A Minimal Solution for Relative Pose with Unknown Focal Length

Henrik Stewénius
Lund University
Sweden
stewe@maths.lth.se

David Nistér
University of Kentucky
Lexington, USA
dnister@cs.uky.edu

Fredrik Kahl
University of California
San Diego, USA
fredrik@maths.lth.se

Frederik Schaffalitzky
University of Oxford
United Kingdom
fsm@robots.ox.ac.uk

Abstract

Assume that we have two perspective images with known intrinsic parameters except for an unknown common focal length. It is a minimally constrained problem to find the relative orientation between the two images given six corresponding points. We present an efficient solution to the problem and show that there are 15 solutions in general (including complex solutions). To the best of our knowledge this was a previously unsolved problem.

The solutions are found through eigen-decomposition of a 15×15 *matrix. The matrix itself is generated in closed form. We demonstrate through practical experiments that the algorithm is correct and numerically stable.*

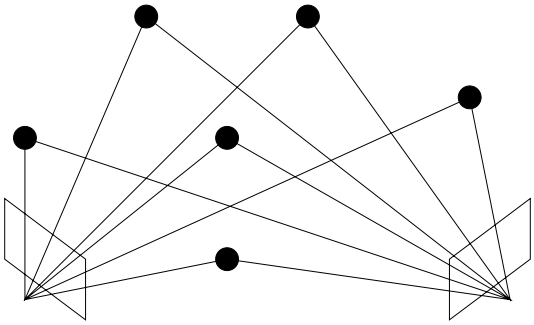

Figure 1: The problem solved here: Relative orientation for 2 cameras with a common, but unknown, focal length f that see 6 unknown points.

1 Introduction

The task of computing a 3D reconstruction from a video sequence is central in computer vision. Different paradigms have been proposed for performing this task and the concept of RANSAC has been quite successful [5]. State-of-the-art in real-time structure from motion uses the five point method, e.g., [8], as a RANSAC engine. While this has proved to be efficient and stable, the cameras need to be pre-calibrated. For uncalibrated cameras, the seven-point method can be applied, e.g., [15], but it is not as stable as the five-point method. We offer an attractive compromise having similar performance characteristics as the five point method and still allowing for unknown focal lengths. By assuming constant and unknown focal lengths, but otherwise known intrinsics, a minimal problem arises for six points. A detailed analysis of this case is given, showing the number of possible solutions, efficient ways to compute them and the stability of the solution with respect to measurement noise.

Minimal case solvers exist for several camera models. The problem for two calibrated cameras and five points was first solved by Kruppa [7] who claimed that there are at most 11 solutions and the false root was then eliminated by [4]. A practical solution was given in [10] and improved in [8]. For three views and four points, the problem is not minimal but as it would be under-constrained with three points, the problem is still of interest and was solved in [9].

Given that the epipolar geometry has been computed in terms of the fundamental matrix, it is well known that it is possible to recover the focal length [5, 14]. However, to the best of our knowledge the relative pose problem for minimal data is still unsolved. Here we present a solver for two cameras and six points. The solver constructs a 15×15 matrix in closed form. Solving the eigen-problem for this matrix gives the 15 (possibly complex) solutions to the relative pose problem. More information on how to build minimal case solvers by studying an an analogouos problem over $\mathbb{Z}_p$ can be found in [13].

We will first list the minimal cases for cameras with a common unknown focal length. Then the equations used will be introduced. The solver was found using Gröbner basis theory but this theory is not necessary to understand the solver. We will also give some numerical results.

2 Background

Suppose we are given m cameras, all calibrated except for a common unknown focal length and n corresponding image points. There are $6m+3n+1-7$ degrees of freedom (6 for each camera, 3 for each point, 1 for focal length and 7 for the unknown coordinate system) and $2mn$ equations, hence

	n						
m	1	2	3	4	5	6	7
1	-1	-2	-3	-4	-5	-6	-7
2	-5	-4	-3	-2	-1	**0**	1
3	-9	-6	-3	**0**	3	6	9
4	-13	-8	-3	2	7	12	17

Table 1: Number of excess constraints for m views and n points with unknown focal length f.

in total, there are $2mn + 6 - 6m - 3n$ excess constraints (Table 1). The minimal case $(m, n) = (2, 6)$ will be solved here. The other possibility $(m, n) = (3, 4)$ is still unsolved.

