A Two–step Robust Approach for Motion Estimation *

Sun Xuening and Minas E. Spetsakis
Department of Computer Science
York University
North York, Ontario, Canada M3J 1P3

Abstract

Estimating 3-D motion parameters of rigid objects from two-frame monocular images is considered in this paper. Our main idea is to regularize the initial problem formulation in oder to obtain stable estimates. In the first step, a linear algorithm is applied to generate an initial guess for the solution. The perturbation bound for the guess is discussed to ensure the stability of the algorithm. In the second step, an iterative method based on Tikhonov regularization approach is used to improve the guess which is close to the true solution in the parameter space. Some theoretical results w.r.t. our algorithm are given without their proofs. The results of experiments are satisfactory with noisy synthetic data. The numerical results also verified the theoretical results.

1 Introduction

Determining the 3–D motion parameters and the locations of rigid objects is a challenging problem in computer vision. This problem is called "motion estimation". So far many algorithms for this problem have been presented. The applications include mobile robot navigation, monitoring of dynamic industrial processes, and tracks of airplanes. Although approaches to motion estimation have been developed for many years, they couldn't meet the requirements of real applications. The main reason for this issue is that the existing algorithms can't deal with the ill–posedness or ill–conditioning of the estimation problem. In other words, in the presence of little noise, the computational results obtained by using these algorithms vary enormously. The worse thing is that the results are inconsistent with the underlying physical processes. Therefore, These algorithms are not robust against noise in image sequences. In this paper, our aim is to investigate robust algorithms based on Tikhonov regularization method. The prior information and stable numerical algorithms are the main considerations in our research. The paper is organized as follows: Section 2 gives a brief statement and initial mathematical formulation of motion estimation. The ill–posedness of this formulation is discussed. In Section 3, a linear estimation approach is presented to get an initial guess for our least–squares estimator, and the perturbation bound for the guess is discussed. In Section 4, a constrained least–squares method and its iteration procedure is proposed to regularize the motion estimation proloem, and related theoretical results are presented without giving their proofs because of the limited space of this paper. Several experimental results and conclusions are given in Section 5.

*This work is supported in part by the York Schalorship and Graduate Assistantship.

2 Statement of the problem

The basic geometry of motion estimation is sketched in Fig.1. The object–space coordinate system is denoted by o-xyz. Let the image plane be located at $z = 1$. o is the focus point. A point p in 3-D is represented by its position vector $(x, y, z)^t$. Its position on the image plane is $p/z = (x/z, y/z, 1)^t = (X, Y, 1)^t$ We now consider a particular point on the surface of a rigid object in the scene. Let $p = (x, y, z)^t$ be the coordinates of the point at time t_1, $p' = (x', y', z')^t$ be the coordinates of the point at time t_2. Correspondingly, the point p projects to $P = (X, Y)^t$ and the point p' projects to $P' = (X', Y')^t$ on the image plane.

As we know, the following relation then holds in terms of Kinematics:

$$(x', y', z')^t = R_0(x, y, z)^t + T_0, \quad (1)$$

Where R_0 is a 3×3 orthonormal matrix, i.e., $R_0^t \cdot R_0 = I$. T_0 is a translation vector. Taking any vector T that is collinear with T_0 and taking cross product with both sides of previous equation, we obtain

$$\frac{z'}{z} T \times (X', Y', 1)^t = T \times [R_0(X, Y, 1)^t]. \quad (2)$$

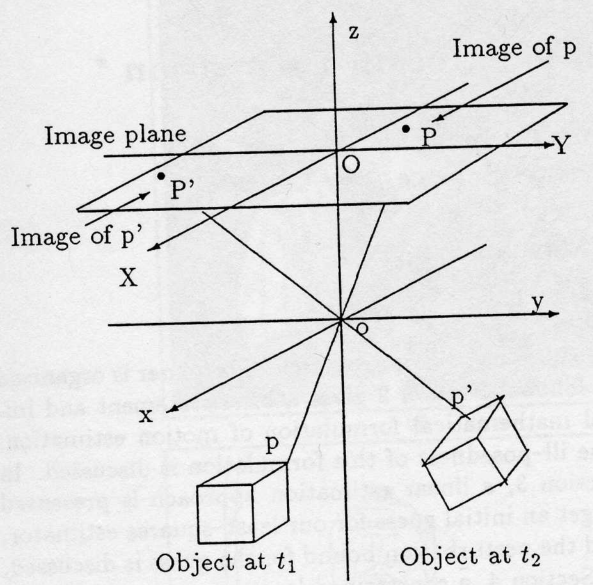

Fig. 1. The basic geometry of motion estimation.

