A GLOBALLY CONVERGENT GAUSS-NEWTON ALGORITHM FOR THE BUNDLE ADJUSTMENT PROBLEM WITH FUNCTIONAL CONSTRAINTS

Niclas Börlin, Per Lindström, Jerry Eriksson
Department of Computing Science, Umeå University, SE-901 87 UMEÅ, Sweden,
niclas.borlin@cs.umu.se

KEY WORDS: Algorithms, Reliability, Bundle Adjustment, Camera Calibration, Mathematics

ABSTRACT
This paper describes a Gauss-Newton-based algorithm for the bundle adjustment problem with functional constraints (GNC). The GNC algorithm has superior theoretical convergence properties compared to the conventional bundle algorithm. Both algorithms were applied to simulated measurements of a sphere with 2–3 cameras and 4–9 points. For 2 cameras and 4–5 points, the GNC converged in substantially more cases. For the other configurations, the convergence properties were similar. The added cost for the GNC algorithm was less than 0.01 iterations on average.
The GNC algorithm need to be evaluated on real-world problems, but the results suggest that the algorithm will be more reliable for minimum data problems and have a minimal overhead for easy problems.

1 INTRODUCTION

Photogrammetry is a science rich in least squares adjustment problems, e.g. bundle adjustment, camera calibration, self-calibration, and template matching. Methods for solving least squares adjustment problems are central to photogrammetry and it is probably a testimony to their success that they are described in appendices rather than in the main text of some textbooks in photogrammetry (Kraus, 1993, Mikhail et al., 2001).

Most methods are based on the Gauss-Markov estimation model that describes how to estimate the parameters of a linear model given observations of it. Non-linear problems are solved as a sequence of linear problems. In each iteration, a correction vector is estimated from the current linear problem. The corrections are added and the process is "repeated until convergence".

However, sometimes the iteration sequence does not converge. In this paper we extend the discussion in (Börlin, 2002) and suggest an algorithm with superior theoretical convergence properties compared to conventional bundle adjustment with functional constraints. The algorithm is not restricted to the bundle adjustment and may be applied to most least squares adjustment problem with non-linear equality constraints, e.g. camera calibration.

2 NON-LINEAR LEAST SQUARES

A weighted non-linear least squares estimation problem with equality constraints may be written as (Gill et al., 1981, Ch. 1 and 4.7) (Björck, 1996, Ch. 9)

$$\min_{\mathbf{x}} f(\mathbf{x}) = \frac{1}{2}\|\mathbf{L}\mathbf{r}(\mathbf{x})\|^2 = \frac{1}{2}\mathbf{r}(\mathbf{x})^{\mathbf{T}}\mathbf{W}\mathbf{r}(\mathbf{x}) \tag{1a}$$

$$\text{subject to } \mathbf{g}(\mathbf{x}) = \mathbf{0}. \tag{1b}$$

The vector $\mathbf{x} \in \Re^n$ contains the n parameters to be estimated. The *objective function* $f(\mathbf{x})$ is a weighted sum of the elements of the *residual function* $\mathbf{r}(\mathbf{x})$ which is the difference between the model and the observations. The matrix $\mathbf{W}$ is a symmetric positive semi-definite (usually positive definite) matrix and $\mathbf{L^T L} = \mathbf{W}$. The constraint function $\mathbf{g}(\mathbf{x})$ is a vector-valued function describing any constraints on the parameters. The residual and constraint functions are assumed to be twice continously differentiable.

2.1 Methods for unconstrained least squares

The classical method for solving the unconstrained problem (1a) is the Gauss-Newton algorithm (Gill et al., 1981, Ch. 4.7.2) (Björck, 1996, Ch. 9.2), where the linearised problem

$$\min_{\mathbf{p}_k} \frac{1}{2} \| \mathbf{L} \left(\mathbf{r}(\mathbf{x}_k) + \mathbf{J}(\mathbf{x}_k)\mathbf{p}_k \right) \|^2 \tag{2}$$

is solved at each iteration k. $\mathbf{J}(\mathbf{x})$ is the Jacobian of the residual function $\mathbf{r}(\mathbf{x})$. Given the solution of Problem (2), the next approximation is given as

