respect to the line passing trough $\boldsymbol{x}'_1$ and $\boldsymbol{x}'_l$ on the corresponding semi-planes containing the points $\boldsymbol{x}_h$ and $\boldsymbol{x}_k$. This also implies that the polygon formed by the properly oriented sequence of vertexes must remain convex. Figure 4.1 describes pictorially the previous condition and displays cases when it

is violated. The test for the sideness constraint is quite simple. First we need to identify the line passing through the points $\boldsymbol{x}_1$ and $\boldsymbol{x}_l$ so that $\boldsymbol{x}_h$ and $\boldsymbol{x}_k$ belong to two different semi-planes. Then we compute a vector orthogonal to such line. A vector that will serve to this purpose is given by $\boldsymbol{n}_{1,l} = \begin{bmatrix} x_{l,2} - x_{1,2} & x_{1,1} - x_{l,1} \end{bmatrix}^T$. To check on which semi-plane the points $\boldsymbol{x}_h$ and $\boldsymbol{x}_k$ lie we just need to evaluate the sign of the expressions $\boldsymbol{n}_{1,l}^T(\boldsymbol{x}_h - \boldsymbol{x}_1)$ and $\boldsymbol{n}_{1,l}^T(\boldsymbol{x}_k - \boldsymbol{x}_1)$ and the analogous expressions for the corresponding points. *If the signs remain consistent then the MSS is acceptable.* Note that in the worst case scenario 32 summations and 16 multiplications are needed. In general the cost of these operations is less than the cost of estimating the homography via the nDLT algorithm. Again, we want to emphasize that this approach can save a lot of computations if it possible to develop a validation test which is (much) faster than the estimation of the parameters utilizing the elements in the MSS. The code that validates the MSS based on the sidenes constraint can be found in Appendix D.2.4.

Error Estimation

The correspondence error between the i^{th} pair of points related by an homography can be expressed using the *symmetric transfer error*:

$$e_i^2 \stackrel{\text{def}}{=} \|\boldsymbol{x}_i' - \boldsymbol{T}_{\boldsymbol{\theta}}(\boldsymbol{x}_i)\|^2 + \|\boldsymbol{x}_i - \boldsymbol{T}_{\boldsymbol{\theta}}^{-1}(\boldsymbol{x}_i')\|^2 \tag{4.9}$$

Assuming that the errors $\boldsymbol{\eta} = \|\boldsymbol{x}_i - \boldsymbol{T}_{\boldsymbol{\theta}}^{-1}(\boldsymbol{x}_i')\|$ and $\boldsymbol{\eta}' = \|\boldsymbol{x}_i' - \boldsymbol{T}_{\boldsymbol{\theta}}(\boldsymbol{x}_i)\|$ can be modeled as Gaussian random vectors whose components are independent and identically distributed $\mathcal{N}(\boldsymbol{0}, \sigma_\eta I)$, then the distribution of e_i^2 is a Chi-square distribution with 4 degrees of freedom (each error contributing for two degrees of freedom). Therefore, from (3.3), we have that $\delta = \sigma_\eta \sqrt{F_{\chi_4^2}^{-1}(P_{inlier})}$. Figure 3.2(b) displays the error threshold for different values of P_{inlier}. As expected from the previous discussion, the larger is the probability to find all the inliers, the larger is the error threshold. The function that computes the fitting error of the data is listed in Appendix D.2.4:

Maximum Number of Iterations

As we saw in equation (3.7), the total number of iterations is essentially proportional to the ratio of the number of inliers over the number of outliers, once the cardinality of the minimum sample set is fixed. For the nDLT algorithm, the caridnality of the

MSS is 4: i.e. we need at least four good point correspondences to estimate the homography. Figure **??** displays the number of iterations for differnt value of the false alarm rate ε and different values of the ratio $\frac{N_I}{N}$. TO BE FINISHED 1: Add figure.

4.3 Frequently Asked Questions

4.3.1 What is the "right" value of σ?

There is no general answer for this question. The value of σ is used to derive the threshold that discriminates between inliers and outliers if the noise affecting the inliers is Gaussian. Thus the "right" value of σ depends on the nature of the data you are dealing with. Sometimes you may want to set directly the noise threshold: to this purpose you may want to use the option `T_noise_squared`.

4.3.2 I want to estimate the parameters of my favourite model. What should I do?

In general there are two main aspects that one needs to consider:

- the model parametrization should be *minimal*, in the sense that the parameters should not be dependent on each other and,
- it should be easy enough to produce a relation that maps the input data contained in the MSS to the model parameters. Note that this operation needs to be repeated a large number of times (see Section 3.4) and therefore has a major impact on the performance of RANSAC.

4.3.3 How do I use the toolbox for image registration purposes?

The toolbox and more specifically the RST and homography estimation routines can be used for registration purposes in order to refine a set of image correspondences. The toolbox itself *does not* provide the routines to detect/describe the features and to establish preliminary matches. This can be done using two packages which can be freely downloaded and which are based on the SIFT framework developed by Dr. Lowe. The original implementation can be found at `http://www.cs.ubc.ca/~lowe/keypoints/`. An excellent implementation with source code developed by Dr. Vedaldi can be obtained from `http://vision.ucla.edu/~vedaldi/code/sift/sift.html`.

4.3.4 Why the behaviour of RANSAC is not repeatable?

Because of the intrinsic nature of the algorithm itself. RANSAC draws the elements composing the MSS randomly (with or without the a bias) from the entire dataset. Therefore at each run the behaviour might change. To obtain repeatable performance you should fix the seed of the internal random number generator. To accomplish this task set the option `fix_seed` to `true`.

4.3.5 What should I do if I find a bug in the toolbox?

Please contact me at `marco.zuliani@gmail.com` sending me an email with a brief description of the bug, the options that you are using and the data that generates the bug. I'll try to get back to you as soon as possible.

4.3.6 Are there any other RANSAC routines for Matlab?

Yes. Here is a list:

- P. Kovesi has an implementation of RANSAC and some routines to estimate an homography and the fundamental matrix: `http://www.csse.uwa.edu.au/~pk/Research/MatlabFns/`.
- K. Marinichev wrote the GML RANSAC Matlab Toolbox: `http://research.graphicon.ru/machine-learning/gml-ransac-matlab-toolbox-2.html`.
- P. Torr developed the Structure and Motion Toolkit for Matlab which contains some implementations of RANSAC: `http://www.mathworks.com/matlabcentral/fileexchange/4576`.

A Notation

- Bold letters indicate vectors, e.g. $\boldsymbol{x} \in \mathbb{R}^n$. Unless otherwise stated, vector components are indicated with the same non bold letter indexed starting from 1, i.e. $\boldsymbol{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$. By default vectors are column vectors.
- The hat over a variable indicates the estimate of such variable, e.g. $\hat{\boldsymbol{\theta}}$ is the estimate of the vector $\boldsymbol{\theta}$.
- To indicate portions of matrices and/or vectors Matlab™ & Octave is used. For example $A_{1:p,1:q}$ indicates the sub-matrix obtained retaining the firs p rows and the first q columns of A.

B

Some Linear Algebra Facts

The results contained in this section are presented in a very accessible way in [Str88, Mey01] and analyzed in greater details in [GV96, HJ99]. Good references for optimization are [Lue03, Ber99] and [Kel99, BV08] which can be downloaded respectively from `http://www.siam.org/books/kelley/fr18/index.php` and `http://www.stanford.edu/~boyd/cvxbook/`. Here we will provide a list of some basic results that have been tailored for the the applications described in the previous chapters.

B.1 The Singular Value Decomposition

Theorem 1 (SVD Decomposition)**.** *Let $A \in \mathbb{R}^{m\times n}$ be a matrix of rank r. Then there exist matrices $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$ and $\Sigma_1 \in \mathbb{R}^{r\times r}$ such that:*

- *V is a* unitary *matrix whose columns form a complete orthonormal basis of eigenvectors of A^TA. V can be partitioned as $\begin{bmatrix} V_1 & V_2 \end{bmatrix}$, where $V_1 \in \mathbb{R}^{n\times r}$, so that:*
  - *$\mathcal{R}(V_1) = \mathcal{R}(A^T)$ and the columns of V_1 form an orthonormal basis of $\mathcal{R}(A^T)$*
  - *$\mathcal{R}(V_2) = \mathcal{N}(A)$ and the columns of V_2 form an orthonormal basis of $\mathcal{N}(A)$*
- *U is a* unitary *matrix whose columns form a complete orthonormal basis of eigenvectors of AA^T. U can be partitioned as $\begin{bmatrix} U_1 & U_2 \end{bmatrix}$, where $U_1 \in \mathbb{R}^{m\times r}$, so that:*

  - $\mathcal{R}(U_1) = \mathcal{R}(A)$ *and the columns of* U_1 *form an orthonormal basis of* $\mathcal{R}(A)$
  - $\mathcal{R}(U_2) = \mathcal{N}(A^T)$ *and the columns of* U_2 *form an orthonormal basis of* $\mathcal{N}(A^T)$

- $\Sigma_r = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r) \in \mathbb{R}^{r \times r}$ *such that* $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$. *The scalars* σ_i *are called* singular values *of* A.

  - A *has a* dyadic expansion*:*

$$A = U_1 \Sigma_1 V_1 = \sum_{i=1}^{r} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^T$$

  - A *has a* singular value decomposition*:*

$$A = U \Sigma V^T = U \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} V^T$$

B.2 Relation Between the SVD Decomposition and the Eigen Decomposition

Lemma 1. *Consider a matrix* $A \in \mathbb{R}^{m \times n}$ *with* $m \geq n$. *Then the squared singular values of* A *coincide with the eigenvalues of* $A^T A$ *and the right singular vectors of* A *coincide with the eigenvectors of* $A^T A$.

