

# Globally Optimal Least Squares Solutions for Quasiconvex 1D Vision Problems

Carl Olsson, Martin Byröd, Fredrik Kahl  
{calle,martin,fredrik}@maths.lth.se

Centre for Mathematical Sciences  
Lund University, Lund, Sweden

**Abstract.** Solutions to non-linear least squares problems play an essential role in structure and motion problems in computer vision. The predominant approach for solving these problems is a Newton like scheme which uses the hessian of the function to iteratively find a local solution. Although fast, this strategy inevitably leads to issues with poor local minima and missed global minima.

In this paper rather than trying to develop an algorithm that is guaranteed to always work, we show that it is often possible to verify that a local solution is in fact also global. We present a simple test that verifies optimality of a solution using only a few linear programs. We show on both synthetic and real data that for the vast majority of cases we are able to verify optimality. Further more we show even if the above test fails it is still often possible to verify that the local solution is global with high probability.

## 1 Introduction

The most studied problem in computer vision is perhaps the (2D) least squares triangulation problem. Even so no efficient globally optimal algorithm has been presented. In fact studies indicate (e.g. [1]) that it might not be possible to find an algorithm that is guaranteed to always work. On the other hand, under the assumption of Gaussian noise the  $\mathbb{L}_2$ -norm is known to give the statistically optimal solution. Although this is a desirable property it is difficult to develop efficient algorithms that are guaranteed to find the globally optimal solution when projections are involved. Lately researchers have turned to methods from global optimization, and a number of algorithms with guaranteed optimality bounds have been proposed (see [2] for a survey). However these algorithms often exhibit (worst case) exponential running time and they can not compare with the speed of local, iterative methods such as bundle adjustment [3–5].

Therefore a common heuristic is to use a minimal solver to generate a starting guess for a local method such as bundle adjustment [3]. These methods are often very fast, however since they are local the success depends on the starting point. Another approach is to minimize some algebraic criteria. Since these typically don't have any geometric meaning this approach usually results in poor reconstructions.

A different approach is the use the maximum residual error rather than the sum of squared residuals. This yields a class of quasiconvex problems where it

is possible to devise efficient global optimization algorithms [6]. This was done in the context of 1D cameras in [7].

Still, it would be desirable to find the statistically optimal solution. In [8] it was shown that for the 2D-triangulation problem (with spherical 2D-cameras) it is often possible to verify that a local solution is also global using a simple test. It was shown on real datasets that for the vast majority of all cases the test was successful. From a practical point of view this is of great value since it opens up the possibility of designing systems where bundle adjustment is the method of choice and only turning to more expensive global methods when optimality can not be verified.

In [9] a stronger condition was derived and the method was extended to general quasiconvex multiview problems (with 2D pinhole cameras).

In this paper we extend this approach 1D multiview geometry problems with spherical cameras. We show that for most real problems we are able to verify that a local solution is global. Further more in case the test fails we show that it is possible to relax the test to show that the solution is global with large probability.

## 2 1D-camera systems

Before turning to the least squares problem we will give a short review of 1D-vision (see [7]). Throughout the paper we will use spherical 1D-cameras.

We start by considering a camera that is located at the origin with zero angle to the  $Y$  axis (see figure 1). For each 2D-point  $(X, Y)$  our camera gives a direction in which the point has been observed. The direction is given in the form of an angle  $\theta$  with respect to a reference axis (see figure 1). Let  $\Pi : \mathbb{R}^2 \mapsto [-\pi/2, \pi/2]$  be defined by

$$\Pi(X, Y) = \text{atan}^2 \frac{X}{Y}. \quad (1)$$

if  $Y > 0$ , (otherwise we let  $\Pi(X, Y) = \infty$ ). The function  $\Pi(X, Y)$  measures the squared angle between the  $Y$ -axis and the vector  $U = (X \ Y)^T$ . Here we have explicitly written  $\Pi(X, Y)$  to indicate that  $\Pi$  always takes values in  $\mathbb{R}^2$ , however throughout the paper we will use both  $\Pi(X, Y)$  and  $\Pi(U)$ . Now, suppose that we have a measurement of a point with angle  $\theta = 0$ . Then  $\Pi$  can be interpreted as the squared angular distance between the point  $(X, Y)$  and the measurement. If the measurement  $\theta$  is not zero we let  $R_{-\theta}$  be a rotation  $-\theta$  then  $\Pi(R_{-\theta}U)$  can be seen as the squared angular distance  $(\phi - \theta)^2$ .

