An Evaluation of the Performance of RANSAC Algorithms for Stereo Camera Calibration

A. J. Lacey, N. Pinitkarn and N. A. Thacker
Imaging Science and Biomedical Engineering,
Stopford Building, University of Manchester,
Oxford Road, Manchester, M13 9PT
a.lacey@man.ac.uk

Abstract

This paper compares the use of RANSAC for the determination of epipolar geometry for calibrated stereo reconstruction of 3D data with more conventional optimisation schemes. The paper illustrates the poor convergence efficiency of RANSAC which is explained by a theoretical relationship describing its dependency upon the number of model parameters. The need for an a-priori estimate of outlier contamination proportion is also highlighted. A new algorithm is suggested which attempts to make better use of the solutions found during the RANSAC search while giving a convergence criteria which is independent of outlier proportion. Although no significant benefit can be found for the use of RANSAC on the problem of stereo camera calibration estimation. The new algorithm suggests a simple way of improving the efficiency of RANSAC searches which we believe would be of value in a wide range of machine vision problems.

1 Introduction

Many problems in machine vision require the reconciliation of sampled data with a known model and specifically the parameterisation of this model. Given a multi-variate non-linear smooth function, denoted by $f(\mathbf{x})$, where $\mathbf{x}$ is the vector of n unknown variables, our task is to search the n-dimensional space for a location which gives the best statistical interpretation of the data. This process is often achieved using an optimisation strategy which attempts to uncover the parameter set which describes the data whilst accounting for the expected noise in the system.

The final goal of the optimisation is to recover the *global* extremity (minimum or maximum) of whatever cost function is being employed. The recovery of *local* extremities is a fundamental issue in optimisation algorithms and usually arises because either the sampled data is insufficient (unlikely to be the case in most well sampled situations) or the starting condition of the algorithm prohibits the algorithm from locating the global extremity. The optimisation process must move from one location to another, on the surface defined by the cost, towards the minimum. If the starting location for the search is not on a direct downhill path to the global minimum it will often not be found. The recovery of optimal solutions is also affected by outlier data which, by definition, cannot

BMVC 2000 doi:10.5244/C.14.65

be accounted for by the model, but also cannot be excluded from the data without extra knowledge. Unfortunately, the majority of image interpretation tasks involve the analysis of data which cannot be perfectly segmented prior to interpretation. Thus the majority of model based scene interpretation algorithms must be able to deal with outliers.

The Random Sample and Consesus (RANSAC) algorithm of Fischler and Bolles [1] searches for suitable solutions directly using the data, repeatedly constructing solutions from randomly sampled minimum subsets which are not related to any concept of an error surface and thus are not restricted to either an assumption of smoothness or the computational form of the objective function. In contrast to most optimisation algorithms which attempt to maximise the quantity of data used to identify a solution, RANSAC constructs solutions from the minimum subset of data necessary (e.g. two points for a line). Provided sufficient repetitions are performed RANSAC is expected to identify solutions computed from outlier free data. To ensure this is the case the objective function used with RANSAC must be robust to outlier data and so the use of a simple least squares metric is unsatisfactory. The scope of applicability for the algorithm is great and potentially encompasses algorithms based on optimisation with or without the use of local derivatives. The use of RANSAC as a wrapper around closed form solutions to vision problems gives potential scope for utility to these otherwise brittle (non-robust) approaches. RANSAC gives us the opportunity to evaluate any estimate of a set of parameters no matter how robust or accurate the method that generated this solution might be. The RANSAC method could thus be considered as an ideal approach to the solution of many machine vision problems. However, the random nature of the search makes direct use of RANSAC as an optimisation algorithm inefficient.

2 The Problem of Epi-Polar Estimation

Over recent years several papers have been published which have defined practical solutions for the automatic estimation of the camera motion parameters [4, 3, 12] either for calibrated stereo or motion estimation. The common goal of these algorithms is to recover the epi-polar geometry or *Essential matrix* which describes the rotation and translation of the single moving camera (ego-motion) or between two spatially separated cameras (stereopsis) from pairings of matched image features. The problem can be expressed as;

$$\mathbf{X}_{1i} = \mathbf{R}\mathbf{X}_{2i} + \mathbf{T} \tag{1}$$

where $\mathbf{X}_{1i}$ is the 3D position of camera 1, $\mathbf{X}_{2i}$ is the 3D position of camera 2 and $\mathbf{R}$ and $\mathbf{T}$ are the rotation and translation matrices necessary to move between these two. The epi-polar geometry is defined by a plane passing through the optical centres of the two cameras and a point in the world. The constraint can be formulated by manipulating equation 1 thus;