Proof. Consider the SVD decomposition $A = U \Sigma V^T$. Then, because of the orthogonality of U, the following chain of equations hold:

$$A^T A = \left(U \Sigma V^T\right)^T U \Sigma V^T = V \Sigma^T \Sigma V^T$$

The previous chain of equations proves the claim regarding the relation between eigenvectors and singular vectors, whereas the relation between eignevalues and singular values is established observing that $\Sigma \Sigma^T = \text{diag}\{\sigma_1^2, \ldots, \sigma_n^2\}$. □

B.3 Fast Diagonalization of Symmetric 2×2 Matrices

Consider a real symmetric matrix $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$. We want to diagonalize A using the least number of operations, so that:

$$A = U\Lambda U^T = \begin{bmatrix} C & -S \\ S & C \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} C & S \\ -S & C \end{bmatrix}$$

where $C = \cos\phi$ and $S = \sin\phi$. The eigenvalues can be computed computing the roots of the characteristic polynomial $P(\lambda) = \lambda^2 - (a-c)\lambda + (ac - b^2)$ first defining:

$$\begin{aligned} \alpha &\stackrel{\text{def}}{=} a + c \\ \beta &\stackrel{\text{def}}{=} (a-c)^2 \\ \gamma &\stackrel{\text{def}}{=} 4b^2 \\ \delta &\stackrel{\text{def}}{=} \sqrt{\beta + \gamma} \end{aligned}$$

and then letting:

$$\lambda_1 = 0.5(\alpha + \delta)$$

$$\lambda_2 = 0.5(\alpha - \delta)$$

for a total number of 5 additions, 5 multiplications and the extraction of a square root.[1] Finally the angle ϕ that defines the eigenvectors univocally (modulo a reflection about the origin) can be obtained recalling the definition of eigenvector:

$$(A - \lambda_1 I)\boldsymbol{u} = \begin{bmatrix} a - \lambda_1 & b \\ b & c - \lambda_1 \end{bmatrix} \begin{bmatrix} C \\ S \end{bmatrix} = 0$$

Since $(\lambda_1 - a)C = bS$ then:

$$\phi = \arctan \frac{\lambda_1 - a}{b}$$

The angle ϕ is the angle formed by the eigenvector corresponding to the largest eigenvalue with the positive x axis. Interestingly enough, the angle can also be

[1] As expected, if $b = 0$ then $\lambda_i = \frac{1}{2}(a + c + |a - c|)$ which returns the diagonal elements of the matrix a and c.

expressed just in terms of the matrix entries,[2] without requiring the computation of the eigenvalues:

$$\phi = -\frac{1}{2}\arctan\frac{2b}{c-a}$$

As a side note, we observe that we can easily compute the 2-norm condition number of A, which is defined as the ratio of the largest singular value of A to the smallest:

$$\kappa(A) = \frac{\alpha+\delta}{\alpha-\delta}$$

Large condition numbers indicate a nearly singular matrix.

B.4 Least Square Problems Solved via SVD

As we saw in the examples in Section 4.2, RANSAC depends on the solution of least square problems of the form $A\boldsymbol{\theta} = \boldsymbol{b}$, where the matrix A and the vector $\boldsymbol{b}$ are obtained from noisy measurements. If $\boldsymbol{b}$ happens to be the null vector it is a common practice to constrain the norm of $\boldsymbol{\theta}$ in order to avoid trivial solutions. In the next sections we will informally discuss how both problems can be solved using the SVD decomposition described in Section B.1.

B.4.1 Solving $A\boldsymbol{\theta} = \boldsymbol{b}$

Let's assume that the matrix $A \in \mathbb{R}^{m\times n}$, with $m \geq n$, is full rank. Since the number of equations in larger than the number of unknowns the solution must be intended in a least square sense, i.e. we are seeking for the vector $\boldsymbol{\theta}^*$ such that:

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta}\in\mathbb{R}^n}{\operatorname{argmin}} \|A\boldsymbol{\theta} - \boldsymbol{b}\|^2 \tag{B.2}$$

The function $\|A\boldsymbol{\theta} - \boldsymbol{b}\|^2$ is convex [BV08] and since $\|\boldsymbol{x}\|^2 = \boldsymbol{x}^T\boldsymbol{x}$, we need to find the point where the gradient of $\boldsymbol{\theta}^T A^T A\boldsymbol{\theta} - 2\boldsymbol{\theta}^T A^T\boldsymbol{b} + \boldsymbol{b}^T\boldsymbol{b}$ is zero. This happens when

[2] The proof requires some algebra and here we will just outline the main steps. First observe that $\tan\theta = \frac{c-a+\delta}{2b}$. Then, because of the double angle tangent identity, observe that:

$$\tan 2\theta = \frac{2\tan\theta}{1-(\tan\theta)^2} = \frac{4b(c-a+\delta)}{4b^2 - [(c-a)+\delta)]^2}$$

It is also true that $\delta^2 = 4b^2 + (c-a)^2$ and therefore the denominator of the previous equation, by adding and subtracting $(c-a)^2$ simplifies to $-2(c-a)(c-a+\delta)$, which allows to immediately prove the result.

$\boldsymbol{\theta}$ satisfies the expression $A^T A\boldsymbol{\theta} = A^T\boldsymbol{b}$. Since we assumed that the matrix A is full rank, $A^T A$ is full rank too, its inverse exists and we can write:

$$\boldsymbol{\theta}^* = (A^T A)^{-1} A^T \boldsymbol{b} = A^\dagger \boldsymbol{b} \tag{B.3}$$

The matrix $A^\dagger$ is called pseudo-inverse (or Moore-Penrose inverse) of A for obvious reasons. If we plug in (B.3) the SVD decomposition of A and recall that the inverse of a unitary matrix coincides with its transpose, we obtain that:

$$\boldsymbol{\theta}^* = V\Sigma^{-1}_{1:n,1:n} U^T \boldsymbol{b}$$

In Matlab™ & Octave the previous expression can be quickly computed either using the function `pinv` (i.e. `theta = pinv(A)*b;`) or the operator \ (i.e. `theta = A\b;`). When $m \geq n$ and A is full rank the solution is the same (but only in this case).

B.4.2 Solving $A\boldsymbol{\theta} = \mathbf{0}$ subject to $\|\boldsymbol{\theta}\| = 1$

Let's assume that the matrix $A \in \mathbb{R}^{m\times n}$, with $m \geq n$, has rank $n-1$ (if the rank is smaller similar considerations hold). In this case we are seeking for the vector that solves the optimization problem:

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta}\in\mathbb{R}^n \text{ such that } \|\boldsymbol{\theta}\|=1}{\operatorname{argmin}} \|A\boldsymbol{\theta}\|_2 \tag{B.4}$$

Since we assumed that the rank of A is $n-1$, the right null space of the matrix has at most dimension 1. Therefore the SVD of such matrix can be written as:

$$A = U\Sigma V^T = \sum_{i=1}^{n} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_j^T = \begin{bmatrix} \boldsymbol{u}_1 & \dots & \boldsymbol{u}_{n-1} & \boldsymbol{u}_n \end{bmatrix} \begin{bmatrix} \sigma_1 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & \sigma_{n-1} & 0 \\ 0 & \dots & 0 & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_1^T \\ \vdots \\ \boldsymbol{v}_{n-1}^T \\ \boldsymbol{v}_n^T \end{bmatrix}$$

where σ_i denotes the i^{th} singular value and $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ are the corresponding left and right singular vectors. We conclude that the non-trivial solution is given by $\boldsymbol{\theta}^* = \boldsymbol{v}_n$ (i.e. the basis for the right null space of A), in fact $\|A\boldsymbol{\theta}^*\|_2 = \|A\boldsymbol{v}_n\|_2 = 0$.

But what happens if the matrix A is obtained from noisy measurements? Most likely the matrix will become full rank and then we should expect $\|A\boldsymbol{\theta}^*\| > 0$. How-

ever the solution still coincides with the right singular vector corresponding to the smallest singular value (in fact now $\sigma_n > 0$). To understand why let's consider again the minimization problem (B.4). We need to minimize $\|U\Sigma V^T\boldsymbol{\theta}\|_2$ or equivalently $\|\Sigma V^T\boldsymbol{\theta}\|_2$, since orthogonal matrices do not modify the 2-norm of a vector (intuitively rotating or reflecting a vector does not modify its length). Because of this we can also write that $\|V^T\boldsymbol{\phi}\|_2 = \|\boldsymbol{\phi}\|_2$. Hence if we let $\boldsymbol{\phi} = V^T\boldsymbol{\theta}$, the minimization problem in (C.2) can be rewritten as:

$$\boldsymbol{\phi}^* = \underset{\boldsymbol{\phi}\in\mathbb{R}^n \text{ such that } \|\boldsymbol{\phi}\|=1}{\operatorname{argmin}} \|\Sigma\boldsymbol{\phi}\|_2$$

But Σ is a diagonal matrix with its entries sorted in decreasing order. Hence the solution is $\boldsymbol{\phi} = \begin{bmatrix} 0 & \dots & 0 & 1 \end{bmatrix}^T$. Since $\boldsymbol{\phi} = V^T\boldsymbol{\theta}$, this is equivalent to pick again the last column of the matrix V, i.e. the right singular vector corresponding to the smallest singular value.