Next we introduce the camera parameters. The camera may be located anywhere in  $\mathbb{R}^2$  with any orientation with respect to a reference coordinate system. In practice we have two coordinate systems, the camera- and the reference- coordinate system. To relate these two we introduce a similarity transformation  $P$  that takes points coordinates in the reference system and transforms then into coordinates in the camera system. We let

$$P = \begin{pmatrix} a & -b & c \\ b & a & d \end{pmatrix} \quad (2)$$

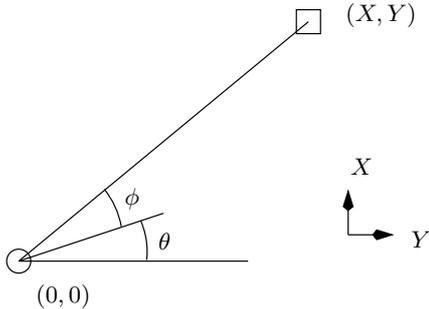


Fig. 1: 1D camera geometry for a calibrated camera.

The parameters  $(a, b, c, d)$  are what we call the inner camera parameters and they determine the orientation and position of the camera. The squared angular error can now be written

$$\Pi \left( R_{-\theta} P \begin{bmatrix} U \\ 1 \end{bmatrix} \right) \quad (3)$$

In the remaining part of the paper the concept of quasiconvexity will be important. A function  $f$  is said to be quasiconvex if its sublevel sets  $S_\phi(f) = \{x; f(x) \leq \phi\}$  are convex. In the case of triangulation (as well as resectioning) we see that the squared angular errors (3) can be written as the composition of the projection  $\Pi$  and two affine functions

$$X_i(x) = a_i^T x + \tilde{a}_i \quad (4)$$

$$Y_i(x) = b_i^T x + \tilde{b}_i \quad (5)$$

(here  $i$  denotes the  $i$ 'th error residual). It was shown in [7] that functions of this type are quasiconvex. The advantage of quasiconvexity is that a function with this property can only have a single local minimum, when using the  $\mathbb{L}_\infty$ -norm. This class of problems include, among others, camera resectioning and triangulation.

In this paper, we will use the theory of quasiconvexity as a stepping stone to verify global optimality under the  $L_2$  Norm. Our approach closely parallels that of [8] and [9]. However while [8] considered spherical 2D cameras only for the triangulation problem and [9] considered 2D-pinhole cameras for general multiview problems, we will consider 1D-spherical cameras.

### 3 Theory

In this section we will give sufficient conditions for global optimality. If  $x^*$  is a global minimum then there is an open set containing  $x^*$  where the Hessian of  $f$  is positive semidefinite. Recall that a function is convex if and only if its Hessian is positive semidefinite. The basic idea which was first introduced in [8] is the following: If we can find a convex region  $\mathcal{C}$  containing  $x^*$  that is large enough to include all globally optimal solutions and we are able to show that the Hessian of  $f$  is convex on this set, then  $x^*$  must be the globally optimal solution.

### 3.1 The set $\mathcal{C}$

The first step is to determine the set  $\mathcal{C}$ . Suppose that for our local candidate solution  $x^*$  we have  $f(x^*) = \phi_{max}^2$ . Then clearly any global optimum must fulfill  $f_i(x) \leq \phi_{max}^2$  for all residuals, since otherwise our local solution is better. Hence we take the region  $\mathcal{C}$  to be

$$\mathcal{C} = \{x \in \mathbb{R}^n, f_i(x) \leq \phi_{max}^2\}. \quad (6)$$

It is easily seen that this set is convex since this is the intersection of the sublevel sets  $S_{\phi_{max}^2}(f_i)$  which are known to be convex since the residuals  $f_i$  are quasiconvex. Hence if we can show that the Hessian of  $f$  is positive definite on this set we may conclude that  $x^*$  is the global optimum.