Geometric Constraints

The fundamental matrix F encodes the epipolar geometry of two views, and corresponding image points x and x' satisfy the coplanarity constraint

$$x'^{\top} F x = 0. \tag{1}$$

Any rank-2 matrix is a possible fundamental matrix, i.e. we have the well known single cubic constraint, e.g. [5]:

Theorem 1 *A fundamental matrix F satisfies*

$$det(F) = 0. \tag{2}$$

An essential matrix has the additional property that the two non-zero singular values are equal. This leads to the following cubic constraints on the essential matrix, adapted from [10]:

Theorem 2 *A real non-zero 3×3 matrix E is an essential matrix if and only if it satisfies the equation*

$$2EE^{\top}E - tr(EE^{\top})E = 0. \tag{3}$$

This constraint previously appeared in [12, 3].

2.1 Gröbner Bases

The *ideal* generated by polynomials $f_1, \ldots, f_n \in \mathbb{C}[x_1, \ldots, x_n]$ is the set I of polynomials $g \in \mathbb{C}[x_1, \ldots, x_n]$ of the form:

$$g = \sum_{i=1}^{n} f_i p_i, \qquad p_i \in \mathbb{C}[x_1, \ldots, x_n]. \tag{4}$$

We also say that the f_i *generate* the ideal I. A *Gröbner basis* of an ideal is a special set of generators, with the property that the leading term of every ideal element is divisible by the leading term of a generator. The notion of leading term is defined relative to a *monomial order*. The Gröbner basis exposes all leading terms of the ideal and leads to the useful notion of *remainder* with respect to (division by) the ideal. Gaussian elimination is a special case of *Buchberger's algorithm* which is a method for calculating a Gröbner basis from any generating set. Gröbner bases, monomial orders and Buchberger's algorithm are explained in [1]. For ideals having a finite set of solutions ("zero-dimensional" ideals) the (vector space) dimension of the quotient ring $A = \mathbb{C}[x_1, \ldots, x_n]/I$ is also finite and the dimension equals the number of solutions, counted with multiplicity. Any polynomial f acts on the quotient ring A by multiplication ($f : g + I \mapsto fg + I$) and this is clearly a linear mapping from A to itself. A natural way to choose a (vector space) basis for A is to take all monomials that are not leading terms of any element of I. The action of a polynomial f is then described by a square matrix m_f called the *action matrix*.

The solutions to a zero-dimensional ideal can be read off directly from the eigen-values and eigen-vectors of appropriate action matrices [2].

3 Solution Procedure for $(2, 6)$

The inner calibration of the camera is assumed to be

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

With observations $\{x_i\}_{i=1}^{6}$ in the first image and $\{x'_i\}_{i=1}^{6}$ in the second image the epipolar constraint (1) gives six linear constraints on the fundamental matrix F ,

$${x_i}^T F x'_i = 0\,, \forall i = 1, \ldots, 6. \tag{5}$$

As a 3×3 matrix has 9 degrees of freedom this determines F up to 3 degrees of freedom, $F = l_0F_0 + l_1F_1 + l_2F_2$ for some scalars l_0, l_1, l_2. The fundamental matrix F can be computed only up to scale so we set $l_0 = 1$.

The fundamental matrix F must fulfill (1), that is, $\det(F) = \det(F_0 + F_1l_1 + F_2l_2) = 0$, which is a third order polynomial equation in (l_1, l_2).