C The Normalized Direct Linear Transform (nDLT) Algorithm

C.1 Introduction

In the ideal case we would like to estimate the parameter vector $\hat{\boldsymbol{\theta}}$ such that for every point i the following homographic relation (expressed in Euclidean coordinates) holds:

$$\boldsymbol{x}'_i = \boldsymbol{T}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{x}_i) = \begin{bmatrix} \frac{\theta_1 x_{1,i} + \theta_4 x_{2,i} + \theta_7}{\theta_3 x_{1,i} + \theta_6 x_{2,i} + \theta_9} \\ \frac{\theta_2 x_{1,i} + \theta_5 x_{2,i} + \theta_8}{\theta_3 x_{1,i} + \theta_6 x_{2,i} + \theta_9} \end{bmatrix}$$

Expanding the previous equation we obtain two equations that are *linear* in the parameter components:

$$\begin{aligned} (\theta_3 x_{1,i} + \theta_6 x_{2,i} + \theta_9)\, x'_{1,i} &= \theta_1 x_{1,i} + \theta_4 x_{2,i} + \theta_7 \\ (\theta_3 x_{1,i} + \theta_6 x_{2,i} + \theta_9)\, x'_{2,i} &= \theta_2 x_{1,i} + \theta_5 x_{2,i} + \theta_8 \end{aligned}$$

The previous equations can be rearranged in matrix form as:

$$\begin{bmatrix} x_{1,i} & 0 & -x_{1,i}x'_{1,i} & x_{2,i} & 0 & -x_{2,i}x'_{1,i} & 1 & 0 & -x'_{1,i} \\ 0 & x_{1,i} & -x_{1,i}x'_{2,i} & 0 & x_{2,i} & -x_{2,i}x'_{2,i} & 0 & 1 & -x'_{2,i} \end{bmatrix} \boldsymbol{\theta} = A(\boldsymbol{x}_i, \boldsymbol{x}'_i)\boldsymbol{\theta} = \mathbf{0} \tag{C.1}$$

Thus we have two equations for nine unknowns (which have just 8 degrees of freedom since the scaling for an homography is immaterial). If we stack one upon the other the matrices $A(\boldsymbol{x}_i, \boldsymbol{x}'_i)$ we obtain the over-determined (i.e. more equations than

unknowns) homogeneous (i.e. the right hand side is zero) linear system:

$$\begin{bmatrix} A(\boldsymbol{x}_1, \boldsymbol{x}'_1) \\ \vdots \\ A(\boldsymbol{x}_N, \boldsymbol{x}'_N) \end{bmatrix} \boldsymbol{\theta} = A\boldsymbol{\theta} = \mathbf{0}$$

Hence one valid solution is obtained solving:

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\mathbb{R}^9 \text{ such that } \|\boldsymbol{\theta}\|=1}{\operatorname{argmin}} \|A\boldsymbol{\theta}\|_2 \tag{C.2}$$

This previous minimization problem can be solved using the singular value decomposition (SVD) as explained in Section B.4.2. The algorithm just illustrated is know as *Direct Linear Transform*.

Remark 3. *The cross product of a vector with itself is zero. In a noise free case we can write* $\boldsymbol{x} \times \boldsymbol{T}_{\boldsymbol{\theta}}(\boldsymbol{x}) = 0$. *Starting from this observation, we can develop an alternative version of the DLT algorithm*[1]. *The solution boils down again to the minimization of an homogeneous over-determined linear system:* $B(\boldsymbol{x}_i, \boldsymbol{x}'_i)\boldsymbol{\theta} = \mathbf{0}$. *The computation of the entries of the matrix* B *is trivial, however the details can be found in [HZ03].*

C.2 Point Normalization

In this paragraph we will try to intuitively assess the problem of numerical stability of the DLT algorithm. The interested reader is referred to [Har97] for a more thorough treatment of the subject, which requires some advanced concepts of numerical linear algebra[2]. The numerical stability of the algorithm refers to its behavior in presence of noise affecting the point position.

Quantitatively, the numerical stability of a problem (or, equivalently, of the algorithm used to solve the problem) is measured by the *condition number*. Such quantity is a measure of that problem's suitability to digital computation. A problem with a small condition number is said to be well-conditioned, while a problem with a

[1] These two different formulations are selected in the function `HomographyDLT` by appropriately setting the parameter `mode`.

[2] An excellent overview of what every computer scientist should know regarding floating-point arithmetic can be found at the following link: `http://docs.sun.com/source/806-3568/ncg_goldberg.html`.

large condition number is said to be ill-conditioned. In the case of the homography estimation, the condition number associated to the problem corresponds to the condition number of the matrix A, which can be defined as the ratio between the largest eigenvalue to the smallest one.

$$\kappa(A) = \frac{\sigma_{max}(A)}{\sigma_{min}(A)}$$

The reason behind this fact is summarized in a well known result from perturbation theory that examines how the solution of a least square problem is affected by perturbations in the coefficient matrix (and possibly in the vector of measurements). For a more general treatment of the subject refer to [GV96], p. 242.

The major reason for the poor condition of the matrix A is the fact that its entries may differ of several orders of magnitude. For our simple analysis it is convenient to consider the matrix $A^T A \in \mathbb{R}^{9\times 9}$. It is easy to prove that the eigenvectors of $A^T A$ coincide with the right singular vectors of A and the relation between eigenvalues and singular values is given by $\lambda_i = \sigma_i^2$. Thus we can write:

$$\kappa(A) = \sqrt{\frac{\lambda_{max}(A^T A)}{\lambda_{min}(A^T A)}}$$

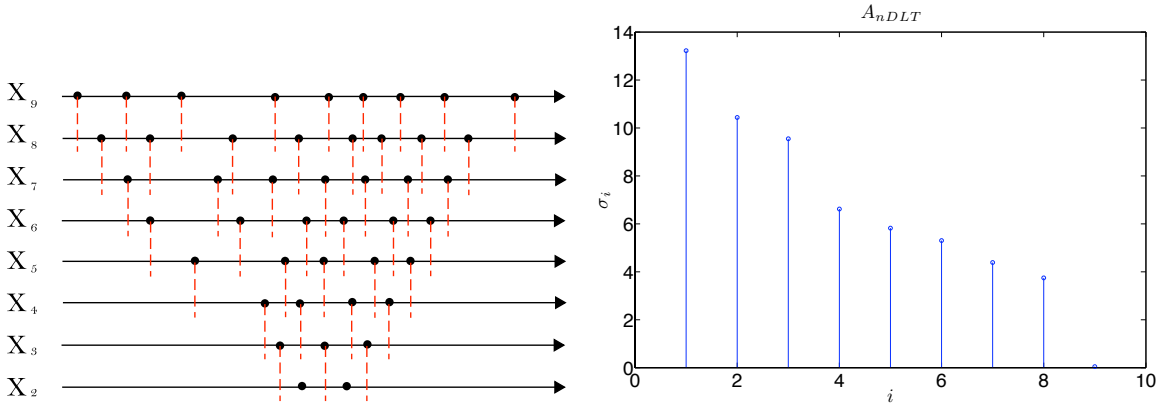
Our goal is now to obtain a rough estimate on the bounds for the condition number. Following [Har97], let's denote by X_r the the $r \times r$ principal submatrix (i.e. the matrix obtained retaining only the first r rows and columns) of $A^T A$. Since $X_9 = A^T A$, then:

$$\kappa(A) = \sqrt{\frac{\lambda_{max}(X_9)}{\lambda_{min}(X_9)}}$$

A very important result of matrix analysis is the *interlacing property of the eigenvalues* (see [GV96], p. 396). This theorem states that the when we remove a row/-column from X, the eigenvalues of the original matrix X and of the reduced matrix X_r interlace, i.e. :

$$\lambda_9(X_9) \leq \lambda_8(X_8) \leq \lambda_8(X_9) \leq \ldots \leq \lambda_2(X_9) \leq \lambda_1(A_8) \leq \lambda_1(X_9)$$

If we iterate the removal process the interlacing pattern of the eigenvalues is depicted in Figure C.1(a). Now let's consider the coordinates of a point in a typical digital



(a) (b)

Figure C.1: Figure (a) shows the interlacing pattern of the eigenvalues for a symmetric matrix. The black dots represent the position of the eigenvalues on the real line and the red vertical lines indicate the bounds on their location when we progressively remove one row and one column from the initial matrix. Figure (b) shows the singular values of the matrix A_{nDLT} for the example in Section C.3. Note that the dynamic range of the singular values is reasonably small and that the ninth one is much smaller than the eighth one ($\sigma_8 \approx 3.7457$ v.s. $\sigma_9 \approx 0.0448$, as expected since the matrix rank is 8).

image: they have an order of magnitude of 10^2. Thus the order of magnitude of the components of the diagonal of the matrix $A(\boldsymbol{x}_i, \boldsymbol{x}_i')^T A(\boldsymbol{x}_i, \boldsymbol{x}_i')$ is:

$$\begin{bmatrix} 10^2 & 10^2 & 10^4 & 10^2 & 10^2 & 10^4 & 1 & 1 & 10^2 \end{bmatrix}$$

Note that the matrix $A^T A$ is obtained summing all the matrices for the different point correspondences (thus the order of magnitude of the coefficients on the diagonal of $A^T A$ does not vary significantly for the purpose of our analysis). It follows that the order of magnitude of the coefficients of the diagonal of X_2 is 10^2. Since the trace of a matrix is the sum of the eigenvalues we conclude that $\text{trace}(X_2) = \lambda_1(X_2) + \lambda_2(X_2)$ has approximately an order of magnitude of 10^2 and since the eigenvalues are always positive (the matrix $A^T A$ is positive semidefinite by construction) because of the interlacing property we can write:

$$\lambda_9(X_9) \leq \lambda_2(X_2) \leq 10^2$$

Moreover the largest eigenvalue of $A^T A$ must be not less than the largest diagonal

entry of A^TA. This immediately follows from the interlacing properties of the singular values. We just need to notice that there exist a permutation matrix that leaves the spectrum of A^TA unaltered and moves the largest diagonal entry at the beginning of the diagonal. If, with a little abuse of notation, we consider X_1 to be the 1×1 matrix containing the largest diagonal entry of A^TA, indeed (from Figure C.1) the largest eigenvalue of A^TA (i.e. of X_9) must be not less than the largest diagonal entry of A^TA. The order of magnitude of the latter is 10^4 and therefore:

$$\lambda_1(X_9) \geq 10^4$$

Therefore the bound that we can set for the condition number is:

$$\kappa(A) = \sqrt{\frac{\lambda_{max}(X_9)}{\lambda_{min}(X_9)}} \geq \sqrt{\frac{10^4}{10^2}} = 10$$

which is approximately the square root of the order of magnitude of the coordinates of the points.

