High Breakdown Bundle Adjustment

Anders Eriksson
School of Computer Science
The University of Adelaide
anders.eriksson@adelaide.edu.au

Mats Isaksson
Centre for Intelligent Systems Research
Deakin University
mats.isaksson@gmail.com

Tat-Jun Chin
School of Computer Science
The University of Adelaide
tat-jun.chin@adelaide.edu.au

Abstract

Identifying the parameters of a model such that it best fits an observed set of data points is fundamental to the majority of problems in computer vision. This task is particularly demanding when portions of the data has been corrupted by gross outliers, measurements that are not explained by the assumed distributions. In this paper we present a novel method that uses the Least Quantile of Squares (LQS) estimator, a well known but computationally demanding high-breakdown estimator with several appealing theoretical properties. The proposed method is a meta-algorithm, based on the well established principles of proximal splitting, that allows for the use of LQS estimators while still retaining computational efficiency. Implementing the method is straight-forward as the majority of the resulting sub-problems can be solved using existing standard bundle-adjustment packages. Preliminary experiments on synthetic and real image data demonstrate the impressive practical performance of our method as compared to existing robust estimators used in computer vision.

1. Introduction

Selecting the model which best fits an observed and inevitably noisy data set is one of the fundamental problems in computer vision. It is well known that if the underlying noise is well modelled by a zero-mean Gaussian distribution then the maximum likelihood estimator is found by a least squares approach for which Bundle Adjustment [22] has become an essential tool. However, this formulation is also extremely sensitive to the presence of data points with gross errors, not explained by the normal distribution. Unfortunately, observations in computer vision are very seldom drawn from a single statistical population, as real world images typically contain multiple confounding light sources, objects, surfaces and/or motions, making ordinary least squares methods impractical in most instances. This formulation also suffers from the issue of masking [18], that is, multiple outliers might be present in the data yet remain undetectable by standard least squares diagnostic procedure.

Robust statistics is a departure from standard multivariate parameter estimation, in that it seeks estimates that are invariant to the presence of data containing *outliers*, or points that significantly deviate from the distributions exhibited by the majority of the data. Outliers have turned out to be more likely in high-dimensional datasets with a large number of observations and are often indistinguishable from the rest of the data by visual inspection.

In computer vision this task of robust statistical estimation is typically formulated in one of two closely related pipelines. Either one attempts to identify and remove the outliers followed by a classical parameter estimation on the remaining data, or one tries to find a fit that is robust to corruptions in the sense that it is close to what we would have obtained without the outliers. The second robust fitting approach is usually preferred as it can be argued that the task of finding *some* of the inliers is easier than finding *all* of the outliers.

Examples of the former approach includes the L_∞ methods of [20, 16] in which outliers are identified as part of the support set of a solution of a quasi-convex optimization problem, and consequently removed. Unfortunately, this approach also typically removes a substantial fraction of the inliers in the process making it unsuitable for cases when the number of outliers is high.

Perhaps the best known example of the latter formulation is RANSAC [7], which has proven to be a very powerful method for model fitting in the presence of outliers. It is

978-1-4799-6683-7/15 $31.00 © 2015 IEEE
DOI 10.1109/WACV.2015.48

a non-deterministic algorithm that samples the model space by randomly selecting a large number of minimal subsets of the data, fitting a model to each and retaining only the model that explains the largest portion of the observations. The method has a number of drawbacks, however, including the lack of a guarantee of optimality, its decreasing efficiency as the problem size grows, and the fact that it requires very efficient minimal solvers, thus limiting its applicability.

In general the robustness of statistical parameter estimation methods has received significant attention from a wide variety of fields. However, the theory behind robust statistics was largely developed as late the 1970s, starting with M-estimators (which RANSAC can be seen as an instance of [14]) and continuing with other bounded influence estimators such as R-, L- and GM-estimators [10]. This broad class of estimators is arguably the most popular robust method currently used. This is motivated in part by their simplicity and computational efficiency. The vast majority of M-estimators do, however, suffer from a low breakdown-point as well as a sensitivity to scale.

The *breakdown-point* [13] is a standard measure of the robustness of an estimator and is defined as the smallest fraction of *arbitrarily large*, incorrect observations (outliers) that can have an *arbitrarily large* effect on the final estimate. For ordinary least squares estimators the breakdown-point is 0 since a single outlier can completely corrupt the result. The highest possible breakdown-point is defined as 0.5 since if more than half the data is corrupted it is impossible, in a worst case scenario, to distinguish between the distribution of the data from that of the contaminations. Note that this is not the same as saying that an estimator can not handle more than 50% outliers.

In this paper we will focus on high-breakdown, scale invariant estimators for computer vision applications, that is, estimators with breakdown-points between 0 and 0.5 that do not require the level of inlier noise to be known. One of the perhaps most important such high-breakdown robust estimators is the **Least Quantile of Squares (LQS)** estimator. Defined as

$$\arg\min_{x} \left\{ \phi_{lqs}(x) = r^2_{(k)}(x) \right\}, \qquad (1)$$

where $r^2_{(k)}(x)$ denotes the k-th largest residual. When $k = n/2$ this method is known as **Least Median of Squares (LMS)** [17], and we label the associated minimisation criteria ϕ_{med}. It can be shown that, the breakdown-point of the above estimators can be set as high as 0.5.