The matrix F can be transformed into an essential matrix by correcting for the intrinsic calibration,

$$E = K^{\top} F K.$$

Set

$$P = f^{-1}K = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & f^{-1} \end{bmatrix} \tag{6}$$

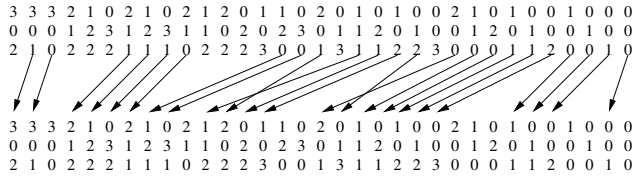

Figure 2: Order of monomials and shifts for multiplication by p. Since the arrows cross this is not a monomial order. The rows are powers of l_1,l_2 and p. The columns are in the order of the monomials.

then Equation (2) is equivalent to

$$2PFPPF^TPPFP - tr(PFPPF^TP)PFP = 0 \quad (7)$$
$$\Leftrightarrow \quad 2FPPF^TPPF - tr(PFPPF^TP)F = 0 \quad (8)$$
$$\Leftrightarrow \quad 2FP^2F^TP^2F - tr(FP^2F^TP^2)F = 0. \quad (9)$$

Notice that f^{-1} only appears in even powers in the above set of polynomial equations and hence one can set $p = f^{-2}$. This is a set of nine fifth order equations in (l_1, l_2, p).

The ten equations, (2) and (9), can be written $AX = 0$, where A is a 10×33 matrix of scalars and X is a vector of monomials

$$\begin{aligned} X=[&l_1^3p^2, l_1^3p^1, l_1^3, l_1^2l_2^1p^2, l_1^1l_2^2p^2, l_2^3p^2, l_1^2l_2^1p^1, \\ &l_1^1l_2^2p^1, l_2^3p^1, l_1^2l_2^1, l_1^1l_2^1p^2, l_1^2p^2, l_2^2p^2, l_1^1p^3, \\ &l_1^1l_2^2, l_2^3, l_1^2p^1, l_2^1p^3, l_1^1l_2^1p^1, l_2^2p^1, l_1^1p^2, l_2^1p^2, \\ &p^3, l_1^2, l_1^1l_2^1, l_2^2, l_1^1p^1, l_2^1p^1, p^2, l_1^1, l_2^1, p^1, 1]^T. \end{aligned} \quad (10)$$

This ordering of the monomials is not a monomial order but it is quite close to the GrevLex monomial order. The reason for not ordering the monomials in GrevLex is that the computations are easier to implement this way. The computed Gröbner basis is the same as we would get using GrevLex. The elimination order is still the one belonging to the GrevLex order. In Figure 2 the order is shown in a way that may be easier to read. The shifts used to perform multiplication by p are also shown.

From this point on all polynomials will be represented by rows in $n \times 33$ matrices. Addition of polynomials is now addition of rows. Multiplying a polynomial with a scalar α corresponds to multiplying the corresponding row with α. Multiplying a polynomial with the monomial p is implemented by shifting elements according to Figure 2. If any non-zero number is in a position that is not shifted it means that multiplication with p was impossible for this vector within this representation. By writing pM_i we mean the polynomial represented by row i in M multiplied by the monomial p. Row indexing starts at 1.

The rows representing $p\det(F)$ and $p^2\det(F)$ are added to the matrix A , or in matrix formulation pM_{10} and p^2M_{10}. This system of 12 rows is seen in Figure 3. This system is reduced using Gauss-Jordan elimination and the rows $pM_{(7,8,9,10)}$ are added. This new system is seen in Figure 4. Again, the system is reduced and the rows $pM_{(8,9)}$ are added. This new system is seen in Figure 5. The system is reduced one last time, and the system now represents a Gröbner basis. This system is seen in Figure 6.

Given the Gröbner basis from the previous step it is now possible to compute the action matrix m_{l_2}, see [2], for multiplication by l_2 by taking the last 15 columns from rows 8,9,11,13,18,10,15 and 16 and changing the sign and then putting ones in selected places, as described in Figure 7.

Dividing each of the 15 eigen-vectors of the transposed action matrix (we are interested in the right eigen-space) by its last element and then selecting elements 12,13 and 14 gives the solutions for (l_1, l_2, p). As there are 15 eigen-vectors there are 15 solutions. The fundamental matrices are then given by $F = F_0 + F_1l_1 + F_2l_2$, as $f^{-2} = p$ is known, the essential matrix can be computed and from the essential matrix the motion (R, t) can be computed.