Then, taking dot product of both sides of Eq.(2) with vector $(X', Y', 1)$,

$$(X', Y', 1)E(X, Y, 1)^t = 0. \qquad (3)$$

Where $E \equiv T_s \cdot R_0^t$, or $E \equiv T \times R_0 = [T \times R_1^t, T \times R_2^t, T \times R_3^t]$; R_1, R_2, R_3 being the rows of R_0. $T = (t_1, t_2, t_3)^t$. Let

$$R_0 = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix}, \qquad T_s = \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix}.$$

After a simple manipulation, we get

$$Ax = 0 \qquad (4)$$

Where

$$A = \begin{bmatrix} X_1'X_1, & X_1'Y_1, & X_1', & Y_1'X_1, & Y_1'Y_1, & Y_1', & X_1, & Y_1, & 1 \\ \vdots & & & & & & & & \vdots \\ X_n'X_n, & X_n'Y_n, & X_n', & Y_n'X_n, & Y_n'Y_n, & Y_n', & X_n, & Y_n, & 1 \end{bmatrix}.$$

$n \geq 8$.

$$x = (e_1, e_2, \cdots e_9)^t.$$

In order to reduce dimensions of the parameter space for our optimization problem, we introduce the Rodrigues parameters b_1, b_2, b_3 to represent the rotation matrix. That is

$$R = \frac{1}{\Delta}\begin{bmatrix} 1 + b_1^2 - b_2^2 - b_3^2, & 2(b_1b_2 - b_3), & 2(b_1b_3 + b_2) \\ 2(b_2b_1 + b_3), & 1 - b_1^2 + b_2^2 - b_3^2, & 2(b_2b_3 - b_1) \\ 2(b_3b_1 - b_2), & 2(b_3b_2 + b_1), & 1 - b_1^2 - b_2^2 + b_3^2 \end{bmatrix}.$$

Where $\Delta = 1 + \sum b_i^2$. The vector (b_1, b_2, b_3) gives the direction of the rotation axis and the rotation angle ϕ satisfies $tan(\frac{1}{2}\phi) = (\sum b_i^2)^{\frac{1}{2}}$.

By using three parameters (b_1, b_2, b_3) to describe the rotation and the other three parameters (t_1, t_2, t_3) to describe the translation, we form the unknown estimate vector $\zeta = (b_1, b_2, b_3, t_1, t_2, t_3)^t$ and try to compute them from (4).

Eq.(4) is a linear equation with nine unknowns that are the elements of E. If we take at least eight such equations, we could almost recover the motion parameters [3]. To increase the stability of the solution, we could take more than eight points and do least squares to minimize a quadratic of the form:

$$\min_{x \in D(A)} \{x^t(A^tA)x\} \qquad (5)$$

Where x is nine dimentional vector in which each element is an element of E and A is a matrix that depends on the various pairs of data points $\{P_i, P_i'\}$.

Unfortunately, we can't solve (4) or (5) in a simple and direct way. In [16], the theoretical analysis on some special situations shows that the objective function (5) is ill–posed in the sense of instability of the solutions. Our experimental results also show that the matrix (A^tA) could be ill–conditioned or degenerated , and the computation for x could fail by only using (5) in many cases. Therefore, new mathematical formulations are strongly needed to ensure stable estimates. Now we concentrate on finding new formulations and robust algorithms for the problem (5).

3 The linear method for an initial guess

In this section, we assume that E is approximately decomposable into R and T. Since we don't expect to obtain a precise solution for the problem in this stage, the approximation is acceptable. Without loss of generality, we restrict x to $||x|| = 1$. From (5), we set up the following minimization problem to compute E:

$$F = (Ax)^t(Ax) + \lambda \cdot (1 - x^tx), \qquad (6)$$

Where A is a coefficient matrix. λ is the Lagrange multiplier. By taking $dF/dx = 0$ and $x^tx = 1$, it is easy to see that when the solution x is the eigenvector corresponding to the smallest eigenvalue of (A^tA), F will achieve its minimun. If the motion of a rigid object on the time interval $[t_1, t_2]$ is relatively small, the matrix A could be ill-conditioned (i.e. the condition number of A could be very, very large). In this case, it's difficult to generate a good guess by using Eq.(6), because the corresponding smallest eigenvalue and the sub-smallest eigenvalue are very close, almost unseperatable. The analysis in [16] implies that eigenvectors w.r.t. poorly seperated eigenvalues are very ill-conditioned because of the ill-conditioning of (A^tA). The other possible approach to this problem is the perturbation method, i.e. we could compute the eigenvectors of $(A+\delta A)^t(A+\delta A)$ instead, if $||\delta A||$ is reasonably small. Let $\tilde{A} = A + \delta A$, then we could obtain an approximate solution by solving the following minimization problem

$$\mathcal{F} = (\tilde{A}x)^t(\tilde{A}x) + \lambda \cdot (1 - x^t x),$$

Where $\tilde{A}$ is a perturbed matrix and the condition number of $\tilde{A}$ is relatively small. It is easy to show analytically that under some circumstances, the previous approximation is reasonable practically. So we could introduce small perturbation in the matrix (A^tA) to deal with the ill-conditioning of it.

Now we come to the question of computing $\zeta = (b_1, b_2, b_3, t_1, t_2, t_3)^t$ from E that has been known by now. Let

$$T = (t_1, t_2, t_3)^t.$$

$$R = \begin{bmatrix} R_1 \\ R_2 \\ R_3 \end{bmatrix}, \quad R' = \begin{bmatrix} R_1' \\ R_2' \\ R_3' \end{bmatrix}, \quad E = \begin{bmatrix} E_1 \\ E_2 \\ E_3 \end{bmatrix}.$$

By manipulating $E = T \times R$ or $E = (-T) \times R'$, T and R can be determined by solving the following equations:

$$E_1 = t_2R_3 - t_3R_2, \; E_2 = t_3R_1 - t_1R_3, \; E_3 = t_1R_2 - t_2R_1 \quad (7)$$

And

$$\begin{aligned} ||E_1||^2 &= t_2^2 + t_3^2, & (E_1, E_2) &= t_1 t_2, \\ ||E_2||^2 &= t_3^2 + t_1^2, & (E_2, E_3) &= t_2 t_3, \\ ||E_3||^2 &= t_1^2 + t_3^2, & (E_3, E_1) &= t_3 t_1, \end{aligned} \quad (8)$$

Actually, we can calculate the solution of T from the following equations:

$$\begin{aligned} ||E_2||^2 + ||E_3||^2 - ||E_1||^2 &= 2t_1^2, \\ ||E_3||^2 + ||E_1||^2 - ||E_2||^2 &= 2t_2^2, \\ ||E_1||^2 + ||E_2||^2 - ||E_3||^2 &= 2t_3^2, \end{aligned} \quad (9)$$

When T are determined , R could be computed by means of Eq.(7). As matter of fact, a simple manipulation leads to

$$\begin{aligned} E_1 \times E_2 &= t_3(t_1R_1 + t_2R_2 + t_3R_3), \\ E_2 \times E_3 &= t_1(t_1R_1 + t_2R_2 + t_3R_3), \\ E_3 \times E_1 &= t_2(t_1R_1 + t_2R_2 + t_3R_3), \end{aligned} \quad (10)$$

In the presence of noise in the image sequence, the above equations are not valid any more. In this case, we consider forming a constrained least squares to decrease the impact of noise. Let $(E_1 \times E_2) = (d_1, d_2, d_3)$, $(E_2 \times E_3) = (d_4, d_5, d_6)$, $(E_3 \times E_1) = (d_7, d_8, d_9)$. From Eq.(11) , we can get the following matrix equation:

$$U \cdot R = V, \quad (11)$$

Where

$$U = \begin{bmatrix} t_1t_3 & t_2t_3 & t_3^2 \\ t_1^2 & t_1t_2 & t_1t_3 \\ t_1t_2 & t_2^2 & t_2t_3 \end{bmatrix}, \quad V = \begin{bmatrix} d_1 & d_2 & d_3 \\ d_4 & d_5 & d_6 \\ d_7 & d_8 & d_9 \end{bmatrix}.$$