Robust estimators having a high breakdown-point tend to be very computationally demanding, and the above formulation is no exception. For instance finding the global minimizer of ϕ_{med} requires an exhaustive search over the entire space of possible estimates generated from the data, a task that is prohibitively expensive even for a moderate number of observations. Even finding local minima to (1) can be challenging since ϕ_{lqs} is non-differentiable hence excluding the application of most standard optimization methods. If in addition the parameter space is constrained the task becomes even more challenging.

Note, however, that there are approaches which attempt to compute such estimators. The *LMedS* algorithm of [17] is a Monte-Carlo technique that randomly draws a large number of minimal subsets of the data, fits a model to each such minimal set and selects from this pool the model that minimizes ϕ_{med}. However, this approach suffers similar limitations to the closely related RANSAC method described above, and inevitably calculates only an approximation to the true result for all but modestly sized problems.

Achieving robustness to outliers in multiview geometry is still very much a relevant and open problem, as evidenced for instance by the work [5]. The main aim of this paper is to contribute to this topic by showing that local solutions of high-breakdown estimators for *general* computer vision problems can be efficiently computed using existing publicly available software packages.

2. Projective Geometry

In this paper we will consider the multiview geometry problem of estimating camera and structure parameters that minimize some aspect of the reprojection error of a number of measured image points.

Let $\pi_{ij}(\theta) = \left[\pi^x_{ij}(\theta) \; \pi^y_{ij}(\theta) \right]$ denote the projection of point i in image j given parameter vector θ and $u_{ij} = \left[u^x_{ij} \; u^y_{ij} \right]$ the observed image location of the same point. The dimension of θ depends entirely on the problem at hand. In the case of triangulation it is simply the coordinate of a single 3D point and for full structure from motion applications θ contains all the camera matrices as well as the locations of all the points in the entire scene.

When the error minimized is the total sum of squares reprojection error

$$\min_{\theta \in \mathcal{Q}} \sum_{ij} r^2(u_i - \pi_{ij}(\theta)) = \min_{\theta \in \mathcal{C}} \sum_{ij} ||u_{ij} - \pi_{ij}(\theta)||^2_2, \qquad (2)$$

then the method of choice is the well known Bundle Adjustment algorithm [22, 6]. Here the structure of the set $\mathcal{Q}$ depends on the setting; for full projective reconstruction $\mathcal{Q}$ is simply $\mathbb{R}^n$ but in the case of euclidean reconstruction it also involves the restriction to rotation matrices.

With

$$\Pi(\theta) = \begin{bmatrix} \pi^x_{11}(\theta) & \pi^y_{11}(\theta) \\ \vdots & \vdots \\ \pi^x_{mn}(\theta) & \pi^y_{mn}(\theta) \end{bmatrix}, \; U = \begin{bmatrix} u^x_{11} & u^y_{11} \\ \vdots & \vdots \\ u^x_{mn} & u^y_{mn} \end{bmatrix}, \qquad (3)$$

we can write (2) compactly as

$$\min_{\theta \in \mathcal{Q}} ||U - \Pi(\theta)||^2_F. \qquad (4)$$

Bundle adjustment method solves (4) iteratively using a nonlinear least squares optimization algorithm. The Levenberg-Marquardt algorithm has proven to be the most successful method, as it is simple to implement, robust to initialization, and its framework makes it very amenable to taking advantage of the forms of sparsity that typically arise in multiview geometry problems. Each step of this algorithm produces an estimate of the parameters that improves upon the previous estimate and the resulting series of estimates can be shown to converge to a local minima of (4).

Bundle adjustment has for some time been an essential part of multiview geometry parameter estimation, and despite its long history within the computer vision community it still receives significant research attention. The focus of these research efforts is primarily towards improving the computational cost involved with applying bundle adjustment to very-large scale reconstruction problems [12, 11, 23, 3]. As a result there exist a large number of publicly available software packages for solving a wide range of least squares problems in computer vision [12, 23, 1].

3. Proximal Splitting Methods

We begin with a brief introduction to proximal splitting methods. The proximity operator $\text{prox}_f : H \rightarrow H$ of a proper, convex and lower semi-continuous function $f : H \rightarrow \mathbb{R}$, with $\rho > 0$ and H a Hilbert space, is defined as

$$\text{prox}_{f/\rho}(y) = \arg\min_{x \in H} \left(f(x) + \frac{\rho}{2}||x - y||^2 \right) \quad (5)$$

This notion was first introduced by [15] as a generalization of the concept of orthogonal projections onto convex sets. The proximity operator plays a central role in a class of convex optimization algorithms known as proximal splitting methods. These schemes are first order optimization methods that are particularly aimed at minimizing a sum of functionals for which it is possible to efficiently compute its proximity operator.