4 Numerical Precision of the Solver

Figure 8 shows the behavior for random image points. There are two critical configurations [14, 6] for determining f from the fundamental matrix F:

- The main axes of the cameras intersect and both cameras have the same distance to the intersection point. Figure 10 shows that even if it is impossible to estimate f it is still possible to estimate F. Figure 9 shows that if the distances from the cameras to the intersection point are not equal then f can be estimated.
- The main axes of the cameras are parallel. Figure 11 shows that even if f is not recovered it is possible to compute F.

It also important to note that planar scenes are degenerate. This is shown in Figure 12.

5 Stability of the Solution Compared to Other Methods

The solver will be compared with other solvers of interest:

- 5 point method as described in [8]. This solver gives 10 solutions.
- 6 point calibrated method by Pizarro described in [11]. This solver gives 6 solutions.
- 7 point method. No assumption on calibration. This gives 3 solutions.

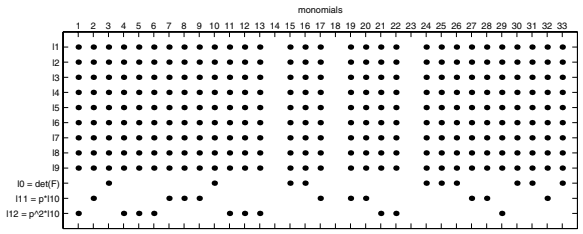

Figure 3: 9 equations from Equation 3 and $\det(F)$, $p\det(F)$ and $p^2\det(F)$.

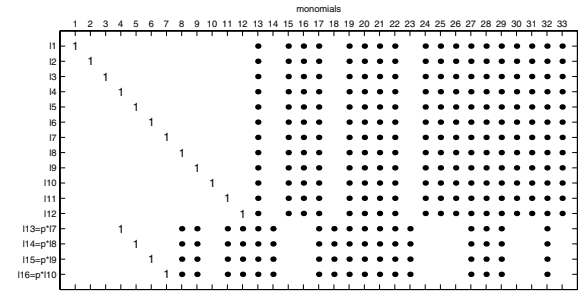

Figure 4: The previous system after a Gauss-Jordan step and adding new equations based on multiples of the previous equations.

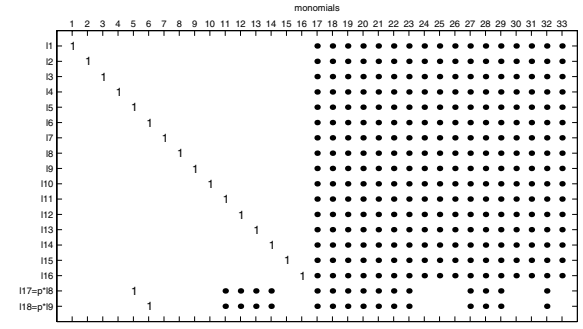

Figure 5: The previous system after a Gauss-Jordan step and adding new equations based on multiples of the previous equations.

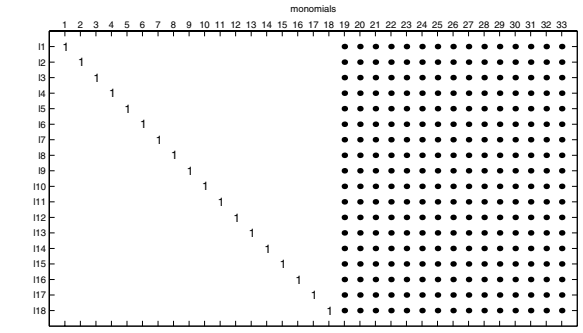

Figure 6: Gauss-Jordan eliminated version of the previous system. This set of equations is a Gröbner basis.