Note that the condition  $f_i(x) \leq \phi_{max}^2$  is somewhat pessimistic. Indeed it assumes that the entire error may occur in one residual which is highly unlikely under any reasonable noise model. In fact we will show that it is possible to replace  $\phi_{max}^2$  with a stronger bound to show that  $x^*$  is with high probability the global optimum.

### 3.2 Bounding the Hessian

The goal of this section is to show that the Hessian of  $f$  is positive semidefinite on the set  $\mathcal{C}$ . To do this we will find a constant matrix  $H$  that acts as a lower bound on  $\nabla^2 f(x)$  for all  $x \in \mathcal{C}$ . More formally we will construct  $H$  such that  $\nabla^2 f(x) \succeq H$  on  $\mathcal{C}$ , that is if  $H$  is positive semidefinite then so is  $\nabla^2 f(x)$ . We begin by studying the 1D-projection mapping  $\Pi$ . The Hessian of  $\Pi$  is

$$\nabla^2 \Pi(X, Y) = \frac{2}{(X^2 + Y^2)^2} \begin{pmatrix} Y^2 - XY \operatorname{atan} \frac{X}{Y} & (X^2 - Y^2) \operatorname{atan} \frac{X}{Y} - XY \\ (X^2 - Y^2) \operatorname{atan} \frac{X}{Y} - XY & X^2 + XY \operatorname{atan} \frac{X}{Y} \end{pmatrix} \quad (7)$$

To simplify notation we introduce the measurement angle  $\phi = \operatorname{atan} \frac{X}{Y}$  and the radial distance to the camera center  $r = \sqrt{X^2 + Y^2}$ . After a few simplifications one obtains

$$\nabla^2 \Pi(X, Y) = \frac{1}{r^2} \begin{pmatrix} 1 + \cos(2\phi) - 2\phi \sin(2\phi) & -\sin(2\phi) - 2\phi \cos(2\phi) \\ -\sin(2\phi) - 2\phi \cos(2\phi) & 1 - \cos(2\phi) + 2\phi \sin(2\phi) \end{pmatrix} \quad (8)$$

In the case of 3D to 2D projections Hartley et.al. [8] obtained a similar  $3 \times 3$  matrix. Using the same arguments it may be seen our matrix can be bounded by the diagonal matrix

$$H(X, Y) = \nabla^2 \Pi(X, Y) = \frac{2}{r^2} \begin{pmatrix} \frac{1}{4} & 0 \\ 0 & -4\phi^2 \end{pmatrix} \quad (9)$$

To see this we need to show that the eigenvalues of  $\nabla^2 \Pi(X, Y) - H(X, Y)$  are all positive. Taking the trace of this matrix we see that the sum of the eigenvalues are  $\frac{1}{r^2}(3/2 + 8\phi^2)$  which is always positive. We also have the determinant

$$\det(\nabla^2 \Pi(X, Y) - H(X, Y)) = -1 + (1 + 16\phi^2)(\cos(2\phi) - 2\phi \sin(2\phi)) \quad (10)$$

It can be shown (see [8]) that this expression is positive if  $\phi \leq 0.3$ . Hence for  $\phi \leq 0.3$ ,  $H(X, Y)$  is a lower bound on  $\nabla^2 \Pi(X, Y)$ .

Now, the error residuals  $f_i(x)$  of our class of problems are related to the projection mapping via an affine change of coordinates

$$f_i(x) = \Pi(a_i^T x + \tilde{a}_i, b_i^T x + \tilde{b}_i). \quad (11)$$

It was noted in [9] that since the coordinate change is affine the Hessian of  $f_i$  is can be bounded by  $H$ . To see this we let  $W_i$  be the matrix containing  $a_i$  and  $b_i$  as columns Using the chain rule we obtain the Hessian

$$\nabla^2 f_i(x) = W_i \nabla^2 \Pi(a_i^T x + \tilde{a}_i, b_i^T x + \tilde{b}_i) W_i^T. \quad (12)$$