For any $y \in H$, $\bar{x}$ is a stationary point of (5) if and only if the inclusion $0 \in \partial f(\bar{x}) + \rho(\bar{x} - y)$ holds. Or equivalently

$$\bar{x} = (I + \frac{1}{\rho}\partial f)^{-1} y. \quad (6)$$

Next consider a general convex optimization problem on the following, primal

$$(P) \quad \min_{x \in H_1} \underbrace{f(x) + g(Ax)}_{=\Phi(x)}, \quad (7)$$

and dual forms

$$(D) \quad -\min_{y \in H_2} \underbrace{f^*(-A^* y) + g^*(y)}_{=\Phi^*(y)}. \quad (8)$$

where $A : H_1 \rightarrow H_2$ is a bounded linear operator and both f and g are proper, convex and lower semi-continuous functions. Here f^* denotes the Fenchel dual of f. If we further assume that a solution to (7) exists then it follows by Fermat's theorem that a minimizer of (7) and (8) must satisfy

$$0 \in \partial f(\bar{x}) + \partial(g \circ A)(\bar{x}), \quad (9)$$

$$0 \in \partial(f \circ -A^*)(\bar{y}) + \partial g(\bar{y}), \quad (10)$$

respectively. The additive structure of this inclusion can then equivalently be recast as a fix-point iteration in a number of different ways, resulting in a variety of different proximal splitting algorithms. Most notably the *Forward-Backward Splitting* and *Douglas-Rachford Splitting* see [4].

We will however in this paper restrict our focus only on the latter, for which the above (dual) inclusion is expressed as

$$\bar{y} + \frac{1}{\rho}\partial(g^*)\bar{y} \in \left(I + \frac{1}{\rho}\partial(f \circ A^*) \right)^{-1} \left(I + \partial(g^*)\right)\bar{y} + \frac{1}{\rho}\partial g^*(\bar{y}). \quad (11)$$

Resulting in the well-known Douglas-Rachford iteration,

$$z^{k+1} = z^k - y^k + \text{prox}_{(f^* \circ A^*)/\rho}(2y^k - z^k), \quad (12)$$

$$y^{k+1} = \text{prox}_{g^*/\rho}(z^{k+1}), \quad (13)$$

here to the dual problem (8).

It has later been established that there are strong connections between proximal splitting methods and a number of already existing algorithms. The Split Bregman and alternating Split Bregman algorithm [9], the augmented Lagrangian methods [2] and Projected Landweber algorithm, to name a few, can all be viewed as special instances of the classical proximal splitting methods, see [4, 19] for more details.

4. Least Quantile of Squares and Proximal Splitting

Proximal splitting methods has almost exclusively been applied to convex problems. The are a number of likely reasons for this. Firstly, the proximal operators of non-convex functions are necessarily no longer firmly non-expansive. Convergence is in general no longer guaranteed and even if the iterates converge they need not do so to a local minima of the associated minimization problem.

We will in this section derive our formulation of a method for finding a local minimizer of least quantile of squares minimization problems on the form (1). Our method is based on proximal splitting and despite the non-convexity of the optimization problem at hand we will show that proximal splitting methods is particularly suitable to the class of robust estimators discussed in this paper.

As a starting point for this presentation we chose the Douglas-Rachford iteration applied to the dual problem (12)-(13). The motivation for this choice was mainly we found that it led to a more intuitive derivation and made the connection to the proximal operator of the associated conjugate function of the function ϕ_{lqs} perhaps more obvious.

However, as indicated by the discussion in the previous section and the work of [19] choosing any of the other interpretations of this approach will without doubt lead to an identical, or very similar, algorithm. For those readers who have an intimate knowledge of these methods, the resulting formulation will probably appear familiar.

Recall again, our intended problem of finding the least quantile of squares of reprojection errors,

$$\min_{\theta \in \mathcal{Q}} \ \phi_{lqs} \{ \frac{1}{2} ||u_{ij} - \pi_{ij}(\theta)||_2^2 \}. \tag{14}$$

Using the indicator function

$$\iota_{\mathcal{S}}(x) = \begin{cases} 0, & x \notin \mathcal{S}, \\ 1, & x \in \mathcal{S}. \end{cases} \tag{15}$$

and $\mathcal{C} = \left\{ x \in \mathbb{R}^{mn \times 2} \ \middle| \ \exists \theta \in \mathcal{Q} \text{ s.t. } x = U - \Pi(\theta) \right\}$ we can reformulate (14) as

$$\min_{x \in \mathbb{R}^{mn \times 2}} \ \phi_{lqs} \{ \ \frac{1}{2} ||x_i||_2^2 \} + \iota_{\mathcal{C}}(x). \tag{16}$$

with x_i denoting the i-th row of x. If we then let $f = \phi_{lqs} \ \{ \ \frac{1}{2} || \cdot ||_2^2 \}$ and $g = \iota_{\mathcal{C}}$ we obtain a formulation in line with the theory of the previous section. Indeed, the Fenchel conjugates of f and g are both convex functions so the Douglas-Rachford splitting (12)-(13) could in theory be applied directly. However, briefly disregarding issues of duality gaps, a requirement for this approach to be computationally efficient is that any subproblem that arise can be solved with little computational effort. Unfortunately, the conjugates f and g are not always easily evaluated. Instead we propose to only find approximate solutions to the associated subproblems, solutions that can be efficiently computed yet are still sufficiently accurate to ensure a high quality solution.