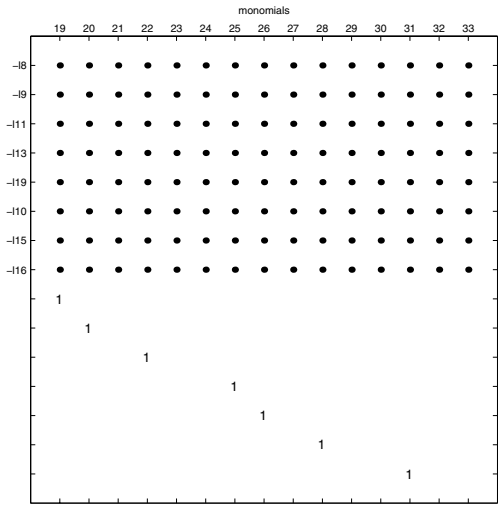

Figure 7: Transpose of the Action-Matrix. Rows 1 to 8 are fetched from the previous matrix and rows 9 to 15 each contains a 1 and the other elements are 0.

A synthetic scene was set up consisting of

- Camera 1 at the origin pointing along the z-axis.
- Camera 2 in a random point 1 unit away from camera 1 and rotated with a random rotation with a mean angle of 5.5^o.
- Scene-points, random points centered on the point $\begin{bmatrix}0 & 0 & 3\end{bmatrix}$ with a Gaussian distribution with standard deviation 1 along each of the axes.

The added noise was calculated for a 1000×1000 camera with a field of view of 40^o and has a normal distribution with standard deviation in pixels given on the horizontal axes of the plots.

The error in the fundamental matrix is computed as $\min_i \|F \pm \tilde{F}_i\|$ where F is the true fundamental matrix and $\{\tilde{F}_i\}$ are the estimated fundamental matrices, all being normalized with Frobenius norm 1.

With the correct value for the focal length f our method beats the 7-point method and the 6-point method of Pizarro but the five-point method is the best, see Figure 13. hen constructing test scenes with focal length $f \neq 1$, the five-point-method and the method of Pizarro run into trouble as they assume that $f = 1$. This is shown in Figures 14 and 15.

There are generally 15 solutions to the problem but some of these solutions can be complex or lead to complex f. In Figure 16 are shown the number of real solutions to the eigen-value problem. Figure 16 shows the number of solutions that have a positive f^2 , which means that it will be possible to compute a positive and real focal length.

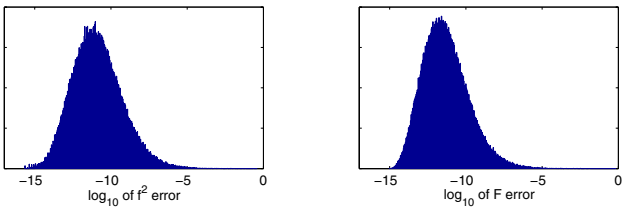

Figure 8: $\log_{10}$ of error in determining f^2 and F for perfect data.

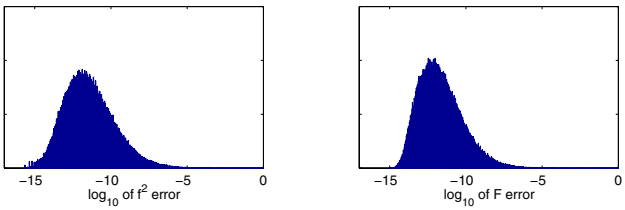

Figure 9: $\log_{10}$ of error in determining f^2 and F for perfect data when the axis intersect.

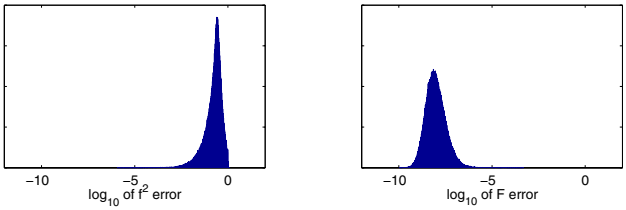

Figure 10: $\log_{10}$ of error in determining f^2 and F. The main axes of the cameras intersect and they have the same distance from principal point to intersection point (some noise added for stability).