And since  $\nabla^2 \Pi$  is bounded by  $H$  we obtain

$$\nabla^2 f(x) \succeq \sum_i W_i H(a_i^T x + \tilde{a}_i, b_i^T x + \tilde{b}_i) W_i^T = \sum_i \frac{2}{r_i^2} \left( \frac{a_i a_i^T}{4} - 4\phi_i^2 b_i b_i^T \right). \quad (13)$$

The matrix appearing on the right hand side of (13) seems easier to handle however it still depends on  $x$  through  $r$  and  $\phi$ . This dependence may be removed by using bound of the type

$$\phi \leq \phi_{\max} \quad (14)$$

$$r_{i,\min} \leq r_i \leq r_{i,\max} \quad (15)$$

The first bound is readily obtained since  $x \in \mathcal{C}$ . In the second one we need to find an upper and lower bound on the radial distance in every camera. We shall see later that this can be cast as a convex problem which can be solved efficiently. As is [9] we now obtain the bound

$$\nabla^2 f(x) \succeq \sum_i \left( \frac{1}{2r_{i,\max}^2} a_i a_i^T - 8 \frac{\phi_{\max}^2}{r_{i,\min}^2} b_i b_i^T \right). \quad (16)$$

Hence if the minimum eigenvalue of the right hand side is non negative the function  $f$  will be convex on the set  $\mathcal{C}$ .

### 3.3 Bounding the Radial Distances $r_i$

In order to be able to use the criterion (13) we need to be able to compute bounds on the radial distances. The  $k$ 'th radial distance may be written

$$r_k(x) = \sqrt{(a_k^T x + \tilde{a}_k)^2 + (b_k^T x + \tilde{b}_k)^2} \quad (17)$$

Since  $x \in \mathcal{C}$  we know that (see [7])

$$(a_k^T x + \tilde{a}_k)^2 + (b_k^T x + \tilde{b}_k)^2 \leq (1 + \tan^2(\phi_{\max})) (b_k^T x + \tilde{b}_k)^2 \quad (18)$$

and obviously

$$(a_k^T x + \tilde{a}_k)^2 + (b_k^T x + \tilde{b}_k)^2 \geq (b_k^T x + \tilde{b}_k)^2 \quad (19)$$

The bound (15) can be obtained by solving the linear programs

$$r_{k,\max} = \max \sqrt{(1 + \tan^2(\phi_{\max}))} (b_k^T x + \tilde{b}_k) \quad (20)$$

$$\text{s.t. } |a_i^T x + \tilde{a}_i| \leq \tan(\phi_{\max})(b_i^T x + \tilde{b}_i), \quad \forall i \quad (21)$$

and

$$r_{k,\min} = \min (b_k^T x + \tilde{b}_k) \quad (22)$$

$$\text{s.t. } |a_i^T x + \tilde{a}_i| \leq \tan(\phi_{\max})(b_i^T x + \tilde{b}_i), \quad \forall i. \quad (23)$$

At first glance this may seem as a quite rough estimate, however since  $\phi_{\max}$  is usually small this bound is good enough. By using SOCP-programming instead of linear programming it is possible to improve these bounds, however since linear programming is faster we prefer to use the looser bounds.

To summarize, the following steps are performed in order to verify optimality:

1. Compute a local minimizer  $x^*$  (e.g. with bundle adjustment).
2. Compute maximum/minimum radial depths over  $\mathcal{C}$ .
3. Test if the convexity condition in (16) holds.

## 4 A Probabilistic Approach

In practice, the constraints  $f_i(x) \leq \phi_{max}^2$  are often overly pessimistic. In fact what is assumed here is that the entire residual error  $\phi_{max}^2$  could (in worst case) arise from a single error residual, which is not very likely. Assume that  $\hat{x}_i$  is the point measurements that would be obtained in a noise free system and that  $x_i$  is the real measurement. Under the assumption of independent Gaussian noise we have

$$\hat{x}_i - x_i = r_i, \quad r_i \sim \mathcal{N}(0, \sigma). \quad (24)$$

Since  $r_i$  has zero mean, an unbiased estimate of  $\sigma$  is given by

$$\hat{\sigma} = \sqrt{\frac{1}{m-d} \phi_{max}}, \quad (25)$$

where  $m$  is the number of residuals and  $d$  denotes the number of degrees of freedom in the underlying problem (for example,  $d = 2$  for 2D triangulation and  $d = 3$  for 2D calibrated resectioning). As before, we are interested in finding a bound for each residual. This time, however, we are satisfied with a bound that holds *with high probability*. Specifically, given  $\hat{\sigma}$ , we would like to find  $L(\hat{\sigma})$  so that