4.1. Approximate Proximal Operators

Since evaluating proximal mappings associated with the conjugate functions f and g can be as demanding as solving the original problem we instead propose to solve approximate variants of these proximal mappings.

We start by stating the following useful theorem.

Lemma 4.1 *Any solution $\bar{x}$ the problem*

$$\min_{x \in \mathbb{R}^m} \ \underbrace{\max_{i \in \mathcal{K}} \{ \ |h_i(x)| \ \}}_{||h_{\mathcal{K}}(x)||_\infty}, \tag{17}$$

with $h : \mathbb{R}^m \to \mathbb{R}^n$ pseudoconvex, and $\mathcal{K} \subset \{1, ..., n\}$, $|\mathcal{K}| = k$, is a local minimizer of

$$\min_{x \in \mathbb{R}^m} \ \phi_{lqs} \ \{ \ |h_i(x)| \ \}, \quad i \in \{1, ..., n\}, \tag{18}$$

if and only if, $|h_i(\bar{x})| \leq |h_j(\bar{x})|$, $\forall i \in \mathcal{K}, \ j \notin \mathcal{K}$.

Hence, for a fixed choice of $\mathcal{K}$, with $h_i(x) = ||x_i||_2^2$ the conjugate function of $\tilde{f}(x) = ||h_{\mathcal{K}}(x)||_\infty = \left| \left| ||x_{\mathcal{K}}||_2^2 \right| \right|_\infty$, becomes

$$\tilde{f}^*(y) = \sup(y^T x - \tilde{f}(x)) = \sup(y^T x - ||x_{\mathcal{K}}||_{2,\infty}^2) \tag{19}$$

$$= \begin{cases} 0, & ||y_{\mathcal{K}}||_{2,1}^2 \leq 1, \ y_i = 0, i \notin \mathcal{K} \\ \infty, & \text{otherwise} \end{cases} \tag{20}$$

Consequently, letting the proximal mapping of $\tilde{f}^*$ at v approximate the proximal mapping of f.

$$\text{prox}_{f^*/\rho}(v) \approx \text{prox}_{\tilde{f}^*/\rho}(v) = \arg\min_y \ ||y - v||_2^2 \tag{21}$$

$$\text{s.t.} \ ||y_{\mathcal{K}}||_{2,1}^2 \leq \rho \tag{22}$$

$$y_i = 0, i \notin \mathcal{K} \tag{23}$$

This convex problem is equivalent to an orthogonal projection onto the $l_{2,\infty}$-ball, for which very efficient solvers exist, see for instance [21]. In addition, the indices in $\mathcal{K}$ needs to be set at each instance of this mapping. We set $\mathcal{K}$ to the indices of the $n/2$ smallest values of$||v_i||_2$. In later sections we will show that this is indeed a sensible choice choice of $\mathcal{K}$, one that leads to an algorithm which behaves as intended.

To then find an approximate solution to prox_{g^*} we first make use of the identity, $\text{prox}_f(v) = v - \text{prox}_{f^*}(v)$, see [4]. We then let

$$\text{prox}_{g^*}(v) = v - \text{prox}_{g^{**}}(v) \approx v - \text{prox}_g(v). \tag{24}$$

This is indeed an approximation, as $g^{**} = g$ if and only if g is convex and lsc. We can write

$$\text{prox}_{g^*/\rho}(v) \approx v - \arg\min_y \iota_{\mathcal{C}}(y) + \frac{\rho}{2} ||y - v||_2^2 = v - \arg\min_{\theta \in \mathcal{Q}} ||U - v - \Pi(\theta)||_2^2. \tag{25}$$

Comparing the above expression with (4) we note that (25) is on a form to which standard bundle adjustment solvers can be applied.

We are now ready to state our proposed algorithm. With $\tilde{f}$ and g as above, and A the identity mapping, applying (21)-(23) and (25) to (12)-(13) we obtain algorithm 1.

Algorithm 1 Least Median of Squares Bundle Adjustment.

Input: U, θ^0, $y^0 = \Pi(\theta^0)$, z^0, ρ, $k = 0$

repeat

$$z^{k+1} = z^k - y^k + \text{prox}_{\tilde{f}^*/\rho}(2y^k - z^k), \tag{26}$$

$$U^k = U - z^{k+1}, \tag{27}$$

$$\theta^{k+1} = \arg\min_{\theta \in \mathcal{Q}} ||U^k - \Pi(\theta)||_2^2 \tag{28}$$

$$y^{k+1} = z^{k+1} - (U - \Pi(\theta^{k+1})), \tag{29}$$

$$k = k + 1 \tag{30}$$

until convergence

This algorithm is very simple in its design and almost as simple in implementation. Algorithm 1 can be realised by essentially inserting a single line (26) at the end of the bundle adjustment loop in the software package of choice. Hence, as our proposed meta-algorithm rests on top of existing total least-squares bundle adjustment solvers, this framework also offers the possibility of seamless transition between a high-breakdown and the standard least squares formulations.

