GLOBALLY OPTIMAL SOLUTION TO EXPLOIT RIGIDITY WHEN RECOVERING STRUCTURE FROM MOTION UNDER OCCLUSION

Pedro M. Q. Aguiar,   João M. F. Xavier,   Marko Stošić

Institute for Systems and Robotics / IST, Lisboa, Portugal
{aguiar,jxavier,mstosic}@isr.ist.utl.pt

ABSTRACT

Widely used SVD-based matrix factorization approaches to the recovery of 3D rigid structure from motion (SFM), require a set of feature points to be visible in a set of images. When there is occlusion, several feature points disappear, the observation matrix misses some entries, there is not equivalent to the SVD, and only suboptimal solutions have been proposed to exploit rigidity. In this paper, we propose a method to complete the trajectories that correspond to a rigid scene, in an optimal way. Our algorithm is not iterative (thus avoiding problems like sensitivity to initialization and local optima); it rather computes in a finite number of steps the globally optimal completion of the observation matrix. We describe experiments that illustrate the gain in accuracy of SFM.

Index Terms— Structure from motion, occlusion, missing data, global optimization, matrix completion, SPOC.

1. INTRODUCTION

A matrix collecting trajectories of feature points of a rigid body is rank deficient in a noiseless scenario. This fact underlies the success of the so-called matrix factorization methods to recover 3D rigid structure from motion (SFM). These methods compute the best rank deficient approximation of a noisy observation matrix from its Singular Value Decomposition (SVD), see *e.g.*, [1, 2].

In practice, it is often the case that, due to occlusion, many feature points become invisible to the camera, and the observation matrix misses several entries, *i.e.*, it is incomplete, preventing thus the use of SVD. In the last decade, several authors addressed the problem of recovering SFM under occlusion, by attempting to complete the observation matrix. Although earlier examples are suboptimal procedures to combine the constrains that arise from the observed submatrices of the original matrix [3], in the recent past, there have been proposed several optimization algorithms that iteratively refine the solution in a local way, by minimizing a nonlinear cost function that measures the residual of the approximation, *e.g.*, [4, 5, 6]. These algorithms are sensitive to the initialization and often fall into local minima.

In this paper, we focus on particular patterns of missing entries, namely those known as a Young diagram, which include the cases that often arise when dealing with occlusion in the recovery of rigid SFM. It was recently shown that inequalities relating the singular values of a matrix with those of its submatrices enable the development of a strategy to compute the globally optimal completion of such matrices, in a finite number of steps [7]. Our experiments in

this paper demonstrate the effectiveness of this approach when recovering SFM. Naturally, the completion algorithm, called SPOC (from SPectrally Optimal Completion), is relevant for many other image/signal processing problems involving partial observations of rank deficient matrices, *e.g.*, separation of style and content in images [8] or localization in a network of acoustic sensors [9].

2. PROBLEM FORMULATION

In the factorization method for the recovery of rigid SFM, a set of N feature points is tracked across F frames and their trajectories are collected in a $2F \times N$ observation matrix $\mathbf{W}$. Usually, each column of $\mathbf{W}$ contains the sequence of (pairs of) coordinates of each feature point. Due to the rigidity of the scene, $\mathbf{W}$ is rank deficient in a noiseless situation – it is rank 4 in general, for the affine camera model, see, *e.g.*, [1, 2]. When $\mathbf{W}$ is completely known, the 3D structure is recovered from the rank 4 matrix that best matches $\mathbf{W}$, which is easily computed from its SVD, by selecting the 4 largest singular values.

When a feature point is occluded (or missed during tracking for any other reason) at a given frame, the corresponding trajectory ends at that frame and the remaining entries of the corresponding column of the observation matrix $\mathbf{W}$ are treated as missing data. The left image of Fig. 1 shows the typical pattern of known/missing entries of $\mathbf{W}$ in this scenario (for $N = 20$ and $F = 6$). For reasons that will become clear in the sequel, without loss of generality, we re-arrange the observation matrix by performing a column re-ordering, in such a way that the pattern of missing entries becomes a Young diagram. This pattern is characterized by a sequence of non-increasing number of missing entries from the first (leftmost) to the last column, see the example in the right plot of Fig. 1.
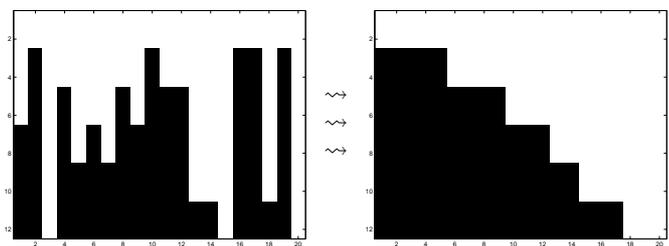

Fig. 1. Example of pattern of missing entries (depicted in black) in the observation matrix obtained when recovering SFM under occlusion. An adequate column re-ordering produces a Young diagram.