To make the DLT algorithm more stable we will lower the condition number by a *normalization procedure* (also known as *preconditioning* by the numerical analysis community). The idea consist in transforming the point coordinates *before applying the DLT algorithm* (i.e. before constructing the matrix A) by centering them about the centroid $\boldsymbol{x}_c \stackrel{\text{def}}{=} \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{x}_i$ and then scaling them so that their average distance from the origin is approximately equal to a small constant (in [HZ03] it is suggested to pick such constant equal to $\sqrt{2}$). This can be done by computing the average distance:

$$d_{av} \stackrel{\text{def}}{=} \frac{1}{N}\sum_{i=1}^{N} \|\boldsymbol{x}_i - \boldsymbol{x}_c\|$$

and finally setting $s \stackrel{\text{def}}{=} \frac{\sqrt{2}}{d_{av}}$ so that $\bar{\boldsymbol{x}}_i \stackrel{\text{def}}{=} s\,(\boldsymbol{x}_i - \boldsymbol{x}_c)$. Once the points have been normalized the DLT algorithm is applied and finally the resulting homography is denormalized (in order to go establish a correct mapping between the non normalized original points). More specifically, if $\bar{H}$ is the homography estimated using the normalized points, the denormalization procedure is implemented by the following

cascade of linear transformations in the projective space:

$$H = T'^{-1}\bar{H}T = \begin{bmatrix} \frac{1}{s'} & 0 & x'_{c,1} \\ 0 & \frac{1}{s'} & x'_{c,2} \\ 0 & 0 & 1 \end{bmatrix} \bar{H} \begin{bmatrix} s & 0 & -sx_{c,1} \\ 0 & s & -sx_{c,2} \\ 0 & 0 & 1 \end{bmatrix} \tag{C.3}$$

Once again note that the normalization procedure is meant to reduce the dynamic range of the entries of the matrix A and consequently, as we discussed before, to improve the conditioning of such matrix. The DLT algorithm plus the normalization procedure is known as *normalizaed DLT algorithm*.

C.3 A Numerical Example

We will illustrate what has been discussed in the previous sections using a numerical example. Suppose that the ground truth homography that describes the mapping between a set of points is given by:

$$H = \begin{bmatrix} 0.57458 & -0.00125 & -3.16022 \\ 0.09153 & 0.65266 & -2.48508 \\ -0.00014 & -0.00025 & 0.52426 \end{bmatrix}$$

We now generate a set of $N = 32$ points uniformly randomly distributed in the interval $[-500, 500] \times [-500, 500]$. The point correspondences are formed by mapping the first set via the homography H and then by perturbing the mapped set with Gaussian noise with zero mean and standard deviation $\sigma_\eta = 3$. We now estimate the homographies using first the DLT algorithm:

$$H_{DLT} = \begin{bmatrix} 0.09147 & -0.00036 & -0.55467 \\ 0.01463 & 0.10341 & -0.81616 \\ -0.00002 & -0.00004 & 0.08338 \end{bmatrix}$$

and then the normalized DLT algorithm, obtaining:

$$H_{nDLT} = \begin{bmatrix} -0.63883 & 0.00216 & 2.78957 \\ -0.10245 & -0.72486 & 3.53780 \\ 0.00016 & 0.00028 & -0.58282 \end{bmatrix}$$

The condition number of the matrices A is respectively $\kappa(A_{DLT}) \approx 412940$ and $\kappa(A_{nDLT}) \approx 295$. As expected the condition number after the normalization is much smaller. The singular values of A_{nDLT} are shown in Figure C.1(b). Note that the dynamic range of the singular values is reasonably small and that the ninth one is much smaller than the eighth one ($\sigma_8 \approx 3.7457$ v.s. $\sigma_9 \approx 0.0448$, as expected since the matrix rank is 8). The better performance obtained via the normalization is also confirmed by the value of the symmetric transfer error, which is $e_{DLT} \approx 1749$ in the first case and $e_{nDLT} \approx 1098$ in the second. The forward and backward mapping is illustrated in Figure C.2.

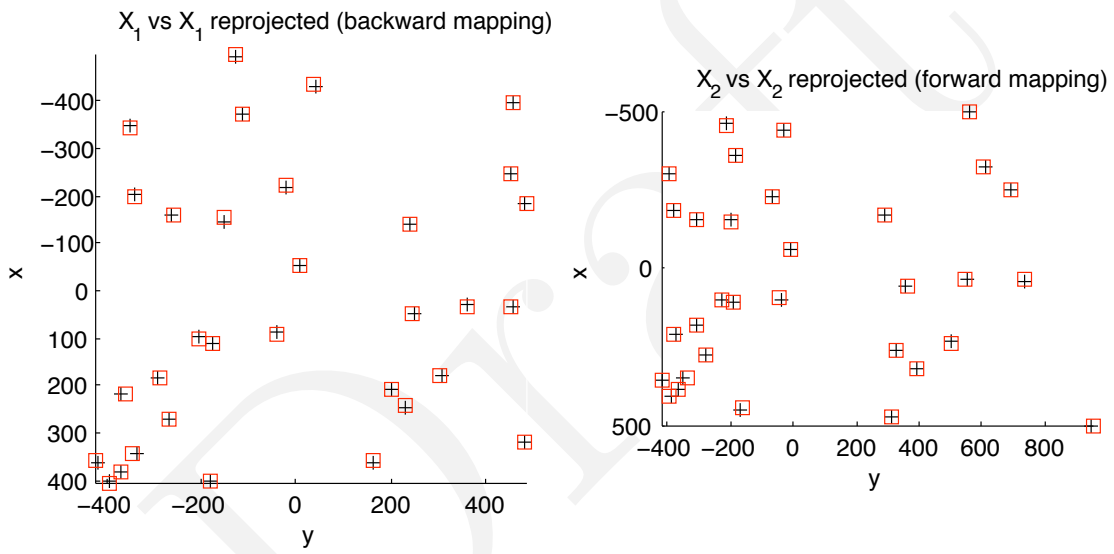


Figure C.2: The figure displays the forward and backward mapping via the homography H_{nDLT}. The crosses represent the original points whereas the squares represent the mapped points.

As a concluding remark, we want to emphasize that the normalization procedure *does not always have such a big impact* in the accuracy of the estimation of the homography, but for sure it will never cause any harm, so it is very important to normalize the points in any case (given also the fact that the additional computational complexity is not too relevant).

C.4 Concluding Remarks About the Normalized DLT Algorithm

The (normalized) DLT algorithm minimizes an algebraic quantity (C.2) that does not have an immediate correspondence with the geometry of the problem. It can be shown that this happens when the homography represents a 2D affine transformation. In the general case a common practice is to minimize the symmetric transfer error (4.9) utilizing iterative descent techniques once an initial estimate of the homography is obtained using the normalized DLT algorithm. In general this approach provides an increased accuracy (especially if the assumption regarding the contamination of Gaussian noise for the point coordinates holds true) at the expense of an increased computational complexity, of the necessity of having an initial estimate of the optimum and of the selection of a stopping criterion that is not always trivial.

A desirable property of the algorithm used to estimate the parameters of the homographic transformation is *to be invariant* with respect to certain class of transformations of the point correspondences. More specifically consider two linear transformations in the projective space, such that:

$$\begin{aligned} \bar{\boldsymbol{p}} &= T\boldsymbol{p} \\ \bar{\boldsymbol{p}}' &= T'\boldsymbol{p}' \end{aligned}$$

and suppose that the points $\boldsymbol{p}$ and $\boldsymbol{p}'$ are related via an homography: $\boldsymbol{p}' \sim H\boldsymbol{p}$. It follows immediately that the transformed points are related via a new homography $\bar{H}$:

$$\bar{\boldsymbol{p}}' \sim \underbrace{T'^{-1}HT^{-1}}_{\bar{H}}\bar{\boldsymbol{p}}$$

We are now interested in answering the following question: if we estimate both the homography H and the homography $\bar{H}$ (for the transformed points) using the DLT algorithm, will we obtain the same result? Unfortunately the answer is no. To show why this is not true is a little bit involved: more details can be found in [HZ03], p. 105. However the good news is that the minimizer of the symmetric transfer error is *invariant under similarity transformations*. To minimize the symmetric transfer error usually people resort to some iterative descent algorithm. The interest reader is once again referred to [HZ03] for more information. Here we suggest the reader to visit the site of Dr. Lourakis for a fast implementation (which can also be called from

Matlab™ & Octave) of the non linear refinement procedure based on the Levenberg-Marquardt algorithm.