An interesting, and perhaps counter-intuitive, interpretation of this algorithm is that it shows that by moving the image measurements U around in a certain way we can dupe the standard bundle adjustment algorithms to find a solution to the non-smooth Least Quantile of Squares problem, a very different problem to the smooth total least-squares problem it was designed for.

4.2. Initialization

Initialization is of crucial importance to refinement method such as bundle adjustment. If the starting point is not good enough the final result will inevitably be poor. This is just as true for standard bundle adjustment as it is for the variant proposed here. Despite this we will not discuss initialization methods in detail here. The reason for this is firstly that appropriate initialization methods are plentiful and well-known. Secondly, even though one of the arguments made throughout this paper in favor of the proposed estimators is their high breakdown point, this does not imply that they suitable only when the the number of outliers is high. Our approach should be viewed as a refinement tool, just as standard bundle adjustment is, but with a higher degree of robustness to gross outliers, and discussions on how to properly initialize them should be identical.

If you believe that RANSAC will successfully remove all outliers and provide a sufficiently accurate initial estimate to your problem, then by all means you should use standard bundle adjustment for the subsequent refinement. But there are countless situation where this is not the case and the presence of gross outliers are almost a certainty. RANSAC can fail or initialization is available by other means, for instance odometry information in SLAM applications, geotagging or through a rough semantic descriptions of the scene. It is in instances such as these we believe that the proposed approach could prove useful. In the following section we simply assume that a rough initial estimate is available and that there are outliers present in the data.

4.3. The Parameter ρ

For convex problem, the Douglas-Rachford splitting can be shown to converge for any choice of $\rho > 0$. This does however not necessarily hold in the non-convex setting. However, we have empirically observed that algorithm 1 typically only converges for moderate values of ρ. A standard approach is to modify this penalty parameter during the progress of the algorithm. The simple such scheme used in this work $\rho^{k+1} = \eta\rho^k$, where typical choices of these parameters were $\rho^0 = 1e - 3$ and $\eta = 1.01$.

4.4. Convergence Analysis

Finally a brief mentioning about the convergence of the proposed algorithm. When applied to non-convex problems, proximal splitting method does in general not need to converge, and even when it does, it is not clear what it will converge to. With the exception of the work in [8], which analyses the Forward-backward splitting on non-convex problems, very little seem to have been published regarding the convergence of Douglas-Rachford splitting applied to non-convex problems.

We have however established the following weaker convergence property of algorithm 1.

Theorem 4.1 *If the sequence of iterates produced by algorithm 1 converges, then $U - \Pi(\theta^k)$ will converge to a local minimizer of* (17).

Proof 4.1 *Let $\bar{y}$, $\bar{\theta}$ and $\bar{z}$ denote the limit of the sequences $\{y^k\}$,$\{\theta^k\}$ and $\{z^k\}$ respectively. As such they must be fix-points of* (21)-(23) *and* (25) *and we have*

$$\bar{z} = \bar{z} - \bar{y} + \text{prox}_{\tilde{f}^*/\rho}(2\bar{y} - \bar{z}) \tag{31}$$

$$\bar{\theta} = \arg\min_{\bar{\theta} \in \mathcal{Q}} ||U - \bar{z} - \Pi(\bar{\theta})||_2^2 \tag{32}$$

$$\bar{y} = \bar{z} - (U - \Pi(\bar{\theta})). \tag{33}$$

Simplifying (31) *and invoking the identity* $\text{prox}_{f^*}(v) = v - \text{prox}_f(v)$ *with* $v = 2\bar{y} - \bar{z}$ *we obtain*

$$\bar{y} - \bar{z} = \text{prox}_{\tilde{f}/\rho}(2\bar{y} - \bar{z}) \Leftrightarrow \partial\tilde{f}(\bar{y} - \bar{z}) + \rho(-\bar{y}) = 0 \tag{34}$$

The first order condition for optimality of (32) *is*

$$\nabla\Pi(\bar{\theta})^T(U - \bar{z} - \Pi(\bar{\theta})) = 0 \tag{35}$$

Multiplying the right-hand side of (34) *by* $\nabla\Pi(\bar{\theta})^T$ *from the left, inserting* (33) *gives us*

$$\nabla\Pi(\bar{\theta})^T \partial \tilde{f}(U - \Pi(\bar{\theta})) = 0 \tag{36}$$

It can then trivially be shown that (assuming the standard regularity conditions hold) these are the necessary conditions for a stationary point of (17). *Since* $\bar{\theta}$ *is a local minimizer of* (17) *and by the convexity of* $\tilde{f}$ *it follows that such a point must be a local minima. Note that we have above assumed that* $\mathcal{Q} = \mathbb{R}^m$, *extending the proof to instances where this is not the case would follow along very similar lines. However, due to space restrictions we omit this part of the proof here.*

Lemma 4.2 *If* $||v_i||_2^2 \leq ||v_j||_2^2,\ \forall\ i \in\leq\ j$ *then* $||x_i||_2^2 \leq ||x_j||_2^2,\ \forall\ i \leq j$, *where*

$$x = v - \text{prox}_{\tilde{f}^*/\rho}(v). \tag{37}$$

That is, since we can without loss of generality assume that v is always sorted according to magnitude, $x = v - \text{prox}_{\tilde{f}^*/\rho}(v)$ is an order-preserving mapping, with respect to magnitude.