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{p}_k. \tag{3}$$

The solution of Problem (2) is the solution of the (weighted) *normal equations*

$$\mathbf{J}_k^\mathbf{T} \mathbf{W} \mathbf{J}_k \mathbf{p} = \mathbf{J}_k^\mathbf{T} \mathbf{W} (-\mathbf{r}_k), \tag{4}$$

where $\mathbf{r}_k = \mathbf{r}(\mathbf{x}_k)$ and $\mathbf{J}_k = \mathbf{J}(\mathbf{x}_k)$.

2.2 Methods for constrained least squares

The Gauss-Newton method may be extended to solve least squares problems with non-linear equality constraints. In this case the constraint function $\mathbf{g}(\mathbf{x})$ is also linearised to form the minimisation problem

$$\min_{\mathbf{p}_k} \frac{1}{2} \| \mathbf{L} \left(\mathbf{r}(\mathbf{x}_k) + \mathbf{J}(\mathbf{x}_k)\mathbf{p}_k \right) \|^2 \tag{5a}$$

$$\text{subject to } \mathbf{g}(\mathbf{x}_k) + \mathbf{K}(\mathbf{x}_k)\mathbf{p}_k = \mathbf{0}, \tag{5b}$$

where $\mathbf{K}(\mathbf{x})$ is the Jacobian of the constraint function $\mathbf{g}(\mathbf{x})$.

The equation system corresponding to Problem (5) is the system equation

$$\begin{bmatrix} -\mathbf{J}_k^\mathbf{T} \mathbf{W} \mathbf{J}_k & \mathbf{K}_k^\mathbf{T} \\ \mathbf{K}_k & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{p}_k \\ \lambda_k \end{bmatrix} = \begin{bmatrix} \mathbf{J}_k^\mathbf{T} \mathbf{W} \mathbf{r}_k \\ -\mathbf{g}_k \end{bmatrix} \tag{6}$$

where λ_k is a vector of Lagrange multipliers of the constraints (5b), $\mathbf{g}_k = \mathbf{g}(\mathbf{x}_k)$, and $\mathbf{K}_k = \mathbf{K}(\mathbf{x}_k)$. Given the solution of Problem (6), the next approximation is calculated as in Equation (3).

2.3 Connection to the bundle adjustment

The formulation (1a) corresponds to the *adjustment of indirect observations* described e.g. in (Cooper and Gordon, 1996, p. 38) and (Mikhail et al., 2001, pp. 394-6) and the solution of Problem (1a) is one half the minimum weighted residual $\mathbf{v^T W v}$ in (Cooper and Gordon, 1996, p. 38) and (Mikhail et al., 2001, p. 394).

Furthermore, Equation (4) is equivalent to Equation (2.46) in (Cooper and Gordon, 1996, p. 38) and Equation (B-17) in (Mikhail et al., 2001, p. 395) and Equation (6) is equivalent to Equation (2.50) in (Cooper and Gordon, 1996, p. 40) and Equation (B-27) in (Mikhail et al., 2001, p. 406). Hence, the bundle adjustment may be interpreted as the Gauss-Newton method applied to the collinearity equations with or without functional constraints.

3 CONVERGENCE

3.1 Local convergence

An optimisation method is said to be *locally convergent* if it converges towards a local minimum given a *close enough* starting point. The Gauss-Newton method for unconstrained problems has been shown to have fast local convergence for problems with small residuals and mild non-linearities (Björck, 1996, Ch. 9.2). For highly non-linear problems or problems with large residuals it may however not be locally convergent, i.e. no matter how close to the minimum we start (except *at* the minimum), the method will not converge!

3.2 Global strategy

In order to guarantee convergence toward a local minimum from a larger region, it is necessary to apply a global strategy on the selection of the iterates. Far from a minimum, i.e. if given a poor starting approximation, the global strategy should act as a safe-guard to avoid divergence. Close to a minimum the global strategy should dampen oscillations. However, it is important that the global strategy does not slow down the algorithm if the linear problems (2) or (5) are good approximations of the original non-linear problem (1).