$$\mathbf{X}_{1i} \cdot (\mathbf{T} \times \mathbf{R}\mathbf{X}_{2i}) = 0 \tag{2}$$

or in terms of image plane co-ordinates;

$$x_1^T \mathbf{E} x_2 = 0 \tag{3}$$

where $\mathbf{E}$ is the Essential matrix and x_1 and x_2 are the vectors of matched feature points in image plane co-ordinates. This constraint is attributable to Longuet-Higgins [10]. If the intrinsic camera parameters are unknown and thus x_1 and x_2 are only known in pixel co-ordinates then the *Fundamental matrix* can be computed where;

$$\mathbf{F} = \mathbf{Q}_1^T \mathbf{E} \mathbf{Q}_2 \quad (4)$$

where $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are the calibration matrices of the two cameras.

Although many of the published methods for estimating $\mathbf{E}$ or $\mathbf{F}$ have practical utility there has been no systematic study into the comparative performance of these approaches. We have therefore decided to use this area as a test ground for the comparative evaluation of RANSAC as an approach to the robust identification of global optima in data with severe outlier contamination. In what follows the optimisation methods described in previous publications are compared in terms of convergence efficiency and a new method is introduced which attempts to make better use of the data available to the RANSAC algorithm.

2.1 Trivedi-Simplex

The Trivedi-Simplex algorithm [5] attempts to recover epi-polar geometry (3 rotation and 3 translation minus one scaling factor) by minimising the epi-polar errors of matched corners on the image plane. This is achieved by comparing the suggested epi-polar solution to all known matches using a robust M-Estimator [7] which accounts for the expected accuracy of the corner features in order to minimise contamination with outlier data, thus;

$$log(ML) = \sum_i MIN(9.0, \frac{-\epsilon_i \Lambda \mathbf{s_{1i}}}{\mathbf{s_{1i}} \Lambda \mathbf{s_{1i}^T} + \mathbf{s_{2i}} \Lambda \mathbf{s_{2i}^T}}) \quad (5)$$

where Λ is the covariance matrix on the feature location u_i, $\epsilon_i = u_{1i}^T \mathbf{F} u_{2i}$, $s_{1i} = u_{1i}^T \mathbf{F}$ and $s_{2i} = \mathbf{F} u_{2i}$. Equation 5 defines the objective function and the *downhill simplex* algorithm [9] is employed to recover its minimum.

2.2 RANSAC

RANSAC can be regarded as an optimisation algorithm since, given an objective function it can determine the parameter set which best positions the model. Provided that the data and the number of attempts are sufficiently large it is expected that model position identified by RANSAC is close to the global solution. RANSAC does not use any local information about the objective function. Instead it samples the objective function at discrete locations by randomly sampling minimum subsets of data and generating candidate solutions which are then evaluated. Consequently, solutions are not restricted by either the initialisation of the algorithm (which is random) or the local smoothness of the objective function (of which RANSAC has no concept).

The RANSAC algorithm used here randomly samples 7 corner matches from the pool of possible matches. These points are then used to construct a minimum estimate of the matrix $\mathbf{E}$ using the epi-polar constraint of equation 3 (possible because the intrinsic camera parameters are available). From the Essential matrix we are then able to compute the rotation-translation parameters using the algorithm of Tsai [11] as summarised in appendix A. This process is only assured to give valid solutions with the minimum number of points when the condition of the $\mathbf{W}$ matrix decomposed from $\mathbf{E}$ fulfills certain conditions [6] which in practice cannot be guaranteed. The obvious solution is to sample more data, however equation 10 of appendix B demonstrates the m-th power dependency of the RANSAC algorithm on the number of samples m. Thus using anything but the minimum number of parameters is computationally expensive. Instead we use the minimum number of points and rely on the RANSAC to disgard the rogue solutions. As before this epi-polar estimate is then compared to all the remaining corner matches by measuring the off epi-polar squared error equation 5 (within the accuracy of the corner location). This error score is used to rank the solution in the RANSAC space.

2.3 RANSAC with fusion

Traditional RANSAC is often used as a means of identifying the inlier datapoints which are then used in a more conventionally optimisation strategy to find the optimal minima. This is because RANSAC itself has no mechanism for searching around data-borne solutions.