Corollary 4.3 *If the sequence of iterates produced by algorithm 1 converges, then* $U - \Pi(\theta^k)$ *will converge to a local minimizer of* (1).

Proof 4.3 *By theorem 4.1 we have that* $U - \Pi(\theta^k)$ *converges to a local minimizer of* (17). *From* (32) *we get*

$$U - \Pi(\bar{\theta}) = (\bar{y} - \bar{z}) = (2\bar{y} - \bar{z}) - \text{prox}_{\tilde{f}^*/\rho}(2\bar{y} - \bar{z}). \tag{38}$$

By the above lemma this implies that

$$||U_i - \Pi_i(\cdot)||_2^2 \leq ||U_j - \Pi_j(\cdot)||_2^2,\ \ i \in \mathcal{K},\ j \notin \mathcal{K} \tag{39}$$

Since by construction $|\mathcal{K}| = k$ *and* $||2\bar{y}_i - \bar{z}_i||_2^2 \leq ||2\bar{y}_j - \bar{z}_j||_2^2,\ i \in \mathcal{K},\ j \notin \mathcal{K}$. *Then letting* $h_i(\cdot) = ||U_i - \Pi_i(\cdot)||_2^2$ *and invoking lemma 4.1 the corollary follows.*

Empirically we have found that the proposed method always converges as long as ρ is made sufficiently large. The quality of the local minima obtained through our proposed method will obviously be dependent on the starting point.

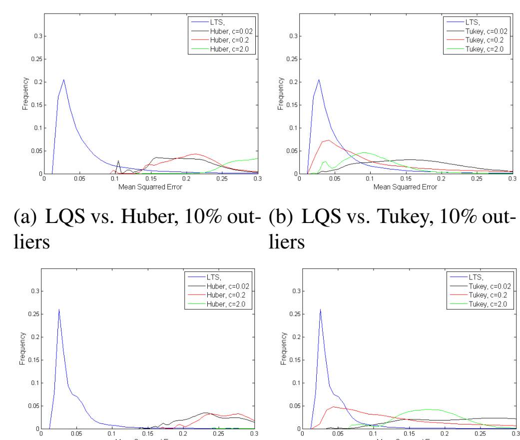
5. Experiments

This section contains an experimental evaluation, as well as a theoretical validation, of our proposed method. We performed experiments on two different datasets, one synthetic and one real. The algorithms were implemented in Matlab and run on standard desktop computers. The implementation of [12] was used as a bundle adjustment solver and the reconstructions carried out in this section are all Euclidean.

5.1. Synthetic Data

In order to evaluate the performance of our proposed method we first tested it on synthetically generated data. We randomly created modest sized instances of 3D scenes containing 5 cameras and 30 3D points. Gaussian noise with a standard deviation of $\sigma = 0.1$ was then added to the image points, finally a varying portion of the observations were replaced by gross outliers. We then compared the errors produced by our high-breakdown algorithm to that of two commonly used M-estimators, the Huber loss function and Tukey's biweight function, see [10]. These M-estimators were implemented using the Iteratively Reweighted Least Squares method [10].

The inlier noise level for both M-estimators was varied as $c = \{0.02,\ 0.2,\ 2.0\}$. We perturbed the ground truth with modest amounts of noise to obtain initialisations for each of the methods. The results, averaged over 500 runs, can be seen in figure 1. This plot shows the distribution of the errors, that is the squared L_2-norms of the reprojection errors with respect to *ground truth*, produced by LQS (with $k = 0.7n$), the Huber- and the Tukey M-estimators for 10% and 30% outliers. We also include the results obtained when initializing the M-estimators with ground truth. The average



(a) LQS vs. Huber, 10% outliers (b) LQS vs. Tukey, 10% outliers

(c) LQS vs. Huber, 30% outliers (d) LQS vs. Tukey, 30% outliers

Figure 1. Histogram of the MSE produced by LQS (k=0.7n) and the Huber and Tukey M-Estimators with varying values of c.

mean squared errors and execution times are given in table 1. As evidenced by the above results LQS perform well across varying degree of outlier corruption and achieves errors significantly smaller than that of the two M-estimators. The Huber loss function M-estimator appears affected by the number of outliers to a greater extent and produces inferior results to the Tukey M-estimator. Improving initialization does not appear to make any difference here, which only serves to validate the understanding we have regarding this low-breakdown point estimator.

The Tukey loss function does produce smaller errors but it is only when initialized with ground truth and a correct noise level estimate that this formulation manages to achieve good results.