3.2.1 Line search One of the most straightforward global strategies is the *Goldstein-Armijo line search* (Gill et al., 1981, Ch. 4.3.2.1) (Björck, 1996, Ch. 9.2.1). The line search calculates a *step length* α_k as the first value of the sequence $\{1, \frac{1}{2}, \frac{1}{4}, \ldots\}$ that produces a "sufficiently better" new approximation

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k. \tag{7}$$

In this context, the update vector $\mathbf{p}_k$ is known as a *search direction*.

For the unconstrained case, "sufficiently better" is judged by the objective function and the step length is accepted if

$$f(\mathbf{x}_k + \alpha_k \mathbf{p}_k) \leq f(\mathbf{x}_k) + \mu \alpha_k \nabla_{\mathbf{x}} f(\mathbf{x}_k)^{\mathbf{T}} \mathbf{p}_k, \tag{8}$$

for a given constant $0 < \mu < 1$. The gradient $\nabla_{\mathbf{x}} f(\mathbf{x}_k) = \mathbf{J}_k^{\mathbf{T}} \mathbf{W} \mathbf{r}_k$. In well-behaved cases the full step length $\alpha_k = 1$ will be accepted and the computational overhead will be minimal. Otherwise, additional residual calculations are necessary to find a suitable α_k. However, since condition (8) is based on a linear approximation of problem (1a), failure to satisfy condition (8) for $\alpha_k = 1$ indicates that the vector $\mathbf{p}_k$ is outside the region around $\mathbf{x}_k$ where the approximation (2) is good and taking the full step $\mathbf{x}_k + \mathbf{p}_k$ may produce arbitrarily bad results.

3.2.2 The merit function The constrained case is more complicated since it is necessary to balance reductions of the objective function with violations of the constraints. One solution is to use a quadratic *merit function* (Gill et al., 1981, Ch. 6.1) (Björck, 1996, Ch. 9.4.2)

$$\psi(\mathbf{x}, \nu_k) = f(\mathbf{x}) + \frac{1}{2} \nu_k \|\mathbf{g}(\mathbf{x})\|^2 = f(\mathbf{x}) + \frac{1}{2} \nu_k \mathbf{g}(\mathbf{x})^{\mathbf{T}} \mathbf{g}(\mathbf{x}) \tag{9}$$

to measure if a point $\mathbf{x}_k + \alpha_k \mathbf{p}_k$ is "sufficiently better" than $\mathbf{x}_k$. The parameter ν_k is the penalty for violating the constraints. The search direction $\mathbf{p}_k$ is calculated from Equation (6) and the step length α_k is accepted if

$$\psi(\mathbf{x}_k + \alpha_k \mathbf{p}_k, \nu_k) \leq \psi(\mathbf{x}_k) + \mu \alpha_k \nabla_{\mathbf{x}} \psi(\mathbf{x}_k, \nu_k)^{\mathbf{T}} \mathbf{p}_k, \tag{10}$$

where the gradient $\nabla_{\mathbf{x}} \psi(\mathbf{x}_k, \nu) = \mathbf{J}_k^{\mathbf{T}} \mathbf{W} \mathbf{r}_k + \nu \mathbf{K}_k^{\mathbf{T}} \mathbf{g}_k$. The basic idea is to start with a low penalty parameter ν_k. This will relax the constraints in the early iterations and allow violations of the constraints given modest reductions of the objective function. Later, the penalty parameter is increased and any deviation from the constraints will require higher and higher reduction of the objective function, forcing the iterates to stay closer to the constraint.

3.3 Calculating the penalty parameter

The calculation of the penalty parameters is critical to the performance of the algorithm; raising it too slowly will allow the iterates to stay away from the constraint too long, raising it too fast will force the iterates to follow the constraint to closely. As in the unconstrainted case we do not want the global strategy to disturb the progress if the problem is nice, i.e. the linear approximation (5) is good.