In order to eliminate the impact of noise, we form the following minmization problem:

$$F = tr((V - U\cdot R)^t(V - U\cdot R)) + tr(\Omega(R^tR - I)), \quad (12)$$

Where $tr(\cdot)$ is the trace of a matrix. Ω denotes Lagrange multiplier matrix (a symmetric matrix) After taking derivative of F w.r.t. the matrix R, we get the following closed-form solution:

$$R = (U^t \cdot VV^t \cdot U)^{-\frac{1}{2}} \cdot U^tV. \quad (13)$$

In the same way, we could establish an equation for R'. The criterion for choosing R or R' is simple one: if the rotation angle $\phi < \pi/2$, then r_1, r_5, r_9 should have the same sign. The previous equation is easily computed by using eigenvalue and eigenvector method which is quite robust against noise in matrices U, V. We have ensured that when $(X_i, Y_i), U, V$ are perturbed to $(X_i + \Delta X_i, Y_i + \Delta Y_i), (U + \Delta U), (V + \Delta V)$, respectively, the solution R of (13) will be perturbed to $(R + \Delta R)$, if $||\Delta X_i||, ||\Delta Y_i||, ||\Delta U||, ||\Delta V||$ are small, then $||\Delta R||$ is quite small.

Now we give a simple and effective method to solve Eq. (13). Let $B = U^tVV^tU$ which is symmetric positive definite. By the similarity transformation, we have

$$P^{-1}BP = \Lambda = diag\{\lambda_1, \cdots, \lambda_n\},$$

Where $\lambda_1, \cdots, \lambda_n$ are the eigenvalues of B. P is a matrix which consists of all the eigenvectors of B. Then

$$B^{-\frac{1}{2}} = P\Lambda^{-\frac{1}{2}}P^{-1},$$

$$B^{-\frac{1}{2}} = P \begin{bmatrix} \lambda_1^{-\frac{1}{2}} & & \\ & \ddots & \\ & & \lambda_n^{-\frac{1}{2}} \end{bmatrix} P^{-1}.$$

Hence

$$R = P \begin{bmatrix} \lambda_1^{-\frac{1}{2}} & & \\ & \ddots & \\ & & \lambda_n^{-\frac{1}{2}} \end{bmatrix} P^{-1} \cdot U^tV.$$

Then it is easy to solve for (b_1, b_2, b_3) from R.

4 Constrained least squares estimator

In the previous section, we have got an initial guess for the parameter vector ζ. ζ^* is denoted to be the guess. Although it is not the precise solution which we want, through the analysis, we are quite sure that the exact solution is not far from the guess we have obtained. We could expect that the true solution lies in a closed sphere $\Omega = \{\zeta : ||\zeta - \zeta^*|| \leq C < \infty\}$.

As we mentioned before, to solve the ill–posed inverse problems like motion estimation effectively, two things are strongly needed: one is inherent constraints or prior information for the underlying physical process; the other is stable numerical methods. At this point, we know that the potential estimates should be within a closed sphere in the parameter space. This fact can serve as prior information. Moreover, we know that the rotation matrix $R(\zeta)$ has to satisfies the orthogonality constraint $R^tR = I$. This constraint can be used as the other prior information. To put these things together, we rewrite the optimization problem as

$$\min\{(Ax(\zeta))^t \cdot (Ax(\zeta))\}, \quad (14)$$
$$s.t.\ ||\zeta - \zeta^*|| \leq C,$$
$$R^t(\zeta)R(\zeta) = I$$

In order to solve the previous problem systematically, we need a proper mathematical method to reformulate the problem and to ensure the stable solutions. As we know, Tikhonov regularization method seems to be the effective one and is our proper choice. In the framework of the regularization approach, we form the following objective function

$$\min_{\zeta\in D(A)} \{||Ax||^2 + ||\Omega_1(R^tR - I)||^2 + ||\Omega_2(\zeta - \zeta^*)||^2\} \quad (15)$$

Where $x = x(\zeta), R = R(\zeta)R(\zeta)$. Ω_1, Ω_2 are regularization parameter matrices. They reflect the impact of the constraints on the solution. The second penalty term in Eq.(18) represents the orthogonality constraint, i.e. $R^tR = I$. The third penalty term represents the constraint that forces the precise solution be within a closed convex set. The previous function integrates all the information we have obtained. Obviously, it is a convex function on a closed convex set regarding our motion estimation problem. The existence of the optimal solution for (15) is guaranteed by our analytical analysis. Now we give the following procedure:

Iteration procedure:

1. Give $\zeta^*, \Omega_1, \Omega_2, \delta_F$ and $n = 1$. Compute the coefficient matrix A from the data pairs $\{P_i, P_i'\}$, $(i = 1, 2, \cdots N)$. Let $\zeta_1 = \zeta^*$.
2. Apply the Levenberg–Marquardt method to solve

$$F(\zeta_{n+1}) = \min_{\zeta_{n+1}\in D(A)} \{||Ax(\zeta_{n+1})||^2 +$$
$$||\Omega_1(R^t(\zeta_{n+1})R(\zeta_{n+1}) - I)||^2 + ||\Omega_2(\zeta_{n+1} - \zeta_n)||^2$$

3. If $||F(\zeta_{n+1}) - F(\zeta_n)|| \leq \delta_F$, then stop; otherwise return to 2.

So far we have built our objective function and the iterative method. To place our algorithm on a solid mathematical ground, we obtain the following theoretical results in the framework of Hilbert space:

- There exists an optimal solution ζ_0 for the problem (15).
- The solution of (15) is robust against noise in image sequences.
- The iteration procedure is convergent and the solution sequence converges to the true value quickly.
- In the presence of noise, the solution sequence for the above iterative procedure is still convergent and depends continuously on the data A.

5 Experimemts and conclusion

Some simulation experiments have been carried out to verify our theoretical results and the effectiveness of our algorithm. A ball that was moving in the half-space $z < 0$ was considered as the object. Some points on the ball were extracted randomly to serve as the feature points.

The conditions of the simulated data are $T_0 = (0.0, 0.0, -1.0)^t$. $(b_1, b_2, b_3) = (0.00, 0.00, -0.414)$. Correspoding rotation matrix is

$$R_0 = \begin{bmatrix} 0.707, & 0.707, & 0.00 \\ -0.707, & 0.707, & 0.00 \\ 0.00, & 0.00, & 1.00 \end{bmatrix}.$$

Experiment 1:

In this experiment, we showed the performance of our algorithm in the presence of noise. As we know, in theory, if noise in the image sequence is relatively small compared with the data, convergence of the solution sequence should be guaranteed. We want to see that it is still true in the numerical computation. We also want to demonstrate that the robustness and accuracy of the estimates are quite good. In this test, 5% of white Gaussian noise was introduced in the simulated data.

We assume that using the linear method, we have obtained the initial guess $g = (0.00, 0.00, -0.71, 0.30, 0.30, -0.70)^t$. Started with this guess, our iterative method gave the following results. They showed that the objective function $\mathcal{F}(\zeta_n)$ decreased monotonically and the solution sequence converged to the true value as the number of iterations increased.

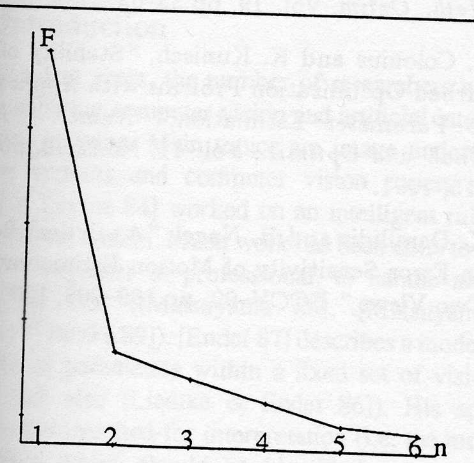

Fig. 2. Plot of the objective function.

We can see that although there are some random factors affecting the trends, the objective function is almost linearly convergent.

Experiment 2:

In this experiment, we show that with different signal–noise ratios and different initial guesses, the algorithm generated a stable solutions for the motion estimation problem. The first group of the results were obtained by using the guess $g = (-0.025, -0.038, -0.354, -0.166, -0.223, -0.612)^t$. The number of iterations was five.

Table 1. The numerical results from the data with 5% perturbation

parameter	estimate	true value	variance
b_1	-0.066	0.00	0.346
b_2	-0.135	0.00	0.342
b_3	-0.408	-0.414	1.438
t_1	-0.183	0.00	0.469
t_2	-0.141	0.00	0.466
t_3	-0.994	-1.00	0.723

Table 2. The numerical results from the data with 10% perturbation

parameter	estimate	true value	variance
b_1	-0.027	0.00	0.226
b_2	-0.066	0.00	0.131
b_3	-0.380	-0.414	0.384
t_1	-0.306	0.00	0.104
t_2	-0.340	0.00	0.102
t_3	-1.015	-1.00	0.113

The second group of the results were obtained by using the guess $g = (0.00, 0.00, -0.241, -0.30, 0.30, -0.70)^t$. The number of iterations was five.