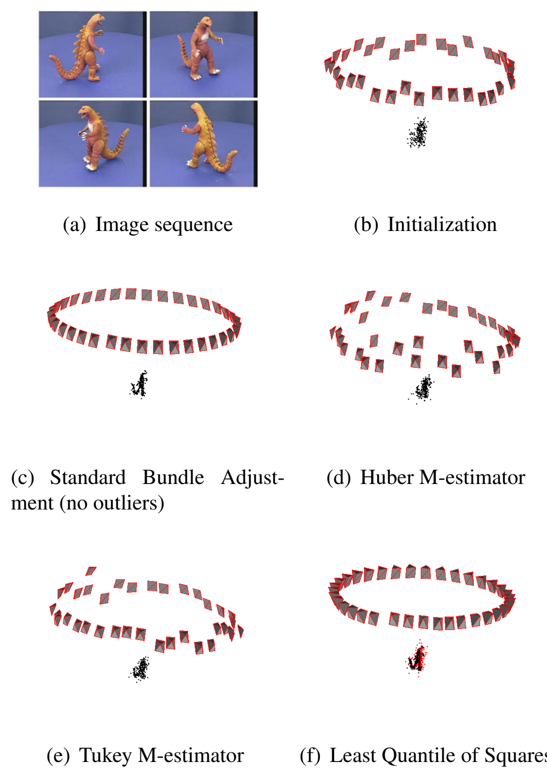
Error Metric		Avg. MSE (10%)	Avg. MSE (30%)	Avg. time
Huber,	$c = 0.02$	0.1962	0.2479	1.13s
	$c = 0.2$	0.2102	0.2693	
Init. by gt	$c = 0.2$	0.2107	0.2622	
	$c = 2.0$	0.3125	0.4364	
Tukey,	$c = 0.02$	0.1630	0.2580	1.28s
	$c = 0.2$	0.1346	0.1938	
Init. by gt	$c = 0.2$	0.0244	0.0398	
	$c = 2.0$	0.1084	0.1848	
LQS,	k=0.7n	0.0388	0.0411	5.21s

Table 1. Average mean squared errors and execution times for 10% and 30% outliers.

Even though the Tukey loss function theoretically has a breakdown point close to 0.5 this formulation does seem to be extremely sensitive to initialization.

5.2. Real Data

Finally we showcase our proposed algorithm on two real world dataset. The dinosaur dataset consisting of a sequence of 36 images taken of a small toy dinosaur on a rotating turn table with 328 points, not containing outliers, tracked partially tracked across these views. And the house dataset [16] consisting of 12 images of a brick building with 1248 tracked points. Outliers were added to 15% of the observations, but the data was not altered in any other way. A full euclidean 3D reconstruction was computed using the two M-estimators as well as the LQS high-breakdown estimator (1). The noise level c was estimated from the uncorrupted data and set to $c = 2.0$ for both M-estimators. Perturbed versions of the ground truth (the result from ordinary least squares estimation on the uncorrupted data) was used for initialization. The result, averaged over 100 runs, is shown in figures 2, 3 and table 2. The average squared L_2 errors on the uncorrupted measurements (similar to the previous experiment), produced by each estimator is shown in table 1. As evidenced, in this instance the reconstructions obtained through high-breakdown estimators are clearly superior to those of the M-estimators.



(a) Image sequence (b) Initialization (c) Standard Bundle Adjustment (no outliers) (d) Huber M-estimator (e) Tukey M-estimator (f) Least Quantile of Squares

Figure 2. Full Euclidean structure-from-motion reconstruction of the dinosaur dataset. We show the reconstruction corresponding to the median MSE of the 100 runs for each method.

6. Conclusions

We have proposed a novel meta-algorithm for high-breakdown parameter estimation applicable to a wide class of computer vision problems. Not only does this scale invariant formulation achieve the theoretically optimal breakdown-point of 50% it also permits a very simple implementation that allows for the leveraging of the computational efficiency of existing state-of-the-art software packages for nonlinear minimization. Preliminary experiments on synthetic as well as real image data demonstrates the superior practical performance of our method over existing robust estimators used in computer vision when the observed data has been corrupted by gross outliers.

Acknowledgements

This research was supported under the Australian Research Councils Discovery Early Career Researcher Award project DE130101775.

References

[1] S. Agarwal and K. Mierle. *Ceres Solver: Tutorial & Reference*. Google Inc.

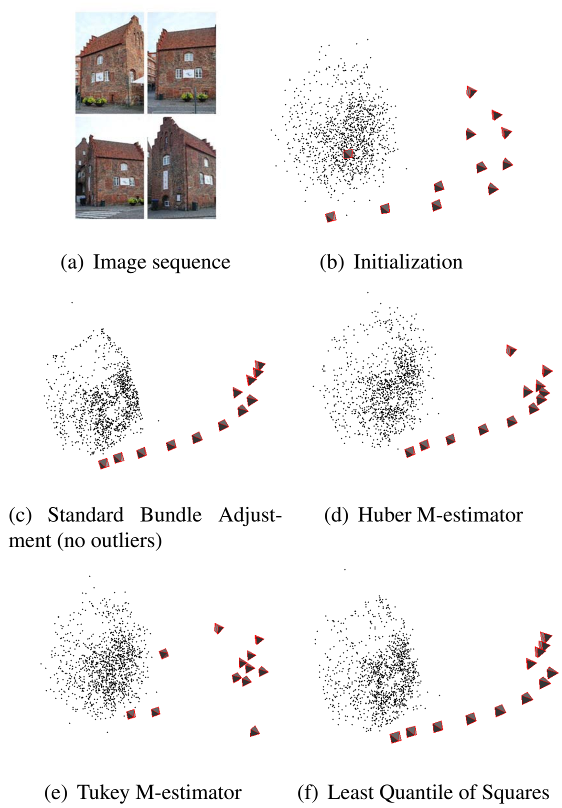
[2] D. P. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods (Optimization and Neural Computation Series)*. Athena Scientific, 1 edition, 1996.