3.3.1 Local update Since the linear least squares model (5) is used to calculate the search direction $\mathbf{p}_k$, a suitable strategy is to approximate the merit function (9) locally around $\mathbf{x}_k$ in the direction of $\mathbf{p}_k$ by a quadratic function consistent with the linear least squares model (5), i.e.

$$\psi(\mathbf{x}_k + \alpha\mathbf{p}_k, \nu_k) \approx \phi(\alpha) = \frac{1}{2}\left(\|\mathbf{L}(\mathbf{r}_k + \mathbf{J}_k\mathbf{p}_k\alpha)\|^2 + \nu_k\|\mathbf{g}_k + \mathbf{K}_k\mathbf{p}_k\alpha\|^2\right), \tag{11}$$

and to choose the penalty parameter ν_k such that $\phi(\alpha)$ has its minimum near $\alpha = 1$.

Since $\phi(\alpha)$ is a quadratic function it takes its minimum for $\alpha = -\phi'(0)/\phi''(0)$, i.e.

$$\alpha_{min}(\nu) = \frac{-\nu\mathbf{g}_k^{\mathbf{T}}\mathbf{K}_k\mathbf{p}_k - \mathbf{r}_k^{\mathbf{T}}\mathbf{W}\mathbf{J}_k\mathbf{p}_k}{\nu\|\mathbf{K}_k\mathbf{p}_k\|^2 + \|\mathbf{L}\mathbf{J}_k\mathbf{p}_k\|^2} = \frac{\nu\|\mathbf{g}_k\|^2 - \mathbf{r}_k^{\mathbf{T}}\mathbf{W}\mathbf{J}_k\mathbf{p}_k}{\nu\|\mathbf{g}_k\|^2 + \|\mathbf{L}\mathbf{J}_k\mathbf{p}_k\|^2}. \tag{12}$$

Solving Equation (12) for ν gives the penalty parameter as a function of the local quadratic approximation of the merit function

$$\nu_{quad}(\alpha) = \frac{-\mathbf{r}_k^{\mathbf{T}}\mathbf{W}\mathbf{J}_k\mathbf{p}_k - \alpha\|\mathbf{L}\mathbf{J}_k\mathbf{p}_k\|^2}{(\alpha - 1)\|\mathbf{g}_k\|^2}. \tag{13}$$

Equation (13) cannot be used for $\alpha = 1$. However, if $\alpha_{min}(\nu_{k-1})$ is close to 1, there is probably no need to recalculate ν. Thus, the following formula is used to suggest an updated penalty parameter based on local information around the point $\mathbf{x}_k$:

$$\nu_{local}(\nu_{k-1}) = \begin{cases} \nu_{k-1} & \text{if } 1 - \delta_1 \le \alpha_{min}(\nu_{k-1}) \le 1 + \delta_2 \\ \nu_{quad}(1 - \delta_1) & \text{if } \alpha_{min}(\nu_{k-1}) < 1 - \delta_1 \\ \nu_{quad}(1 + \delta_2) & \text{if } \alpha_{min}(\nu_{k-1}) > 1 + \delta_2 \end{cases}, \tag{14}$$

for $0 < \delta_i < 1,\ i = 1, 2$.

3.3.2 Global update Given ν_{k-1} and $\nu_{local}(\nu_{k-1})$, the new penalty parameter is conventionally taken as $\nu_k = \max(\nu_{k-1}, \nu_{local}(\nu_{k-1}))$, i.e. forming a non-decreasing sequence.

However, convergence can be unnecessarily slow if an occasional large penalty parameter ν_j leads to subsequent large penalty parameters ν_k even when $\nu_{local}(\nu_{k-1})$ is not large. However, it is possible to allow a temporary reduction in the penalty parameter without losing any important convergence properties provided that the sequence $\{\nu_k\}_{k=0}^{\infty}$ contains a non-decreasing subsequence (Lindström and Wedin, 1988). This may be accomplished by selecting

$$\nu_k = \max(\nu_{local}(\nu_{k-1}), \eta_r), \tag{15}$$

where η_r is the rth largest penalty parameter seen so far.