In the RANSAC with fusion algorithm (RANSAC-f) we maintain an ordered list of the best n solutions evaluated by RANSAC. After a number of iterations this list will be populated with outlier free solutions. However, the longer the list is maintained the less likely it is that new sampled results will be competitive enough to replace the top ranked solution; a problem commonly referred to a *elitism*. Therefore, once this list is populated with good solutions the algorithm begins combining results in order to generate improved solutions. When a new solution is ranked high enough to enter the list an attempt is first made to combine it with solutions already present. Starting with the top of the list the parameters from both solutions are averaged and a new solution found. This solution is then evaluated using the same robust, epi-polar error score as used with the datapoint solutions. If the score for the fused solution is better than both the new and listed solution then the combined result is placed in a second list and the listed element removed. Before a solution is entered into this second list an attempt is made to combine it with any solutions already present. Again if the combined result is better than both the source data results then this solution is entered into a third list and so on until no further combination is possible. At any point, if a combination solution is not better than both source solutions it is simply placed in the current list at its ordered location.

The process of fusion enables solutions not defined directly by sampled data to be explored, reusing good solutions in an attempt to better the top-ranked result. Combination of results is appropriate provided that they are of a similar accuracy, i.e. have equivalent covariances; it is not worthwhile combining accurate measurements with inaccurate measurements. This can be avoided by allowing the list to develop before fusion begins and constructing the the list from relatively few solutions. The generation of a hierchical list structure is also necessary if combination is to be performed using source data of equivalent accuracies. Computationally the process of combining results is little different from sampling the data.

For a particular dataset it is possible to estimate the number of trials required by RANSAC to arrive at a concensus using the equation below (equation 6 is derived in appendix B).

$$n = \frac{\log\left(1-z\right)}{\log\left(1-(1-\epsilon)^m\right)} \tag{6}$$

where n is the number of trials, z is the confidence level, m is the number of points selected and ϵ is the outlier proportion. Therefore, it is necessary to know the proportion of outliers in the dataset which, in most cases, is difficult if not impossible to determine.

In the case of the RANSAC-f algorithm it is believed that a self termination criteria could be specified using the list update rate. By taking the ratio of the rate of update of the top list element to the rate of list entries it is possible to define a termination threshold which is independent of outliers, simply because the probability of an outlier contaminated tuple producing a good solution reduces as the search proceeds.

3 Experiments

In the experiments which follow the different algorithms are used to estimate the epi-polar geometry of 4 pairs of stereo (spatially separated) images. Point correspondences between the images are established using the corner detection and matching algorithm as described in [4]. Any remaining outliers (less than 10%) are removed by hand, leaving the numbers of correct matches as; (a) Head (figure 1) 200, (b) House (figure 2) 63, (c) Saucer (figure 3) 95, (d) Shaft Assembly (figure 4) 65. The measures which are most appropriate for evaluating the performance of a stereo camera calibration system are the error on the verge angle and the epi-polar error. The first determines the accuracy of 3D reconstruction and the latter determines the accuracy of the epi-polar constraint used during feature matching [2]. Although we cannot measure the verge error in these experiments, because we do not have a 'gold standard' answer, we can compute the off epi-polar error which is correlated with verge angle accuracy. The RMS off epi-polar error is plotted against the number of iterations in the graphs of figures 1 to 4. Although the ranking of solutions by RANSAC is done using the off epi-polar error compared with all (inlier and outlier) data, the RMS error plots are computed with inlier data only and thus reflect the deviation away from the 'true' answer.