[3] M. Byrod and K. Astrom. Conjugate gradient bundle adjustment. In *Proceedings of the 11th European conference on Computer vision: Part II*, ECCV'10, pages 114–127, Berlin, Heidelberg, 2010. Springer-Verlag.

[4] P. L. Combettes and J.-C. Pesquet. Proximal Splitting Meth-

	Dinosaur		House	
Error Metric	Avg. MSE	Avg. time	Avg. MSE	Avg. time
Huber, $c = 2.0$	3.15	9.28	18.84	32.92
Tukey, $c = 2.0$	10.07	9.24	22.67	32.32
LQS, k=0.8n	0.095	36.73	0.62	210.31
L_2 *sol. (no outl.)*	*0.027*	-	0.035	-

Table 2. The average mean squared error on inliers on the dinosaur and house datasets with 15% of the measurements corrupted by gross outliers.



(a) Image sequence (b) Initialization

(c) Standard Bundle Adjustment (no outliers) (d) Huber M-estimator

(e) Tukey M-estimator (f) Least Quantile of Squares

Figure 3. Full Euclidean structure-from-motion reconstruction of the house dataset.

ods in Signal Processing. In H. Bauschke, R. Burachik, P. Combettes, V. Elser, D. Luke, and H. E. Wolkowicz, editors, *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, pages 185–212. Springer, 2011.

[5] D. Crandall, A. Owens, N. Snavely, and D. P. Huttenlocher. Discrete-continuous optimization for large-scale structure from motion. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2011.

[6] C. Engels, H. Stewénius, and D. Nistér. Bundle adjustment rules. In *Photogrammetric Computer Vision (PCV)*. ISPRS, Sept. 2006.

[7] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381–395, June 1981.

[8] M. Fukushima and H. Mine. A generalized proximal point algorithm for certain non-convex minimization problems. *International Journal of Systems Science*, 12(8):989–1000, 1981.

[9] T. Goldstein and S. Osher. The split bregman method for l1-regularized problems. *SIAM Journal on Imaging Science*, 2(2):323–343, Apr. 2009.

[10] P. Huber. *Robust Statistics*. Wiley, New York, 1974.

[11] Y. Jeong, D. Nistér, D. Steedly, R. Szeliski, and I.-S. Kweon. Pushing the envelope of modern methods for bundle adjustment. *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(8):1605–1617, 2012.

[12] M. A. Lourakis and A. Argyros. SBA: A Software Package for Generic Sparse Bundle Adjustment. *ACM Trans. Math. Software*, 36(1):1–30, 2009.

[13] R. A. Maronna, D. R. Martin, and V. J. Yohai. *Robust Statistics: Theory and Methods*. John Wiley and Sons, New York, 2006.

[14] P. Meer. Robust techniques for computer vision. *Emerging Topics in Computer Vision*, pages 107–190, 2004.

[15] J. J. Moreau. Fonctions convexes duales et points proximaux dans un espace hilbertien. *Comptes Rendus de l'Académie des Sciences (Paris), Série A*, 255:2897–2899, 1962.

[16] C. Olsson, A. Eriksson, and R. Hartley. Outlier removal using duality. 2010.

[17] P. Rousseeuw. Least median of squares regression. *Journal of the American Statistical Association*, 79:871–880, 1984.

[18] P. J. Rousseeuw and B. C. v. Zomeren. Unmasking multivariate outliers and leverage points. *Journal of the American Statistical Association*, 85(411):pp. 633–639, 1990.

[19] S. Setzer. Split bregman algorithm, douglas-rachford splitting and frame shrinkage. *Scale space and variational methods in computer vision*, pages 464–476, 2009.

[20] K. Sim and R. Hartley. Removing outliers using the l_∞ norm. pages 485–494, 2006.

[21] S. Sra. Fast projections onto mixed-norm balls with applications. *Data Mining and Knowledge Discovery*, 25(2):358–377, 2012.

[22] B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon. Bundle adjustment - a modern synthesis. In *Proceedings of the International Workshop on Vision Algorithms: Theory and Practice*, ICCV '99, pages 298–372, London, UK, UK, 2000. Springer-Verlag.

[23] C. Wu, S. Agarwal, B. Curless, and S. M. Seitz. Multicore bundle adjustment. In *CVPR*, pages 3057–3064. IEEE, 2011.