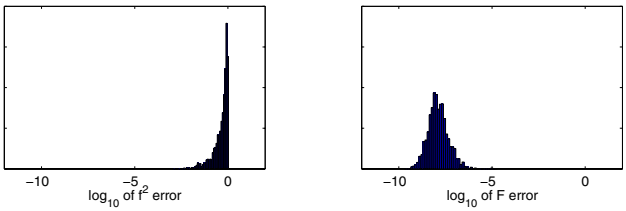

Figure 11: $\log_{10}$ of error in determining f^2 and F. The main axes of the cameras are parallel (some noise added for stability).

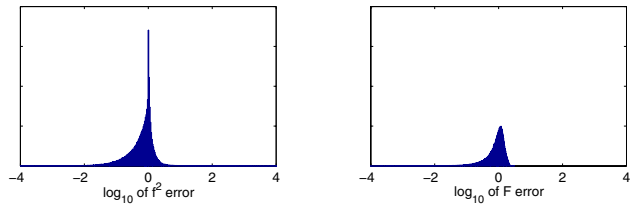

Figure 12: $\log_{10}$ of error in determining f^2 and F. All points on a plane.

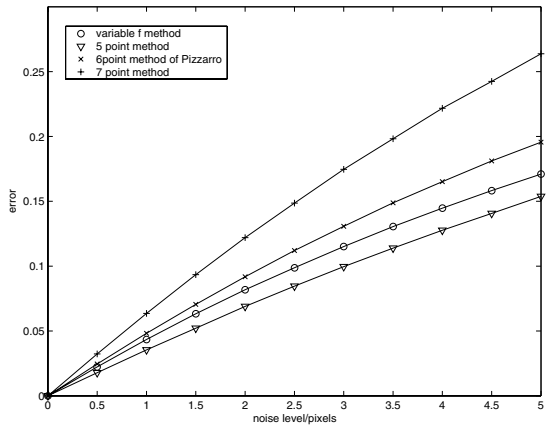

Figure 13: $\log_{10}$ of error in determining the fundamental matrix. Assuming correct calibration. Noise in pixels.

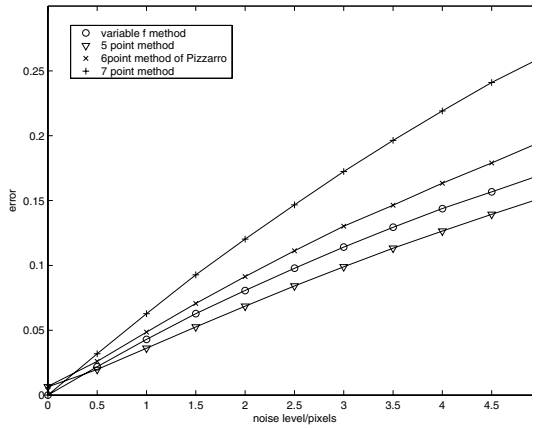

Figure 14: $\log_{10}$ of error in determining the fundamental matrix. Data was generated for $f = 1.01$ but the solutions were computed assuming $f = 1$. Note that for noise smaller than 0.5 our method is better than the five-point.

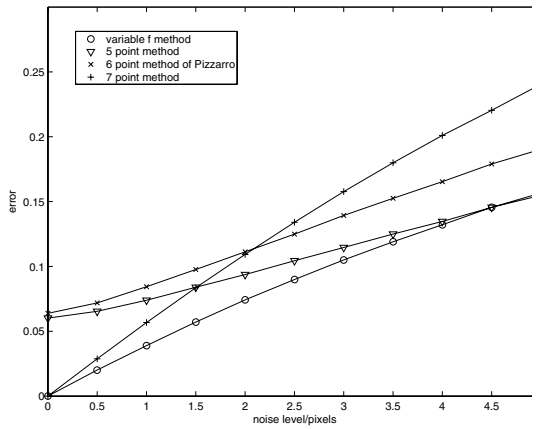

Figure 15: $\log_{10}$ of error in determining the fundamental matrix. Data was generated for $f = 1.1$ but the solutions were computed assuming $f = 1$.