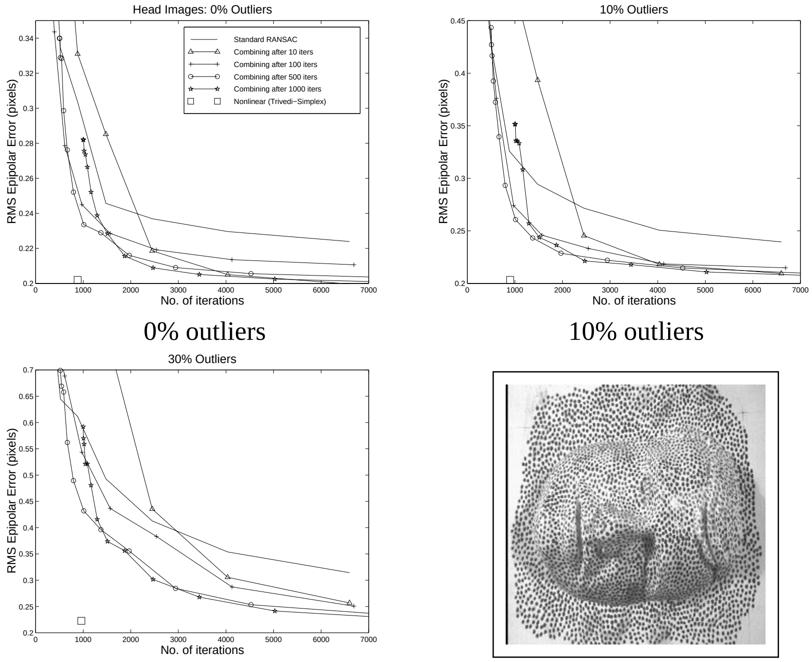
An iteration was counted for every comparison of an epi-polar solution to the inlier data. Outlier data was introduced by include mis-match corners at the prescribed percentage into the datasets.

The Trivedi-Simplex algorithm requires an initial estimate of the epi-polar geometry, chosen as parallel in all cases. Also required is a starting scale for the simplex which is problem dependent and has been selected for each image pair so as to reduce the number of failures in the outlier free case.

The epi-polar estimation was repeated 30 times and the average RMS error plotted. The RANSAC-f algorithm was executed with 4 different starting points for the fusion process, delaying it by 10, 100, 500 and 10 00 iterations.

4 Discussion

The performance of the RANSAC-f algorithm is noticeably improved over the standard RANSAC approach, often achieving comparable results with a fifth of the computational cost (iterations). Although the delay in the fusion process has little bearing on the ultimate RMS residual error, it has significant impact on the rate of decay. Likewise, however, the same is true of the Trivedi-simplex algorithm when compared to RANSAC-f. The Trevidi-simplex algorithm out-performs RANSAC-f in all cases, often achieving superior results in far fewer iterations. This is not suprising, since the Trivedi-simplex algorithm



0% outliers

10% outliers

30% outliers

Original Image

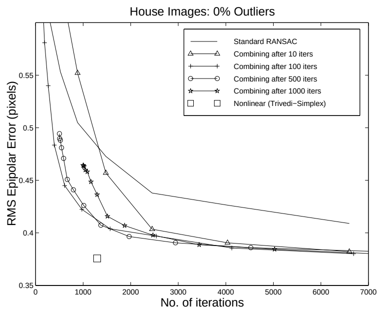
Figure 1: Head Image

is using the maximum quantity of data to estimate the solution. However, the Trivedi-simplex needs to be intialised, prior to execution, with an appropriate simplex scale if convergence is to be reliably achieved.

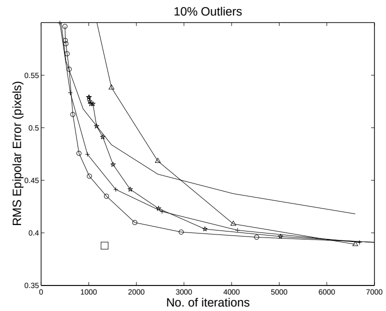
5 Conclusions

RANSAC is commonly utilised to identify an outlier free subset of data. This is only possible if a robust objective function is used and the algorithm is allowed to iterate for sufficient cycles. The traditional method of determining the termination point requires the proportion of outliers present in the data, information which is not commonly available. We have presented a variant on the RANSAC algorithm which maintains ordered lists of competitive solutions. The lists are used to uncover solutions which are not available to traditional RANSAC by combining the parameters of previous good solutions. This has been proved to be an effective technique in order to overcome the $O(m)$ dependency of traditional RANSAC.

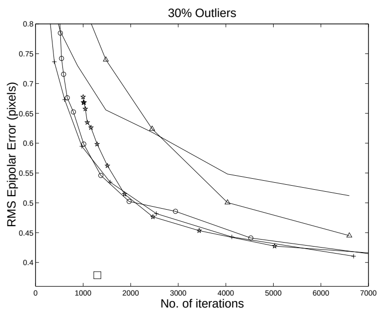
It is also suggested that maintaining such lists could be useful in self termination. Ultimately however we see that in terms of residual per iterations the conventional technique remains superior, provided that a suitable start location can be identified. For the case of stereo camera calibration it is generally quite easy to specify a good starting point. Therefore the use of RANSAC for this applcation does not appear to be justified. How-