3.4 Termination criteria

For most problems, the termination criteria for the algorithm could be

$$\|\mathbf{L}\mathbf{J}_k\mathbf{p}_k\| \le \varepsilon_r\|\mathbf{L}\mathbf{r}_k\| \text{ and } \max|g_i(\mathbf{x}_k)| \le \varepsilon_g \tag{16a}$$

for some constants $\varepsilon_r > 0$, $\varepsilon_g > 0$. Criteria (16a) correspond to that the constraints are satisfied and the residual is close to orthogonal to the intersection of the hyperplanes spanned by the columns of

the jacobians $\mathbf{J}_k$ and $\mathbf{K}_k$. However, for problems with very small residuals, the following termination criteria

$$\|\mathbf{L}\mathbf{J}_k\mathbf{p}_k\| \leq \varepsilon_r(1 + \|\mathbf{L}\mathbf{r}_k\|) \text{ and } \max|g_i(\mathbf{x}_k)| \leq \varepsilon_g \tag{16b}$$

is more suitable. A geometric illustration of the termination critera (16) for the unconstrained case is given in (Börlin, 2002)[1].

3.5 The complete algorithm

In summary, the complete GNC (Gauss-Newton for constrained problems) algorithm is

- Initialize. Select ν_{-1}. Reset top penalty parameter list $\eta_1 = \ldots = \eta_r = \nu_{-1}$. Select x_0.
- Repeat for $k = 0, 1, 2, \ldots, max_{iter}$
  - Calculate $\mathbf{r}_k$, $\mathbf{J}_k$, $\mathbf{g}_k$, and $\mathbf{K}_k$.
  - Calculate the search direction $\mathbf{p}_k$ from Equation (6).
  - Stop if the termination criterion (16) is satisfied.
  - Calculate the penalty parameter ν_k from Equation (15).
  - If $\nu_k > \nu_{k-1}$,
    - Replace η_r by ν_k and sort η_i such that $\eta_1 \geq \ldots \geq \eta_r$.
  - Calculate a step length α_k satisfying (10).
  - Update $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k\mathbf{p}_k$.

In our implementation we use $\delta_1 = \delta_2 = 0.25$, $\mu = 0.25$, $\nu_{-1} = 0.1$, $r = 4$, $\varepsilon_r = 10^{-6}$, $\varepsilon_g = 10^{-8}$.

3.6 Convergence properties

Provided that the level set $\{\mathbf{x} \in \Re^n : f(\mathbf{x}) \leq f(\mathbf{x}_0)\}$ is closed and bounded and that the jacobians $\mathbf{J}_k$ and $\mathbf{K}_k$ all have full rank, the GNC algorithm may be proven to be *globally convergent* (Lindström and Wedin, 1988, Gulliksson et al., 1997) toward a stationary point, i.e. a point where $g(\mathbf{x}_k) = \mathbf{0}$ and $\nabla_{\mathbf{x}} f(\mathbf{x}_k)^{\mathbf{T}}\mathbf{K}(\mathbf{x}_k) = \mathbf{0}$. Except in rare circumstances, such as that the starting point $\mathbf{x}_0$ is a stationary point, the GNC algorithm will indeed converge towards a constrained minimum.

If the level set is unbounded, the iteration sequence may diverge, with or without a global strategy, depending on the problem and the starting approximation (Gill et al., 1981, Ch. 4.3).

4 TEST PROBLEM

As a photogrammetric test problem we considered simulated measurements of part of a spherically shaped dome with radius 15m and center at 25m above ground level, see Figure 1. Cameras were placed at ground level symmetrically in a circle with radius 10m and aimed at the top of the dome. The inner orientation of the cameras were assumed to be known with a camera constant of 50mm.

Symmetrically placed points on a circle on the sphere about 36.7m above ground level were projected into the cameras. The point at the top of the sphere was also projected. The resulting simulated image coordinates were in the range $\pm$12.3mm.

The projection model was the collinearity equations with the rotation matrices parameterized by the Euler angles ω, ϕ, κ. Inner constraints were maintained by fixating the position and orientation of camera 1 and the baseline between cameras 1 and 2. The sphere constraint was formulated as

$$(X_i - S_X)^2 + (Y_i - S_Y)^2 + (Z_i - S_Z)^2 - S_r^2 = 0, \tag{17}$$

for all points (X_i, Y_i, Z_i) with the sphere center (S_X, S_Y, S_Z) and radius S_r as unknowns.