$$P[\forall i : -L(\hat{\sigma}) \leq r_i \leq L(\hat{\sigma})] \geq P_0 \quad (26)$$

for a given confidence level  $P_0$ . To this end, we make use of a basic theorem in statistics which states that  $\frac{X}{\sqrt{Y/\gamma}}$  is  $T$ -distributed with  $\gamma$  degrees of freedom, when  $X$  is normal with mean 0 and variance 1,  $Y$  is a chi squared random variable with  $\gamma$  degrees of freedom and  $X$  and  $Y$  are independent. A further

basic fact from statistics states that,  $\hat{\sigma}^2(m-d)/\sigma^2$  is chi squared distributed with  $\gamma = m - d$  degrees of freedom. Thus,

$$\frac{r_i}{\hat{\sigma}} = \frac{r_i/\sigma}{\sqrt{\hat{\sigma}^2/\sigma^2}} \quad (27)$$

fulfills the requirements to be  $T$  distributed apart from a small dependence between  $r_i$  and  $\hat{\sigma}$ . This dependence, however, vanishes with enough residuals and in any case leads to a slightly more conservative bound.

Given a confidence level  $\beta$  we can now *e.g.* do a table lookup for the  $T$  distribution to get  $t_\gamma^\beta$  so that

$$P[-t_\gamma^\beta \leq \frac{r_i}{\hat{\sigma}} \leq t_\gamma^\beta] \geq \beta. \quad (28)$$

Multiplying through with  $\hat{\sigma}$  we obtain  $L(\hat{\sigma}) = \hat{\sigma}t_\gamma^\beta$ . Given a confidence level  $\beta_0$  for all  $r_i$ , we assume that the  $r_i/\hat{\sigma}$  are independent and thus set  $\beta = \beta_0^{1/m}$  to get

$$P[\forall i : -t_\gamma^\beta \leq \frac{r_i}{\hat{\sigma}} \leq t_\gamma^\beta] \geq \beta_0. \quad (29)$$

The independence assumption is again only approximately correct, but similarly yields a slightly more conservative bound than necessary.

## 5 Experiments

In this section we demonstrate our theory on a few experiments. We used two real datasets to verify the theory. The first one is measurements of measurements performed at a ice hockey rink. The set contains 70 1D-images (with 360 degree field of view) and 14 reflectors. Figure 2 shows the setup, the motion of the cameras and the position of the reflectors.

The structure and motion was obtained using the  $L_\infty$  optimal methods from [7]. We first picked 5 cameras and solved structure and motion for these cameras and the viewed reflectors. We then added the remaining cameras and reflectors using alternating resection and triangulation. Finally we did bundle adjustment to obtain locally optimal  $L_2$  solutions. We then ran our test on all (14) triangulation and (70) resectioning subproblems in this and in every case we were able to verify that these subproblems were in fact globally optimal. Figure reffhockey2 shows one instance of the triangulation problem and one instance of the resectioning problem. The  $L_2$  angular errors where roughly the same ( $\approx 0.1$ - $0.2$  for both triangulation and resectioning) throughout the sequence.

In the hockey rink dataset the the cameras are placed so that the angle measurements can take roughly any value in  $[-\pi, \pi]$ . In our next dataset we wanted to test what happens if the measurements are restricted to a smaller interval. It is well known that for example resectioning is easier if one has measurements in wide spread directions. Therefore we used a data set where the the cameras do not have a 360 field of view and where there are not reflectors in every direction. Figure 4 shows the setup. We refer to this data set as the the coffee room