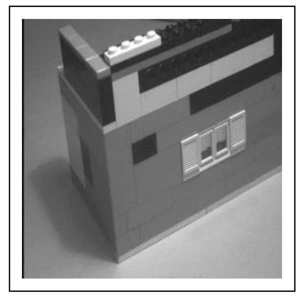
0% outliers



10% outliers



30% outliers



Original Image

Figure 2: Brick House

ever, for motion estimation, where the data may also be significantly contaminated with outliers, this argument may no longer hold. Under these circumstances RANSAC will typically require typically four times the computatinal resource to attain the same level of calibration accuracy. Ultimately the strength of the RANSAC approach lies in its ability to filter outlier data with little in the way of prior information. Thus, a more sophisticated approach which may be suitable in all circumstances would involve using the results from RANSAC-f to initialise a simplex search algorithm.

The software (TINA [13]) and data used in this research are available for free download from our website at `http://www.niac.man.ac.uk/Tina`.

A Motion Parameter Estimation

Given an Essential matrix $\mathbf{E}$ formed under the constraint (see section 2);

$$x_1^T \mathbf{E} x_2 \;=\; 0 \tag{7}$$

the motion parameters $\mathbf{R}$ and $\mathbf{T}$ can be estimated by first decomposing $\mathbf{E}$ thus;

$$\mathbf{E} \;=\; \mathbf{U}\mathbf{W}\mathbf{V}^T \tag{8}$$

and then reconstructing thus;

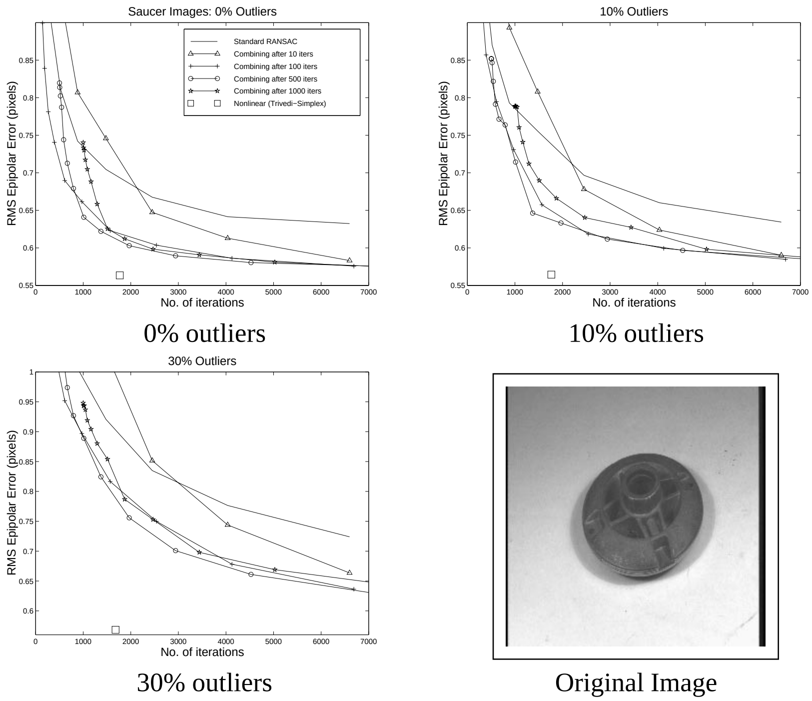


Figure 3: Saucer

$$\mathbf{R}_1 \;=\; \mathbf{U}\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & t \end{bmatrix}\mathbf{V}^T \quad \mathbf{R}_2 \;=\; \mathbf{U}\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & t \end{bmatrix}\mathbf{V}^T \quad \mathbf{T} \;=\; s\begin{bmatrix} e_1.e_2/e_2.e_3 \\ e_1.e_2/e_1.e_3 \\ 1 \end{bmatrix} \tag{9}$$

where R_1 and R_2 are the camera 1-to-2 and 2-to-1 rotation matrices, $t \;=\; det(\mathbf{U}).det(\mathbf{V})$, s is a scale factor and e_i is the i th row of the Essential matrix.