[1] However, the second termination criteria in (Börlin, 2002) contains a typo and should read $\|\mathbf{L}\mathbf{J}_k\mathbf{p}_k\| \leq \varepsilon(1 + \|\mathbf{L}\mathbf{F}_k\|)$.

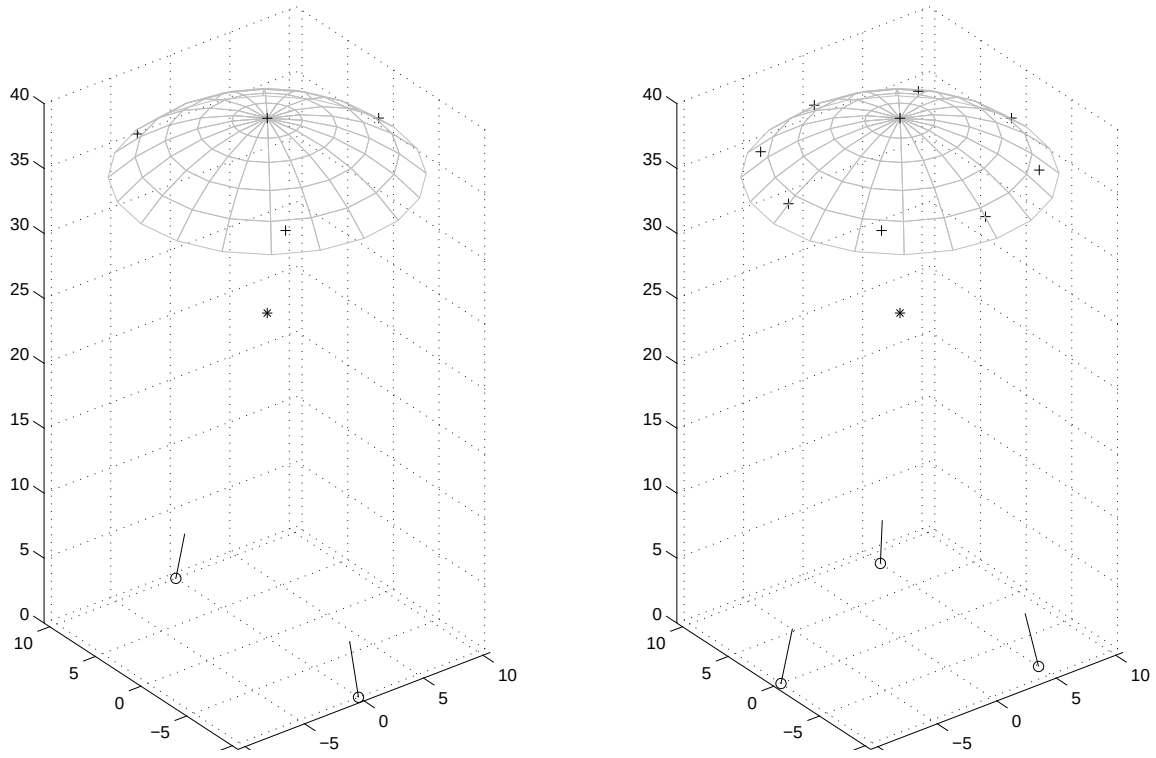

Figure 1: Three-dimensional setup of the simulated problem. The minimum setup (left) has 4 points (+) and 2 cameras (o). The maximum setup (right) has 9 points and 3 cameras. The camera viewing directions and sphere center (*) are also indicated.

4.1 Simulations

Camera networks with 2 and 3 cameras were simulated, varying the total number of points between 4 and 9. For each camera/point combination, 1000 simulations were performed.