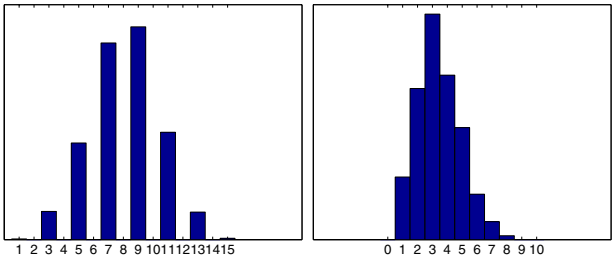

Figure 16: Left: The number of real solutions for f^2. Right:The number of positive f^2 leading to real f.

6 Implementation and Speed

Apart from computing the solutions to the eigen-problem, all the steps are in closed form and can be heavily optimized. For a fully optimized version the eigen-problem is likely to be the bottleneck.

Our current implementation -which is not optimized for speed - runs at 1000 Hz. One implementation is available from [16].

7 Conclusions

We have presented a solution to the minimal problem of six points in two views with unknown focal length. Using Gröbner basis techniques we show how the problem can be solved in an efficient manner, making it an attractive solution for semi-calibrated structure and motion computations. We have also shown that the method gives surprisingly good stability under noise, competitive even with the five point method and more stable than the seven point method.

Acknowledgements

Research support from the Swedish Research Council and U.C. MICRO Program is gratefully acknowledged.

References

[1] D. Cox, J. Little and D. O'Shea, *Ideals, Varieties, and Algorithms*, ISBN 0-387-94680-2, Springer-Verlag, 1997.

[2] D. Cox, J. Little and D. O'Shea, *Using Algebraic Geometry*, ISBN 0-387-98492-5, Springer-Verlag, 1998.

[3] M. Demazure, *Sur Deux Problemes de Reconstruction*, Technical Report No 882, INRIA, Rocquencourt, France, 1988.

[4] O. Faugeras and S. Maybank, Motion from Point Matches: Multiplicity of Solutions, *International Journal of Computer Vision*, 4(3):225-246, 1990.

[5] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, ISBN 0-521-62304-9, Cambridge University Press, 2000.

[6] F. Kahl and B. Triggs, Critical Motions in Euclidean Structure from Motion, *IEEE International Conference on Computer Vision and Pattern Recognition*, Volume 2, pp. 366–372, 1999.

[7] E. Kruppa, Zur Ermittlung eines Objektes aus zwei Perspektiven mit Innerer Orientierung, *Sitz.-Ber. Akad. Wiss., Wien, Math. Naturw. Kl., Abt. IIa.*, 122:1939-1948, 1913.

[8] D. Nistér, An Efficient Solution to the Five-Point Relative Pose Problem, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(6):756-770, 2004.

[9] D. Nistér and F. Schaffalitzky, What do four points in two calibrated images tell us about the epipoles? *Proceedings of the 8th European Conference on Computer Vision*, Volume 2, pp. 41–57, 2004.

[10] J. Philip, A Non-Iterative Algorithm for Determining all Essential Matrices Corresponding to Five Point Pairs, *Photogrammetric Record*, 15(88):589-599, October 1996.

[11] O. Pizarro, R. Eustice and H. Singh, Relative Pose Estimation for Instrumented, Calibrated Platforms, *VIIth Digital Image Computing: Techniques and Applications*, pp. 601–612, 2003.

[12] P. Stefanovic, Relative Orientation - a New Approach, *I. T. C. Journal*, 3:417-448, 1973.

[13] H. Stewénius, Gröbner Basis Methods for Minimal Problems in Computer Vision, *PhD-thesis, Lund University*, 2005.

[14] P. Sturm, On Focal Length Calibration from Two Views, *IEEE International Conference on Computer Vision and Pattern Recognition*, Volume 2, pp. 145–150, 2001.

[15] P. Torr, A. Fitzgibbon and A. Zisserman, The Problem of Degeneracy in Structure and Motion Recovery from Uncalibrated Image Sequences, *International Journal of Computer Vision*, 32(1):27–44, 1999.

[16] Variable f solver code, `http://www.robots.ox.ac.uk/~fsm/var_f_code/`