A Matlab™ & Octave function that implements the normalized DLT algorithm is listed below:

```
1  function [H A] = HomographyDLT(X1, X2, mode, normalization)
2
3  % [H A] = HomographyDLT(X1, X2, mode)
4  %
5  % DESC:
6  % computes the homography between the point pairs X1, X2
7  %
8  % AUTHOR
9  % Marco Zuliani - marco.zuliani@gmail.com
10 %
11 % VERSION:
12 % 1.0.1
13 %
14 % INPUT:
15 % X1, X2        = point matches (cartesian coordinates)
16 % mode          = 0 -> Hartley Zisserman formulation
17 %                 1 -> Zuliani formulation (default)
18 % normalization = true (default) or false to enable/disable point
19 %                  normalzation
20 %
21 % OUTPUT:
22 % H             = homography
23 % A             = homogenous linear system matrix
24
25 % HISTORY
26 % 1.0.0         ??/??/04 - intial version
27 % 1.0.1         07/22/04 - small improvements
28
29 if (nargin < 3)
30     mode = 'MZ';
31 end;
32
33 if (nargin < 4)
34     normalization = true;
35 end;
```

```
36
37 N = size(X1, 2);
38
39 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
40 % checks
41 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
42 if (size(X2, 2) ≠ N)
43     error('HomographyDLT:inputError', ...
44         'The set of input points should have the same cardinality')
45 end;
46 if N < 4
47     error('HomographyDLT:inputError', ...
48         'At least 4 point correspondences are needed')
49 end;
50
51 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
52 % normalize the input
53 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
54 if normalization
55     % fprintf('\nNormalizing...')
56     [X1, T1] = normalize_points(X1);
57     [X2, T2] = normalize_points(X2);
58 end;
59
60 % compute h
61 switch mode
62     case 'HZ'
63     A = get_A_HZ(X1, X2);
64     case 'MZ'
65     A = get_A_MZ(X1, X2);
66 end;
67 try
68 [U S V] = svd(A);
69 catch
70     keyboard
71 end;
72 h = V(:, 9);
73
74 % reshape the output
75 switch mode
76     case 'HZ'
77     H = [h(1:3)'; h(4:6)'; h(7:9)'];
```

```
    case 'MZ'
    H = reshape(h, 3, 3);
end;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% de-normalize the parameters
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
if normalization
    H = T2\H*T1;
end;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% re-normalize the homography
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
H = H/norm(H(:));

return

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Matrix construction routine
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Hartley Zisserman formulation
function A = get_A_HZ(X1, X2)

X1 = cart2homo(X1);
X2 = cart2homo(X2);

N = size(X1, 2);

A = zeros(2*N, 9);
zero = [0; 0; 0];

row = 1;
for h = 1:N

    a = X2(3,h)*X1(:,h)';
    b = X2(2,h)*X1(:,h)';
    c = X2(1,h)*X1(:,h)';
    A(row, :) = [zero' -a b];
    A(row+1, :) = [a zero' -c];

```

```
120     row = row + 2;
121
122 end;
123
124 % Zuliani's formulation
125 function A = get_A_MZ(X1, X2)
126
127 N = size(X1, 2);
128
129 A = zeros(2*N, 9);
130
131 row = 1;
132 for h = 1:N
133
134     A(row, :)   = [...
135         X1(1,h) 0 -X1(1,h)*X2(1,h) ...
136         X1(2,h) 0 -X1(2,h)*X2(1,h) ...
137         1 0 -X2(1,h) ...
138         ];
139     A(row+1, :) = [...
140         0 X1(1,h) -X1(1,h)*X2(2,h) ...
141         0 X1(2,h) -X1(2,h)*X2(2,h) ...
142         0 1 -X2(2,h) ...
143         ];
144
145     row = row + 2;
146
147 end;
148
149 return
```

D Some Code from the RANSAC Toolbox

D.1 Function Templates

D.1.1 MSS Validation

```
function flag = validateMSS_foo(X, s)

% flag = validateMSS_foo(X, s)
%
% DESC:
% Validates the MSS obtained via the sampling of the data before computing
% the parameter vector Theta
%
% INPUT:
% X                 = samples on the manifold
% s                 = indices of the MSS
%
% OUTPUT:
% flag              = true if the MSS is valid

% perform here the check on the MSS

return;
```

D.1.2 Parameter Estimation

```
function [Theta, k] = estimate_foo(X, s)

% [Theta k] = estimate_foo(X, s)
%
% DESC:
% Template for the estimation function to be used inside RANSAC
%
% INPUT:
% X                 = 2D point correspondences
% s                 = indices of the points used to estimate the
%                     parameter vector. If empty all the points
%                     are used
%
% OUTPUT:
% Theta             = estimated parameter vector Theta = H(:). If
%                     the estimation is not successful return an
%                     empty vector. i.e. Theta = [];
% k                 = dimension of the minimal subset

% here we define the size of the MSS
k = ;

% check if the input parameters are valid
if (nargin == 0) || isempty(X)
    Theta = [];
    return;
end;

% select the points to estimate the parameter vector
if (nargin == 2) && ¬isempty(s)
    X = X(:, s);
end;

% check if we have enough points
N = size(X, 2);
if (N < k)
    error('estimate_foo:inputError', ...
        'At least k point correspondences are required');
```

```
39 end;
40
41 % call the estimation function foo
42 Theta = foo(X);
43
44 % here you may want to check if the results are meaningful.
45 % If not return an empty vector
46
47 return;
```

D.1.3 Parameter Validation

```
1  function flag = validateTheta_foo(X, Theta, s)
2
3  % flag = validateMSS_foo(X, Theta, s)
4  %
5  % DESC:
6  % Validates the parameter vector
7  %
8  % INPUT:
9  % X                 = samples on the manifold
10 % Theta             = parameter vector
11 % s                 = indices of the MSS
12 %
13 % OUTPUT:
14 % flag              = true if the MSS is valid
15
16 % perform here the check on the parameter vector Theta
17
18
19 return
```

D.1.4 Fitting Error

```
function [E T_noise_squared d] = error_foo(Theta, X, sigma, P_inlier)

% [E T_noise_squared d] = error_foo(Theta, X, sigma, P_inlier)
%
% DESC:
% Template to estimate the error due to the foo constraint. To
% return only the error threshold the function call should be:
%
% [dummy T_noise d] = error_foo([], [], sigma, P_inlier);
%
% INPUT:
% Theta             = foo parameter vector
% X                 = samples on the manifold
% sigma             = noise std
% P_inlier          = Chi squared probability threshold for inliers
%                     If 0 then use directly sigma.
%
% OUTPUT:
% E                 = squared error
% T_noise_squared   = squared noise threshold
% d                 = degrees of freedom of the error distribution

% compute the error obtained by the orthogonal projection of
% the data points X onto the model manifold instantiated with the
% parameters Theta
E = [];
if ¬isempty(Theta) && ¬isempty(X)

    % error computation

end;

% compute the error threshold
if (nargout > 1)

    if (P_inlier == 0)
        % in this case the parameter sigma coincides with the noise
        % threshold
```

```
39          T_noise_squared = sigma;
40      else
41          % otherwise we compute the error threshold given the standard
42          % deviation of the noise assuming that the errors are normally
43          % distributed. Hence the sum of their squares is Chi2
44          % distributed with d degrees of freedom
45          d = ;
46
47          % compute the inverse probability
48          T_noise_squared = sigma^2 * chi2inv_LUT(P_inlier, d);
49
50      end;
51
52  end;
53
54  return;
```

D.2 Source Code for the Examples

D.2.1 Line Estimation

Parameter Estimation

```
1  function [Theta, k] = estimate_line(X, s)
2
3  % [Theta k] = estimate_line(X, s)
4  %
5  % DESC:
6  % estimate the parameters of a 2D line given the pairs [x, y]^T
7  % Theta = [a; b; c] where a*x+b*yc = 0
8  %
9  % AUTHOR
10 % Marco Zuliani - marco.zuliani@gmail.com
11 %
12 % VERSION:
13 % 1.0.0
14 %
15 % INPUT:
16 % X                 = 2D points
17 % s                 = indices of the points used to estimate the parameter
18 %                     vector. If empty all the points are used
19 %
20 % OUTPUT:
21 % Theta             = estimated parameter vector Theta = [a; b; c]
22 % k                 = dimension of the minimal subset
23
24 % HISTORY:
25 % 1.0.0             = 01/26/08 - initial version
26
27 % cardinality of the MSS
28 k = 2;
29
30 if (nargin == 0) || isempty(X)
31     Theta = [];
32     return;
33 end;
34
```

```
35 if (nargin == 2) && ¬isempty(s)
36     X = X(:, s);
37 end;
38
39 % check if we have enough points
40 N = size(X, 2);
41 if (N < k)
42     error('estimate_line:inputError', ...
43         'At least 2 points are required');
44 end;
45
46 % compute the mean
47 mu = mean(X, 2);
48 % center the points
49 Xn = X - repmat(mu, 1, N);
50
51 % populate the matrix A^T A = [ a b ; b c ]
52 a = dot(Xn(1, :), Xn(1, :));
53 b = dot(Xn(1, :), Xn(2, :));
54 c = dot(Xn(2, :), Xn(2, :));
55
56 % Sacrifice accuracy for speed: compute the eigendecomposition of A^T A
57 alpha   = a+c;
58 temp    = a-c;
59 beta    = temp*temp;
60 gamma   = 4*b*b;
61 Δ    = sqrt(beta + gamma);
62 lambda  = 0.5*(alpha+Δ);
63 phi     = atan2(lambda-a,b);
64
65 % Get the eigenvector corresponding to the smallest eigenvalue
66 Theta(1,1) = -sin(phi);
67 Theta(2,1) = cos(phi);
68 Theta(3,1) = -Theta(1)*mu(1) - Theta(2)*mu(2);
69
70 return;
```