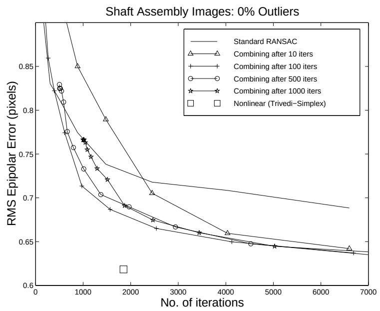
B Derivation of RANSAC Iteration Length

Given a dataset of N points of which g are inliers the probability that a randomly selected subset of m points will contain only inliers is;

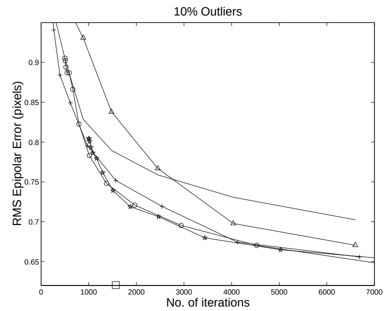
$$P(m \in g) \;=\; \frac{{}^gC_m}{{}^NC_m} \;=\; \frac{g(g-1)(g-2)...(g-m-1)}{N(N-1)(N-2)...(N-m-1)}$$

When g is much larger than m this can be approximated as;

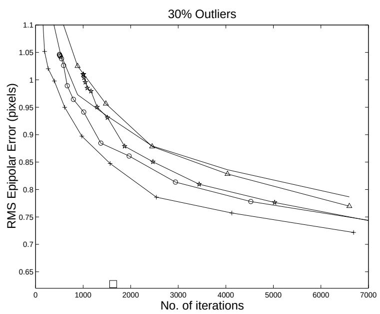
$$P(m \in g) \;\approx\; \left(\frac{g}{N}\right)^m \;=\; \left(\frac{N-b}{N}\right)^m \;=\; (1-\epsilon)^m \tag{10}$$



0% outliers



10% outliers



30% outliers

Original Image

Figure 4: Shaft Assembly

where $\epsilon = b/N$ and is the outlier proportion. The probability of finding an outlier free subset at the x-th trial is;

$$P(x) = q^{x-1}p$$

where $p = P(m \in g)$ is the probability of success and $q = 1 - p$. In n trials the probability of encountering at least one outlier free subset is;

$$z = P(1 \geq x \geq n) = \sum_{x=1}^{n} q^{x-1}p = \frac{p(q^n - 1)}{q - 1} = 1 - (1-p)^n \qquad (11)$$

Substituting p from the relationship in equation 10 and rearrange in terms of the number of trials n gives;

$$n = \frac{\log(1-z)}{\log(1-(1-\epsilon)^m)} \qquad (12)$$

References

[1] M. A. Fischler and R. C. Bolles, ”Random sample consensus: a paradigm for model fitting with application to image analysis and automated cartography”, Communication Association and Computing Machine, 24(6), pp.381-395, 1981

[2] A.J.Harris, N.A.Thacker and A.J.Lacey, Modelling Feature Based Stereo Vision for Range Sensor Simulation. proc of the European Simulation Multiconference, 417-421, Manchester, 1998.

[3] P. H. S. Torr and D. W. Murray, ”The development and comparison of robust methods for estimating fundamental matrix ”, International Journal of Computer Vision, 24(3), pp.271-300, 1997

[4] N. A. Thacker and J. E. Mayhew, ”Optimal combination of stereo camera calibration from arbitrary stereo images”, Image and Vision Computing, 9(1), pp.27-32, 1991

[5] H. P. Trivedi, ”Estimation of stereo and motion parameters using a variational principal”, Image and Vision Computing, 5(2), pp.181-183, 1987

[6] K. Kanatani, ”Statistical Optimization for Geometric Computation: Theory and Practice”, Elsevier Science, Amsterdam, 1996

[7] P. J. Huber, ”Robust Statistics”, John Wiley & Sons, New York, 1981

[8] P. J. Rousseeuw, ”Robust Regression and Outlier Detection”, Wiley, New York, 1987

[9] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, ”Numerical Recipes in C, The Art of Scientific Computing”, Cambridge University Press, second edition, New York, 1992

[10] H. C. Longuet-Higgins, ”A computer algorithm for reconstructing scene from two projections”, Nature,293(10), pp.133-135, 1981

[11] R. Y. Tsai and T. S. Huang, ”Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 6(1), pp.13-27, 1984

[12] J. Weng, T. S. Huang and N. Ahuja, ”Motion and structure from two perspective views: algorithms, error analysis and error estimation”, IEEE Trans. PAMI, 11(5), pp. 451-476, 1989

[13] TINA: Algorithm Development Libraries, http://www.niac.man.ac.uk/Tina