In each simulation, white noise with standard deviation of 50 microns was added to the projected points, simulating measurement errors. Camera position and orientation uncertainty were simulated by perturbing the true positions and Euler angles by white noise with a standard deviation of 0.1m in each direction and 5 degrees about each axis, respectively. Initial object point approximations were calculated by triangulation of the measured points and approximated cameras. The sphere parameters were approximated from the initial object points and the linearised equation (17)

$$\min_{S_{X0},S_{Y0},S_{Z0},d} \sum_i \left(2X_iS_{X0} + 2Y_iS_{Y0} + 2Z_iS_{Z0} + d - X_i^2 - Y_i^2 - Z_i^2\right)^2, \tag{18}$$

where $r_0 = \sqrt{d + S_{X0}^2 + S_{Y0}^2 + S_{Z0}^2}$.

The undamped bundle adjustment algorithm and the damped GNC algorithm were given the same measurements and starting approximations. A maximum of 20 iterations were allowed for each algorithm. The number of iterations to convergence or the failure to converge was recorded for each algorithm.

5 RESULTS

The simulation results are shown in Table 1. With 2 cameras and 4–5 points, the GNC has substantially fewer failures, and when both algorithms converge, the GNC is faster in a few more cases. With more points or another camera, both algorithms converge in almost all cases and the Bundle converges faster in a few more cases. In all 11616 cases where both algorithms converge, the average number of iterations for the Bundle and GNC algorithms were 5.433 and 5.431 iterations, respectively.

Table 1: Results of 1000 simulations on each camera/point level. Failure counts for either or both algorithms are shown to the left. Success counts are given by which algorithm was fastest.

		Failure counts			Success counts		
Cameras	Points	Bundle only	GNC only	Both	Bundle faster	GNC faster	Tied
2	4	222	12	85	29	40	612
	5	52	0	2	17	27	902
	6	2	0	1	8	1	988
	7	1	0	0	7	2	990
	8	0	1	0	11	2	986
	9	0	0	0	3	2	995
3	4	2	1	0	6	4	987
	5	0	1	1	12	1	985
	6	1	0	0	13	1	985
	7	0	0	0	12	0	988
	8	0	0	0	16	1	983
	9	0	0	0	25	1	974

6 DISCUSSION

When approaching the photogrammetric community with an optimisation background it was a surprise to the authors that the prevailing algorithms do not use any global strategies, something known in the optimisation community to be necessary for difficult problems. We assumed the reason was that in most cases the data and the starting approximations are of high quality, in which case a locally convergent algorithm is sufficient.

Our results does indicate this to be generally true. However, for the minimal data configurations, the GNC algorithm is substantially more reliable. Furthermore, the overhead of the GNC algorithm compared to the Bundle is minimal.

The objection may be raised that should the bundle adjustment fail to converge, the project planning was poor and should be improved. We concur that an improvement in a numerical routine is by no means a substitute for good planning and carefully implementation of measurements. However, in cases when it is not possible to improve the data, e.g. due to restricted access to the object or poor spatial distribution of markers, we argue that the improvement in reliability introduced by the GNC algorithm will be useful.

REFERENCES

Björck, A., 1996. Numerical methods for least squares problems. SIAM.

Börlin, N., 2002. Improving the robustness of least squares template matching with a line-search algorithm. International Archives of Photogrammetry and Remote Sensing 34(5), pp. 7–11.

Cooper, S. A. R. and Gordon, S., 1996. Theory of close range photogrammetry. In: K. B. Atkinson (ed.), Close Range Photogrammetry and Machine Vision, Whittles, chapter 2, pp. 9–51.

Gill, P. E., Murray, W. and Wright, M. H., 1981. Practical Optimization. Academic Press.

Gulliksson, M., Söderkvist, I. and Wedin, P.-Å., 1997. Algorithms for constrained and weighted nonlinear least squares. SIAM J. OPTIM. 7(1), pp. 208–224.

Kraus, K., 1993. Photogrammetry, Volume 1. Ferd. Dümmler, Bonn.

Lindström, P. and Wedin, P.-Å., 1988. Methods and software for nonlinear least squares problems. Technical Report UMINF-133.87, Inst. of Information Processing, University of Umeå.

Mikhail, E. M., Bethel, J. S. and McGlone, J. C., 2001. Introduction to Modern Photogrammetry. Wiley.