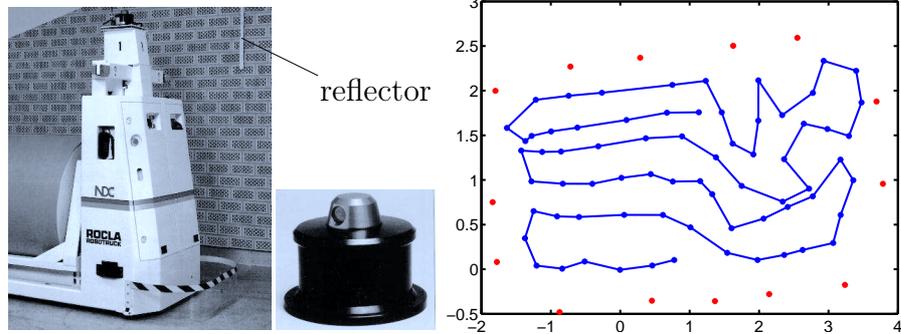


Fig. 2: Left: A laser guided vehicle. Middle: A laser scanner or angle meter. Right: positions of the reflectors and motion for the vehicle.

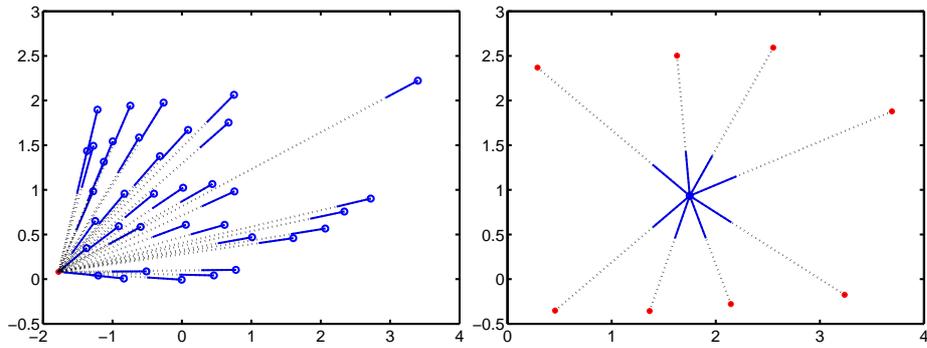


Fig. 3: Left: An instance of the triangulation problem. The reflector is visible from 36 positions with the total angular  $L_2$ -error of 0.15 degrees. Right: An instance of the resectioning problem. The camera detected 8 reflectors with the total angular  $L_2$ -error of 0.12 degrees.

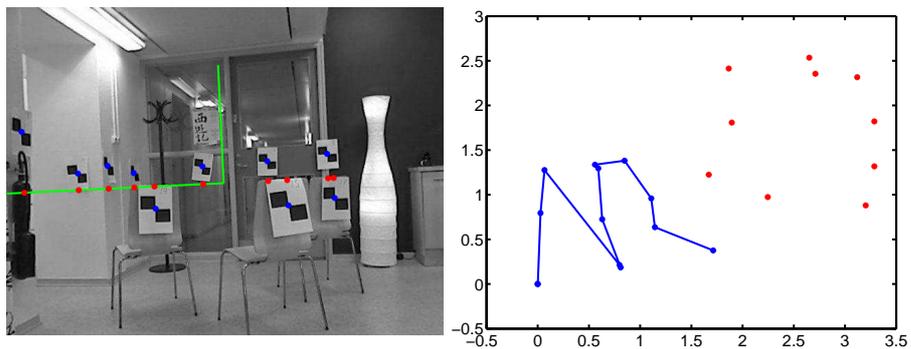


Fig. 4: Left: An images from the coffee room sequence. The green lines are estimated horizontal and vertical directions in the image, the blue dots are detected markers and red dots are the estimated bearings to the markers. Right: Positions of the markers and motion for the camera.

sequence since it was taken in our coffee room. Here we have places 10 markers in in various positions and used regular 2D-cameras to obtain 13 images. (Some of the images are difficult to make out in figure 4 since they where taken close together only varying orientation.) The to estimate the angular bearings to the markers we first estimated the vertical and horizontal green lines in the figures. The detected 2D-marker positions was then projected onto the horizontal line and the angular bearings was computed. This time we computed the the structure and motion using a minimal case solver (3-cameras 5-markers) and then alternated resection-intersection followed by bundle adjustment. We then ran all the triangulation and resectioning subproblems and in all cases we where able to verify optimality. This time the  $L_2$  angular errors was more varied. For triangulation most of the errors where around 0.5-1 degree whereas for resectioning the most of error where smaller ( $\approx 0.1-0.2$ ). Although in one camera  $L_2$ -error was as large as 3.2 degrees, however we were still able to verify that the resection was optimal.