Table 3. The numerical results from the data with 5% perturbation

parameter	estimate	true value	variance
b_1	0.084	0.00	0.532
b_2	-0.050	0.00	0.442
b_3	-0.418	-0.414	1.153
t_1	-0.248	0.00	0.391
t_2	-0.283	0.00	0.375
t_3	-1.002	-1.00	0.688

Table 4. The numerical results from the data with 10% perturbation

parameter	estimate	true value	variance
b_1	-0.149	0.00	0.180
b_2	-0.164	0.00	0.206
b_3	-0.448	-0.414	0.961
t_1	-0.293	0.00	0.215
t_2	-0.294	0.00	0.250
t_3	-0.976	-1.00	0.287

The problem of estimating 3–D motion parameters of objects from two frames of noisy monocular images is discussed from the standpoint of solving nonlinear inverse problems. In the research reported here, we have presented two-step robust approach towards the problem. The main idea behind our algorithm is to introduce more prior information and to apply Tikhonov regularization approach to the parameter estimation problem. The theory presented here ensured the existence of an optimal solution for the minimization problem and guaranteed convergence and stability of our iterative method. The results of simulation experiments demonstrated the feasibility of the method. More statistical analysis is needed in further study. We hope that we can apply this approach to robot vision.

References

1. M. E. Spetsakis and J. Aloimonos, "Optimal Motion Estimation," *Proc. of Workshop on Visual Motion,* pp. 229–237, 1989.

2. M. E. Spetsakis and J. Aloimonos, "Optimal Computing of Structure from Motion Using Point Correspondences in Two Frames," *Proc. of IEEE conference on CVPR,* pp. 449–453, 1988.

3. R. Y. Tsai and T. S. Huang, "Uniqueness and Estimation of 3–D Motion Parameters of Rigid Objects with Curved surfaces," *IEEE Trans. PAMI-6,* pp.13–27, 1984.

4. X. Zhuang, T. S. Huang and R. M. Haralick, "Two-view Motion Analysis: a Unified Algorithm," *J. Opt. Soc. Am. A,* Vol.3, No.9, pp.1492–1500, 1986.

5. J. Weng, T. S. Huang and N. Ahuja, "A Two Step Approach to Optimal Motion and Struture Estimation," *Proc. of IEEE Computer Society Workshop on Computer Vision,* 1987.

6. O. Faugeras, F. Lustman and G. Toscani, " Motion and Strucuture from Point and Line Matches," *Proc. ICCV,* London, England, 1987.

7. T. J. Broida and R. Chellappa, " Performance Bounds for Estimating 3–D Motion Parameters from a Sequence of Noisy Images," *J. Opt. Soc. Am. A,* Vol.6, No.6, pp.879–889, 1989.

8. T. J. Broida, S. Chandrushekhar and R. Chellappa, " Recursive Estimation of 3–D Motion from a Monocular Image Sequence," *IEEE Trans. on AES,* Vol. 26, No.4, pp.639–655, 1990.

9. O. Bottema and B. Roth, *Theoretical Kinematics,* North Holland, New York, 1979.

10. J. H. Wilkinson, *The Algebraic Eigenvalue Problem,* Oxford Science Publications, Oxford, 1965.

11. G. W. Steward, "Perturbation Bounds for the Definite Generalized Eigenvalue Problem," *Linear Algebra and its Applications,* Vol. 23, pp.69–85, 1979.

12. J. J. More, "The Levenberg Marquardt Algorithm: Implementation and Theory," *Lecture Notes in Mathematics 630, Numerical Analysis,* pp.105–116.

13. D. H. Griffel, *Applied Functional Analysis,* Ellis Horwood Ltd, 1981.

14. F. Colonius and K. Kunisch, "Output Least Squares Stability in Elliptic Systems," *Appl. Math. Optim,* Vol. 19, pp.33–63, 1989.

15. F. Colonius and K. Kunisch, "Stability of Perturbed Optimization Prolems with Applications to Parameter Estimation," *Numer. Funct. Anal. and Optimiz.,* Vol. 11, No.9&10, pp.873–915, 1991.

16. K. Daniilidis and H. Nagel, "Analytical Results on Error Sensitivity of Motion Estimation from Two Views," ECCV-90, pp.199–208, 1990.