We address the problem of recovering rigid SFM under occlusion as a search for the best completion of the observation matrix $\mathbf{W}$. Since $\mathbf{W}$ is rank 4 in a noiseless situation, we formalize the

---

problem as the minimization of the fifth largest singular value over all possible completions. This leads to the optimization problem

$$\widehat{\mathbf{X}} = \arg\min_{\mathbf{X} \in \mathcal{S}_W} \sigma_5(\mathbf{X}), \tag{1}$$

where $\mathcal{S}_W$ is the set of possible completions of $\mathbf{W}$ and $\sigma_5(\cdot)$ denotes the fifth largest singular value. $\mathbf{X}$ denotes a candidate, *i.e.*, a matrix whose entries corresponding to observed entries in $\mathbf{W}$ are fixed at the observation values, and $\widehat{\mathbf{X}}$ is the best completion, which minimizes the spectral norm of the matrix of residuals [10]; for this reason, we call it the SPectrally Optimal Completion (SPOC).

At this point, we must emphasize that, although other authors have addressed the recovery of SFM under occlusion by seeking optimal completions of the observation matrix, *e.g.*, [4, 5, 6], they used a fully general pattern of missing entries and ended up proposing iterative optimization methods that often fall in local optima. In opposition, we focus on the class of missing data patterns that often arise when dealing with occlusion, *i.e.*, the Young diagram illustrated in Fig. 1, and derive a non-iterative globally optimal solution.

3. COMPUTING THE GLOBAL OPTIMUM

As a building block for our SPOC algorithm, we start by addressing the case of a $m \times n$ observation matrix $\mathbf{W}$ that misses only the last entry of the first column, *i.e.*, the bottom left entry. The problem (1) becomes

$$\hat{x} = \arg\min_{x \in \mathbb{R}} \sigma_5\left(\begin{bmatrix} \mathbf{a} & \mathbf{C} \\ x & \mathbf{b}^T \end{bmatrix}\right), \tag{2}$$

where the $(m-1) \times 1$ vector $\mathbf{a}$, the $(n-1) \times 1$ vector $\mathbf{b}$, and the $(m-1) \times (n-1)$ matrix $\mathbf{C}$, are known (they contain the observed entries of $\mathbf{W}$) and $x \in \mathbb{R}$ denotes a candidate value for the unknown entry, thus the matrix inside brackets represents a generic element $\mathbf{X}$ of the set of possible completions, $\mathcal{S}_W$.

Our closed-form solution for the problem (2) is rooted on the theorem of interlacing inequalities for singular values, see *e.g.*, [10]. This theorem provides the two following lower bounds for the cost in (2), valid for any choice of $x \in \mathbb{R}$:

$$\sigma_5(\mathbf{X}) \geq \sigma_5\left(\begin{bmatrix} \mathbf{a} & \mathbf{C} \end{bmatrix}\right), \qquad \sigma_5(\mathbf{X}) \geq \sigma_5\left(\begin{bmatrix} \mathbf{C} \\ \mathbf{b}^T \end{bmatrix}\right).$$

It can be shown that the conjunction of these two inequalities provides a tight bound for the cost, *i.e.*, that the minimum value for the cost can be obtained in closed-form as:

$$\sigma_{\min} = \min_{x \in \mathbb{R}} \sigma_5(\mathbf{X}) = \max\left\{\sigma_5\left(\begin{bmatrix} \mathbf{a} & \mathbf{C} \end{bmatrix}\right), \sigma_5\left(\begin{bmatrix} \mathbf{C} \\ \mathbf{b}^T \end{bmatrix}\right)\right\}. \tag{3}$$

Furthermore, in general, the optimization problem (2) is solvable, *i.e.*, there is a value $\hat{x}$ for the missing entry that produces the infimum $\sigma_{\min}$ in (3), see [11] for details:

$$\sigma_5\left(\begin{bmatrix} \mathbf{a} & \mathbf{C} \\ \hat{x} & \mathbf{b}^T \end{bmatrix}\right) = \sigma_{\min}. \tag{4}$$

Now, computing the optimal solution for the missing entry, $\hat{x}$, becomes straight forward because (4) implies that $\sigma_{\min}^2$, obtained from (3), is an eigenvalue (the fifth largest one) of the matrix

$$\begin{bmatrix} \mathbf{a} & \mathbf{C} \\ \hat{x} & \mathbf{b}^T \end{bmatrix}\begin{bmatrix} \mathbf{a} & \mathbf{C} \\ \hat{x} & \mathbf{b}^T \end{bmatrix}^T = \begin{bmatrix} \mathbf{a}\mathbf{a}^T + \mathbf{C}\mathbf{C}^T & \mathbf{a}\hat{x} + \mathbf{C}\mathbf{b} \\ \hat{x}\mathbf{a}^T + \mathbf{b}^T\mathbf{C}^T & \hat{x}^2 + \mathbf{b}^T\mathbf{b} \end{bmatrix},$$