### 5.1 Probabilistic Verification of Optimality

In this section we study the effect of the tighter bound one obtains by accepting a small, but calculable risk of missing the global optimum. Here, we would like to see how varying degrees of noise affects the ability to verify a global optimum and hence set up a synthetic experiment with randomly generated 1D cameras and points. For the experiment, 20 cameras and 1 point were generated uniformly at random in the square  $[0.5, 0.5]^2$  and noise was added. The experiment was repeated 20 times at each noise level with noise standard deviation from 0 to 3.5 degrees and for each noise level we recorded the proportion of instances where the global optimum could be verified. We performed the whole procedure once with the exact bound and once with a bound set at a 95% confidence level. The result is shown in Figure 5. As expected, the tighter 95% bound allows one to verify a substantially larger proportion of cases.

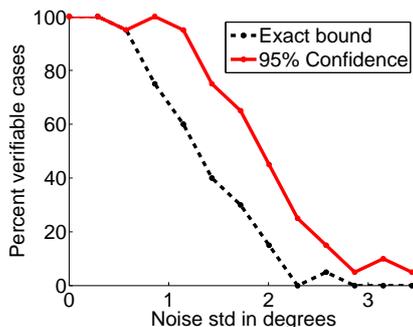


Fig. 5: Proportion of instances where global optimality could be verified versus image noise.

## 6 Conclusions

Global optimization of the reprojection errors in  $L_2$  norm is desirable, but difficult and no really practical general purpose algorithm exists. In this paper we have shown in the case of 1D vision how local optima can be checked for global optimality and found that in practice, local optimization paired with clever initialization is a powerful approach which often finds the global optimum. In particular our approach might be used in a system to filter out only the truly difficult local minima and pass these on to a more sophisticated but expensive global optimizer.

## 7 Acknowledgments

This work has been funded by the European Research Council (GlobalVision grant no. 209480), the Swedish Research Council (grant no. 2007-6476) and the Swedish Foundation for Strategic Research (SSF) through the programme Future Research Leaders. Travel funding has been received from The Royal Swedish Academy of Sciences and the Foundation Stiftelsen J.A. Letterstedts resesitpendiefond.

## References

1. Stewénius, H., Schaffalitzky, F., Nistér, D.: How hard is three-view triangulation really? In: Int. Conf. Computer Vision, Beijing, China (2005) 686–693
2. Hartley, R., Kahl, F.: Optimal algorithms in multiview geometry. In: Asian Conf. Computer Vision, Tokyo, Japan (2007)
3. Triggs, B., McLauchlan, P., R.I. H., Fitzgibbon, A.: Bundle adjustment - a modern synthesis. In: Vision Algorithms'99, in conjunction with ICCV'99, Kerkyra, Greece (1999) 298–372
4. Engels, C., Stewénius, H., Nistér, D.: Bundle adjustment rules. In: Photogrammetric Computer Vision (PCV). (2006)
5. Kai, N., Steedly, D., Dellaert, F.: Out-of-core bundle adjustment for large-scale 3d reconstruction. In: Conf. Computer Vision and Pattern Recognition, Minneapolis, USA (2007)
6. Hartley, R., Kahl, F.: Critical configurations for projective reconstruction from multiple views. Int. Journal Computer Vision **71** (2007) 5–47
7. Åström, K., Enqvist, O., Olsson, C., Kahl, F., Hartley, R.: An  $L_\infty$  approach to structure and motion problems in 1d-vision. In: Int. Conf. Computer Vision, Rio de Janeiro, Brazil (2007)
8. Hartley, R., Seo, Y.: Verifying global minima for  $L_2$  minimization problems. In: Conf. Computer Vision and Pattern Recognition, Anchorage, USA (2008)
9. Olsson, C., Kahl, F., Hartley, R.: Projective least squares: Global solutions with local optimization. In: Proc. Int. Conf. Computer Vision and Pattern Recognition. (2009)