Error Estimation

```
1  function [E T_noise_squared d] = error_line(Theta, X, sigma, P_inlier)
2
3  % [E T_noise_squared d] = error_line(Theta, X, sigma, P_inlier)
4  %
5  % DESC:
6  % estimate the squared fitting error for a line expresed in cartesian form
7  % ax + by + c =0
8  %
9  % AUTHOR
10 % Marco Zuliani - marco.zuliani@gmail.com
11 %
12 % VERSION:
13 % 1.0.0
14 %
15 % INPUT:
16 % Theta             = line parameter vector
17 % X                 = samples on the manifold
18 % sigma             = noise std
19 % P_inlier          = Chi squared probability threshold for inliers
20 %                     If 0 then use directly sigma.
21 %
22 % OUTPUT:
23 % E                 = squared error
24 % T_noise_squared    = noise threshold
25 % d                 = degrees of freedom of the error distribution
26
27 % HISTORY
28 %
29 % 1.0.0             - 01/26/08 initial version
30 % 1.0.1             - 02/22/09 updated error threshold
31
32 % compute the squared error
33 E = [];
34 if ¬isempty(Theta) && ¬isempty(X)
35
36     den = Theta(1)^2 + Theta(2)^2;
37
38     E = ( Theta(1)*X(1,:) + Theta(2)*X(2,:) + Theta(3) ).^2 / den;
```

```
39
40 end;
41
42 % compute the error threshold
43 if (nargout > 1)
44
45     if (P_inlier == 0)
46         T_noise = sigma;
47     else
48         % Assumes the errors are normally distributed. Hence the sum of
49         % their squares is Chi distributed (with 2 DOF since we are
50         % computing the distance of a 2D point to a line)
51         d = 2;
52
53         % compute the inverse probability
54         T_noise_squared = sigma^2 * chi2inv_LUT(P_inlier, d);
55
56     end;
57
58 end;
59
60 return;
```

D.2.2 Plane Estimation

Parameter Estimation

```
1  function [Theta, k] = estimate_plane(X, s)
2
3  % [Theta k] = estimate_plane(X)
4  %
5  % DESC:
6  % estimate the parameters of a 3D plane given the pairs [x, y, z]^T
7  % Theta = [a; b; c; d] where:
8  %
9  % a*x1+b*y1+c*z1+d = 0
10 % a*x2+b*y2+c*z2+d = 0
11 % a*x3+b*y3+c*z3+d = 0
12 %
13 % AUTHOR
14 % Marco Zuliani - marco.zuliani@gmail.com
```

```
15 %
16 % VERSION:
17 % 1.0.0
18 %
19 % INPUT:
20 % X                 = 3D points
21 % s                 = indices of the points used to estimate the parameter
22 %                     vector. If empty all the points are used
23 %
24 % OUTPUT:
25 % Theta             = estimated parameter vector Theta = [a; b; c; d]
26 % k                 = dimension of the minimal subset
27
28 % HISTORY:
29 % 1.0.0             = 07/05/08 - initial version
30
31 % cardinality of the MSS
32 k = 3;
33
34 if (nargin == 0) || isempty(X)
35     Theta = [];
36     return;
37 end;
38
39 if (nargin == 2) && ¬isempty(s)
40     X = X(:, s);
41 end;
42
43 % check if we have enough points
44 N = size(X, 2);
45 if (N < k)
46     error('estimate_plane:inputError', ...
47         'At least 3 points are required');
48 end;
49
50 A = [transpose(X(1, :)) transpose(X(2, :)) transpose(X(3, :)) ones(N, 1)];
51 [U S V] = svd(A);
52 Theta = V(:, 4);
53
54 return;
```

Error Estimation

```
1  function [E T_noise_squared d] = error_plane(Theta, X, sigma, P_inlier)
2
3  % [E T_noise_squared d] = error_plane(Theta, X, sigma, P_inlier)
4  %
5  % DESC:
6  % estimate the squared fitting error for a plane expresed in cartesian form
7  %
8  % a*x1+b*y1+c*z1+d = 0
9  % a*x2+b*y2+c*z2+d = 0
10 % a*x3+b*y3+c*z3+d = 0
11 %
12 % AUTHOR
13 % Marco Zuliani - marco.zuliani@gmail.com
14 %
15 % VERSION:
16 % 1.0.1
17 %
18 % INPUT:
19 % Theta             = plane parameter vector
20 % X                 = samples on the manifold
21 % sigma             = noise std
22 % P_inlier          = Chi squared probability threshold for inliers
23 %                     If 0 then use directly sigma.
24 %
25 % OUTPUT:
26 % E                 = squared symmetric reprojection error
27 % T_noise_squared   = squared noise threshold
28 % d                 = degrees of freedom of the error distribution
29
30 % HISTORY
31 %
32 % 1.0.0             - 07/05/08 initial version
33 % 1.0.1             - 02/22/09 updated error threshold
34
35 % compute the squared error
36 E = [];
37 if ¬isempty(Theta) && ¬isempty(X)
38
```

```
39      den = Theta(1)^2 + Theta(2)^2 + Theta(3)^2;
40
41      E = ( ...
42          Theta(1)*X(1,:) + ...
43          Theta(2)*X(2,:) + ...
44          Theta(3)*X(3,:) + ...
45          Theta(4)...
46          ).^2 / den;
47
48  end;
49
50  % compute the error threshold
51  if (nargout > 1)
52
53      if (P_inlier == 0)
54          T_noise_squared = sigma;
55      else
56          % Assumes the errors are normally distributed. Hence the sum of
57          % their squares is Chi distributed (with 3 DOF since we are
58          % computing the distance of a 3D point to a plane)
59          d = 3;
60
61          % compute the inverse probability
62          T_noise_squared = sigma^2 * chi2inv_LUT(P_inlier, d);
63
64      end;
65
66  end;
67
68  return;
```

D.2.3 RST Estimation

Parameter Estimation

```
1  function [H s phi T] = RSTLS(X1, X2, normalization)
2
3  % [H s phi T] = RSTLS(X1, X2, normalization)
4  %
5  % DESC:
6  % computes the RST transformation between the point pairs X1, X2
7  %
8  % VERSION:
9  % 1.0.1
10 %
11 % INPUT:
12 % X1, X2         = point matches (cartesian coordinates)
13 % normalization = true (default) or false to enable/disable point
14 %                  normalzation
15 %
16 % OUTPUT:
17 % H              = homography representing the RST transformation
18 % s              = scaling
19 % phi            = rotation angle
20 % T              = translation vector
21
22
23 % AUTHOR:
24 % Marco Zuliani, email: marco.zuliani@gmail.com
25 % Copyright (C) 2011 by Marco Zuliani
26 %
27 % LICENSE:
28 % This toolbox is distributed under the terms of the GNU GPL.
29 % Please refer to the files COPYING.txt for more information.
30
31
32 % HISTORY
33 % 1.0.0          08/27/08 – intial version
34 % 1.0.1          06/09/09 – implemented closed form for the LS estimation
35 %                           routines
36
37 if (nargin < 3)
```

```
38      normalization = true;
39  end;
40
41  N = size(X1, 2);
42
43  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
44  % checks
45  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
46  if (size(X2, 2) ≠ N)
47      error('RSTLS:inputError', ...
48          'The set of input points should have the same cardinality')
49  end;
50  if N < 2
51      error('RSTLS:inputError', ...
52          'At least 2 point correspondences are needed')
53  end;
54
55  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
56  % normalize the input
57  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
58  if (normalization) && (N > 2)
59      % fprintf('\nNormalizing...')
60      [X1, T1] = normalize_points(X1);
61      [X2, T2] = normalize_points(X2);
62  end;
63
64  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
65  % estimation
66  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
67  if (N == 2)
68
69      % fast estimation
70      Theta = zeros(4,1);
71
72      % MM ≝ M_{:,1} = y^(1) − y^(2) = [ y_1^(1) − y_1^(2) ; y_2^(1) − y_2^(2) ]
73      % 2 additions
74      MM = X1(:,1) − X1(:,2);
75      % detMM ≝ |M|
76      % 1 additions, 2 multiplication
77      detMM = MM(1)*MM(1) + MM(2)*MM(2);
```

```
78      % MMi = [ [M^-1]_{1,1} ; -[M^-1]_{2,1} ]
79      % 2 multiplications
80      MMi = MM / detMM;
81
82      % Delta = T_θ(y^(1)) - T_θ(y^(2))
83      % 2 additions
84      Delta = X2(:,1) - X2(:,2);
85
86      % Theta(1:2) = M^-1 (T_θ(y^(1)) - T_θ(y^(2)))
87      % 1 additions, 2 multiplications
88      Theta(1) = MMi(1)*Delta(1) + MMi(2)*Delta(2);
89      % 1 additions, 2 multiplications
90      Theta(2) = MMi(1)*Delta(2) - MMi(2)*Delta(1);
91      % Theta(3:4) = -S^(2) θ_{1:2} + T_θ(y^(2))
92      % 2 additions, 2 multiplications
93
94      Theta(3) = X2(1,2) - Theta(1)*X1(1,2) + Theta(2)*X1(2,2);
95      % 2 additions, 2 multiplications
96      Theta(4) = X2(2,2) - Theta(1)*X1(2,2) - Theta(2)*X1(1,2);
97
98      % total: 11 additions, 12 multiplications
99  else
100
101     % Closed form LS solution. Using the tutorial notation.
102
103     % Notation semplification:
104     % p^(i) = ȳ^(i)  and  q^(i) = conj(T_θ(y^(i)))
105     % a = sum_{i=1}^N ( (p_1^(i))^2 + (p_2^(i))^2 )
106     a = sum(X1(:).^2);
107
108     % Explicit LS expansion:
109     % θ_1 = (1/a) sum_{i=1}^N p_1^(i) q_1^(i) + p_2^(i) q_2^(i)
110     % θ_2 = (1/a) sum_{i=1}^N -p_2^(i) q_1^(i) + p_1^(i) q_2^(i)
111     % θ_3 = 0
112     % θ_4 = 0
113     Theta(1) = sum( X1(1, :).*X2(1, :) + X1(2, :).*X2(2, :) ) / a;
114     Theta(2) = sum( -X1(2, :).*X2(1, :) + X1(1, :).*X2(2, :) ) / a;
115     Theta(3) = 0;
116     Theta(4) = 0;
117
```

```
118     % Traditional LS
119     %
120     %     A = zeros(2*N, 4);
121     %     b = zeros(2*N, 1);
122     %
123     %     ind = 1:2;
124     %     for n = 1:N
125     %
126     %         A(ind, 1:2) = [X1(1,n) -X1(2,n); X1(2,n) X1(1,n)];
127     %         A(ind, 3:4) = eye(2);
128     %
129     %         b(ind) = X2(1:2, n);
130     %
131     %         ind = ind + 2;
132     %
133     %     end;
134     %
135     %     % solve the linear system in a least square sense
136     %     Theta = A\b;
137
138 end;
139
140 % compute the corresponding homography
141 H = [Theta(1) -Theta(2) Theta(3); Theta(2) Theta(1) Theta(4); 0 0 1];
142
143 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
144 % de-normalize the parameters
145 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
146 if (normalization) && (N > 2)
147     H = T2\H*T1;
148 end;
149 H = H/H(9);
150
151 % prepare the output
152 if nargout > 1
153
154     s       = sqrt(H(1,1)*H(1,1) + H(2,1)*H(2,1));
155     phi     = atan2(H(2,1), H(1,1));
156     T       = H(1:2, 3);
157
158 end;
159
```

```
160 return
```

D.2.4 Homography Estimation

Parameter Estimation

```
1  function [Theta, k] = estimate_homography(X, s)
2
3  % [Theta k] = estimate_homography(X, s)
4  %
5  % DESC:
6  % estimate the parameters of an homography using the normalized
7  % DLT algorithm. Note that Theta = H(:)
8  %
9  % AUTHOR
10 % Marco Zuliani - marco.zuliani@gmail.com
11 %
12 % VERSION:
13 % 1.0.1
14 %
15 % INPUT:
16 % X                 = 2D point correspondences
17 % s                 = indices of the points used to estimate the parameter
18 %                     vector. If empty all the points are used
19 %
20 % OUTPUT:
21 % Theta             = estimated parameter vector Theta = H(:)
22 % k                 = dimension of the minimal subset
23
24 % HISTORY:
25 % 1.0.0             = ??/??/05 - initial version
26 % 1.0.1             = 27/08/08 - minor improvements
27
28 % cardinality of the MSS
29 k = 4;
30
31 if (nargin == 0) || isempty(X)
32     Theta = [];
33     return;
34 end;
35
36 if (nargin == 2) && ~isempty(s)
37     X = X(:, s);
```

```
38 end;
39
40 % check if we have enough points
41 N = size(X, 2);
42 if (N < k)
43     error('estimate_homography:inputError', ...
44         'At least 4 point correspondences are required');
45 end;
46
47 H = HomographyDLT(X(1:2, :), X(3:4, :));
48 Theta = H(:);
49
50 return;
```

```
1  function [H A] = HomographyDLT(X1, X2, mode, normalization)
2
3  % [H A] = HomographyDLT(X1, X2, mode)
4  %
5  % DESC:
6  % computes the homography between the point pairs X1, X2
7  %
8  % AUTHOR
9  % Marco Zuliani - marco.zuliani@gmail.com
10 %
11 % VERSION:
12 % 1.0.1
13 %
14 % INPUT:
15 % X1, X2         = point matches (cartesian coordinates)
16 % mode           = 0 -> Hartley Zisserman formulation
17 %                  1 -> Zuliani formulation (default)
18 % normalization = true (default) or false to enable/disable point
19 %                  normalzation
20 %
21 % OUTPUT:
22 % H              = homography
23 % A              = homogenous linear system matrix
24
25 % HISTORY
26 % 1.0.0          ??/??/04 - intial version
27 % 1.0.1          07/22/04 - small improvements
```

```
28
29 if (nargin < 3)
30     mode = 'MZ';
31 end;
32
33 if (nargin < 4)
34     normalization = true;
35 end;
36
37 N = size(X1, 2);
38
39 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
40 % checks
41 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
42 if (size(X2, 2) ~= N)
43     error('HomographyDLT:inputError', ...
44         'The set of input points should have the same cardinality')
45 end;
46 if N < 4
47     error('HomographyDLT:inputError', ...
48         'At least 4 point correspondences are needed')
49 end;
50
51 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
52 % normalize the input
53 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
54 if normalization
55     % fprintf('\nNormalizing...')
56     [X1, T1] = normalize_points(X1);
57     [X2, T2] = normalize_points(X2);
58 end;
59
60 % compute h
61 switch mode
62     case 'HZ'
63     A = get_A_HZ(X1, X2);
64     case 'MZ'
65     A = get_A_MZ(X1, X2);
66 end;
67 try
68 [U S V] = svd(A);
69 catch
```

```
    keyboard
end;
h = V(:, 9);

% reshape the output
switch mode
    case 'HZ'
    H = [h(1:3)'; h(4:6)'; h(7:9)'];
    case 'MZ'
    H = reshape(h, 3, 3);
end;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% de-normalize the parameters
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
if normalization
    H = T2\H*T1;
end;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% re-normalize the homography
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
H = H/norm(H(:));

return

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Matrix construction routine
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Hartley Zisserman formulation
function A = get_A_HZ(X1, X2)

X1 = cart2homo(X1);
X2 = cart2homo(X2);

N = size(X1, 2);

A = zeros(2*N, 9);
zero = [0; 0; 0];

row = 1;
```

```
112 for h = 1:N
113
114     a = X2(3,h)*X1(:,h)';
115     b = X2(2,h)*X1(:,h)';
116     c = X2(1,h)*X1(:,h)';
117     A(row, :) = [zero' -a b];
118     A(row+1, :) = [a zero' -c];
119
120     row = row + 2;
121
122 end;
123
124 % Zuliani's formulation
125 function A = get_A_MZ(X1, X2)
126
127 N = size(X1, 2);
128
129 A = zeros(2*N, 9);
130
131 row = 1;
132 for h = 1:N
133
134     A(row, :)   = [...
135         X1(1,h) 0 -X1(1,h)*X2(1,h) ...
136         X1(2,h) 0 -X1(2,h)*X2(1,h) ...
137         1 0 -X2(1,h) ...
138         ];
139     A(row+1, :) = [...
140         0 X1(1,h) -X1(1,h)*X2(2,h) ...
141         0 X1(2,h) -X1(2,h)*X2(2,h) ...
142         0 1 -X2(2,h) ...
143         ];
144
145     row = row + 2;
146
147 end;
148
149 return
```

MSS Validation

```
1  function flag = validateMSS_homography(X, s)
2
3  % flag = validateMSS_homography(X, s)
4  %
5  % DESC:
6  % Validates the MSS obtained sampling the data using the sideness
7  % constraintbefore computing the parameter vector Theta
8  %
9  % INPUT:
10 % X                 = samples on the manifold
11 % s                 = indices of the MSS
12 %
13 % OUTPUT:
14 % flag              = true if the MSS is valid
15
16 % HISTORY:
17 %
18 % 1.1.0             – 10/12/08 – Initial version
19
20
21 % set this to true to display invalid MSS (just for debug/didactic
22 % purposes)
23 graphic = false;
24
25 % Check if the points are in pathological configurations. Compute the
26 % covariance matrix and see if the determinant is too small (which implies
27 % the point are collinear)
28 ind = [1 2];
29 for h = 1:2
30     mu = mean(X(ind, s), 2);
31     Xzm = X(ind, s) – repmat(mu, 1, length(s));
32     C1 = Xzm(1, :)*transpose(Xzm(1, :));
33     C2 = Xzm(1, :)*transpose(Xzm(2, :));
34     C3 = Xzm(2, :)*transpose(Xzm(2, :));
35     % compute the condition number
36     alpha = C1+C3;
37     temp  = C1–C3;
38     beta  = temp*temp;
```

```
39      gamma = 4*C2*C2;
40      Δ = sqrt(beta+gamma);
41      kappa = (alpha+Δ)/(alpha-Δ);
42      if ( kappa > 1e9 )
43          flag = false;
44          return;
45      end;
46      ind = ind + 2;
47  end;
48
49  % Generate all the possible pairings for the line that determines the
50  % semi-planes. The anchor point is the first one, i.e. the one with index
51  % 1. Thus the line that defines the semiplanes can be the line passing
52  % through the points:
53  %
54  % (1,2) or
55  % (1,3) or
56  % (1,4)
57  %
58  % The remaining rows define the points that should lie in different
59  % semiplanes
60  ind = s([...
61      2 3 4; ...
62      3 2 2; ...
63      4 4 3]);
64
65  % Setting the flag to false should guard against collinearity
66  flag = false;
67  for l = 1:3
68
69      % compute the normal vector n_{1,l}
70      % 2 summations
71      n(1) = X(2, ind(1,l))-X(2, s(1));
72      n(2) = X(1, s(1))-X(1, ind(1,l));
73
74      % get the projection of the other two points
75      % 6 summations, 4 multiplications
76      p1 = n(1)*( X(1,ind(2,l)) - X(1, s(1)) ) +...
77          n(2)*( X(2,ind(2,l)) - X(2, s(1)) );
78      p2 = n(1)*( X(1,ind(3,l)) - X(1, s(1)) ) +...
79          n(2)*( X(2,ind(3,l)) - X(2, s(1)) );
80
```

```
    % if they lie on the same side then select next arrangement
    if sign(p1) == sign(p2)
        continue;
    end;

    % compute the normal vector n'_{1,l} for the corresponding
    % points
    % 2 summations
    np(1) = X(4, ind(1,l))-X(4, s(1));
    np(2) = X(3, s(1))-X(3, ind(1,l));

    % get the projection of the other two corresponding points
    % 6 summations, 4 multiplications
    pp1 = np(1)*( X(3,ind(2,l)) - X(3, s(1)) ) +...
        np(2)*( X(4,ind(2,l)) - X(4, s(1)) );
    pp2 = np(1)*( X(3,ind(3,l)) - X(3, s(1)) ) +...
        np(2)*( X(4,ind(3,l)) - X(4, s(1)) );

    % verify the sideness
    flag = (sign(p1) == sign(pp1)) && (sign(p2)==sign(pp2));

    if (graphic) && (flag == false)

        color = 'gr';
        d = 16;

        figure;

        offset = 0;
        for kk = 1:2
            subplot(1,2,kk)
            hold on
            plot(X(1+offset, s), X(2+offset, s), ...
                'o','MarkerSize', 8, ...
                'MarkerEdgeColor', 'k', ...
                'MarkerFaceColor', color(kk))
            % draw the line that separates the planes
            plot([X(1+offset, s(1)) X(1+offset, ind(1, l))], ...
                [X(2+offset, s(1)) X(2+offset, ind(1, l))], '--k');

            for hh = 1:4
                text(X(1+offset, s(hh))+d, ...
```

```
123                     X(2+offset, s(hh))+d, num2str(hh))
124             end;
125             xlabel('x');
126             ylabel('y');
127             axis equal
128             offset = offset + 2;
129         end;
130
131         pause
132     end;
133
134     break;
135
136 end;
137
138
139 return;
```

Error Estimation

```
1  function [E T_noise_squared d] = error_homography(Theta, X, sigma, P_inlier)
2
3  % [E T_noise_squared d] = error_homography(Theta, X, sigma, P_inlier)
4  %
5  % DESC:
6  % estimate the squared symmetric transfer error due to the homographic
7  % constraint
8  %
9  % AUTHOR
10 % Marco Zuliani - marco.zuliani@gmail.com
11 %
12 % VERSION:
13 % 1.0.1
14 %
15 % INPUT:
16 % Theta             = homography parameter vector
17 % X                 = samples on the manifold
18 % sigma             = noise std
19 % P_inlier          = Chi squared probability threshold for inliers
20 %                     If 0 then use directly sigma.
21 %
22 % OUTPUT:
23 % E                 = squared symmetric transfer error
24 % T_noise_squared    = squared noise threshold
25 % d                 = degrees of freedom of the error distribution
26
27 % HISTORY
28 %
29 % 1.0.0             - 11/18/06 initial version
30 % 1.0.1             - 02/22/09 updated error threshold
31
32
33 % compute the squared symmetric reprojection error
34 E = [];
35 if ¬isempty(Theta) && ¬isempty(X)
36
37     N = size(X, 2);
38
```

```
39      X12 = zeros(2, N);
40      [X12(1, :) X12(2, :)] = mapping_homography(X(1,:), X(2,:), true, Theta);
41
42      X21 = zeros(2, N);
43      [X21(1, :) X21(2, :)] = mapping_homography(X(3,:), X(4,:), false, Theta);
44
45      E1 = sum((X(1:2, :)-X21).^2, 1);
46      E2 = sum((X(3:4, :)-X12).^2, 1);
47
48      E = E1 + E2;
49  end;
50
51  % compute the error threshold
52  if (nargout > 1)
53
54      if (P_inlier == 0)
55          T_noise_squared = sigma;
56      else
57          % Assumes the errors are normally distributed. Hence the sum of
58          % their squares is Chi distributed (with 4 DOF since the symmetric
59          % distance contributes for two terms and the dimensionality is 2)
60          d = 4;
61
62          % compute the inverse probability
63          T_noise_squared = sigma^2 * chi2inv_LUT(P_inlier, d);
64
65      end;
66
67  end;
68
69  return;
```

GNU Free Documentation License

Preamble

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "**Document**", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "**you**". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "**Modified Version**" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "**Secondary Section**" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "**Invariant Sections**" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "**Cover Texts**" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "**Transparent**" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent

file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "**Opaque**".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "**Title Page**" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

A section "**Entitled XYZ**" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "**Acknowledgements**", "**Dedications**", "**Endorsements**", or "**History**".) To "**Preserve the Title**" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.

B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.

C. State on the Title page the name of the publisher of the Modified Version, as the publisher.

D. Preserve all the copyright notices of the Document.

E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.

F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.

G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.

H. Include an unaltered copy of this License.

I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.

J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.

K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.

L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.

M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.

N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.

O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties–for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements".

6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided for under this License. Any other attempt to copy, modify, sublicense or distribute the Document is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See http://www.gnu.org/copyleft/.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation.

ADDENDUM: How to use this License for your documents

To use this License in a document you have written, include a copy of the License in the document and put the following copyright and license notices just after the title page:

> Copyright © YEAR YOUR NAME. Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled "GNU Free Documentation License".

If you have Invariant Sections, Front-Cover Texts and Back-Cover Texts, replace the "with ... Texts." line with this:

> with the Invariant Sections being LIST THEIR TITLES, with the Front-Cover Texts being LIST, and with the Back-Cover Texts being LIST.

If you have Invariant Sections without Cover Texts, or some other combination of the three, merge those two alternatives to suit the situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.

Bibliography

[Ber99] D. P. Bertsekas, *Nonlinear programming*, second ed., Athena Scientific, 1999.

[BV08] S. Boyd and L. Vandenberghe, *Convex optimization*, sixth ed., Cambridge University Press, 2008.

[CM02] O. Chum and J. Matas, *Randomized RANSAC with $T_{d,d}$ test*, 13th British Machine Vision Conference, September 2002.

[CM05] _____, *Matching with PROSAC - progressive sample consensus*, Proceedings of Conference on Computer Vision and Pattern Recognition (San Diego), vol. 1, June 2005, pp. 220–226.

[FB81] M. A. Fischler and R. C. Bolles, *Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography*, Communications of the ACM **24** (1981), 381–395.

[GV96] G. H. Golub and C. F. Van Loan, *Matrix computations*, The John Hopkins University Press, 1996.

[Har97] R. I. Hartley, *In defence of the eight-point algorithm*, IEEE Transaction on Pattern Recognition and Machine Intelligence **16** (1997), no. 6, 580–593.

[HJ99] R. Horn and C. R. Johnson, *Matrix analisys*, Cambridge University Press, 1999.

[Hub81] P. J. Huber, *Robust statistics*, Wiley, 1981.

[HZ03] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*, second ed., Cambridge University Press, 2003.

[Kel99] C. T. Kelley, *Iterative methods for optimization*, SIAM, 1999.

[KK04] Y. Kanazawa and H. Kawakami, *Detection of planar regions with uncalibrated stereo using distribution of feature points*, British Machine Vision Conference (Kingston upon Thames, London), vol. 1, September 2004, pp. 247–256.

[Lue03] D. G. Luenberger, *Linear and nonlinear programming*, second ed., Addison-Wesley, 2003.

[Men95] J. Mendel, *Lessons in estimation theory for signal processing, communication and control*, Prentice-Hall, Englewood-Cliffs, 1995.

[Mey01] C. D. Meyer, *Matrix analysis and applied linear algebra*, SIAM, 2001.

[Nab03] I. Nabney, *Netlab: Algorithms for pattern recognition*, Springer, 2003.

[Nis03] D. Nistér, *Preemptive RANSAC for live structure and motion estimation*, IEEE International Conference on Computer Vision (Nice, France), October 2003, pp. 199–206.

[RL87] P. J. Rousseeuw and A. M. Leroy, *Robust regression and outlier detection*, Wiley, 1987.

[Ste95] C. V. Stewart, *MINPRAN: A new robust estimator for computer vision*, IEEE Transactions on Pattern Analysis and Machine Intelligence **17** (1995), no. 10, 925–938.

[Ste99] ———, *Robust parameter estimation in computer vision*, SIAM Review **41** (1999), no. 3, 513–537.

[Str88] G. Strang, *Linear algebra and its applications*, Academic Press, 1988.

[TF08] R. Toldo and A. Fusiello, *Robust multiple structures estimation with j-linkage*, European Conference on Computer Vision (Marseille, France), October 2008, pp. 537–547.

[TM05] B. J. Tordoff and D. W. Murray, *Guided-MLESAC: Faster image transform estimation by using matching priors*, IEEE Transactions on Pattern Analysis and Machine Intelligence **27** (2005), no. 10, 1523–1535.

[TZ00] P.H.S. Torr and A. Zisserman, *MLESAC: A new robust estimator with application to estimating image geometry*, Journal of Computer Vision and Image Understanding **78** (2000), no. 1, 138–156.

[VJFS05] A. Vedaldi, H. Jin, P. Favaro, and S. Soatto, *KALMANSAC: Robust filtering by consensus*, Proceedings of the International Conference on Computer Vision (ICCV), vol. 1, 2005, pp. 633–640.

[VL01] E. Vincent and R. Laganière, *Detecting planar homographies in an image pair*, 2nd International Symposium on Image and Signal Processing and Analysis (Pula, Croatia), June 2001, pp. 182–187.

[WS04] H. Wang and D. Suter, *Robust adaptive-scale parametric model estimation for computer vision.*, IEEE Transactions on Pattern Analysis and Machine Intelligence **26** (2004), no. 11, 1459–1474.

[Zha97] Z. Zhang, *Parameter estimation techniques: A tutorial with application to conic fitting*, Image and Vision Computing Journal **25** (1997), no. 1, 59–76.

[ZKM05] M. Zuliani, C. S. Kenney, and B. S. Manjunath, *The MultiRANSAC algorithm and its application to detect planar homographies*, IEEE International Conference on Image Processing, September 2005.