thus we have:

$$\begin{vmatrix} \mathbf{a}\mathbf{a}^T + \mathbf{C}\mathbf{C}^T - \sigma_{\min}^2\mathbf{I}_{m-1} & \hat{x}\mathbf{a} + \mathbf{C}\mathbf{b} \\ \hat{x}\mathbf{a}^T + \mathbf{b}^T\mathbf{C}^T & \hat{x}^2 + \mathbf{b}^T\mathbf{b} - \sigma_{\min}^2 \end{vmatrix} = 0, \tag{5}$$

where $\mathbf{I}_{m-1}$ stands for the $(m-1) \times (m-1)$ identity matrix and $|\cdot|$ represents the determinant. Since the determinant is a multilinear operator (*i.e.*, linear in each column separately), equality (5) is simply a quadratic equation in the unknown $\hat{x}$, which is easily solved (the solution is unique, under general assumptions of solvability [11]).

We now address the Young diagram, *i.e.*, the general pattern of missing data for an observation matrix typical of SFM under occlusion, see Fig. 1. The interesting fact about this class of incomplete matrices is that the optimal completion, *i.e.*, the completion minimizing σ_5, can be found in a sequential way, by solving, for each missing entry, a one-dimensional problem such as the one in (2).

Our SPOC procedure is the following: we start at the first (the topmost) incomplete row and complete it, from right to left, then we move to the following row and complete it, from right to left, and so on, until we have completed the last row. For example, for the following 8×9 observation matrix $\mathbf{W}$ (in SFM, it would correspond to a scenario where 9 feature points are seen at the first image, two of them are lost in the second image, three other points are lost in the third image, and one more is lost in the fourth image, *i.e.*, only 3 points are seen in the 4 images), the order by which the missing entries are filled in is indicated as $\checkmark_1, \checkmark_2, ..., \checkmark_{26}$:

$$\begin{bmatrix}
w_{11} & w_{12} & w_{13} & w_{14} & w_{15} & w_{16} & w_{17} & w_{18} & w_{19} \\
w_{21} & w_{22} & w_{23} & w_{24} & w_{25} & w_{26} & w_{27} & w_{18} & w_{19} \\
\checkmark_2 & \checkmark_1 & w_{33} & w_{34} & w_{35} & w_{36} & w_{37} & w_{18} & w_{19} \\
\checkmark_4 & \checkmark_3 & w_{43} & w_{44} & w_{45} & w_{46} & w_{47} & w_{18} & w_{19} \\
\checkmark_9 & \checkmark_8 & \checkmark_7 & \checkmark_6 & \checkmark_5 & w_{56} & w_{57} & w_{18} & w_{19} \\
\checkmark_{14} & \checkmark_{13} & \checkmark_{12} & \checkmark_{11} & \checkmark_{10} & w_{66} & w_{67} & w_{18} & w_{19} \\
\checkmark_{20} & \checkmark_{19} & \checkmark_{18} & \checkmark_{17} & \checkmark_{16} & \checkmark_{15} & w_{77} & w_{18} & w_{19} \\
\checkmark_{26} & \checkmark_{25} & \checkmark_{24} & \checkmark_{23} & \checkmark_{22} & \checkmark_{21} & w_{87} & w_{18} & w_{19}
\end{bmatrix}.$$

Computing $\checkmark_i$ corresponds to solving (2) for the largest submatrix of $\mathbf{W}$ that has the entry corresponding to $\checkmark_i$ in its bottom left position. Note that, due to the filling order just described, all the other entries of this submatrix are available: they are either observed entries of $\mathbf{W}$ or $\checkmark_j$, with $j \leq i-1$, which were computed before.

A simple example provides insight over why the SPOC procedure just described yields the globally optimal completion. Consider the problem of completing a matrix $\mathbf{W}$ with two missing entries, located at the first and second elements of the last row:

$$\{\hat{x}, \hat{y}\} = \arg\min_{(x,y) \in \mathbb{R}^2} \sigma_5\left(\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{C} \\ x & y & \mathbf{b}^T \end{bmatrix}\right). \tag{6}$$

As for problem (2), the theorem of interlacing inequalities for singular values [10] provides a bound for the minimum value $\sigma_{\min}$ of $\sigma_5(\cdot)$ in (6):

$$\sigma_{\min} \geq \max\left\{\sigma_5\left(\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{C} \end{bmatrix}\right), \sigma_5\left(\begin{bmatrix} \mathbf{C} \\ \mathbf{b}^T \end{bmatrix}\right)\right\}. \tag{7}$$

We will see that the sequential procedure described above produces the solution $(\hat{x}, \hat{y})$ that attains this bound.

According to the SPOC procedure, we first solve a problem of the form (2) for the unknown entry y, *i.e.*, we compute $\hat{y}$ such that

$$\sigma_5\left(\begin{bmatrix} \mathbf{a}_2 & \mathbf{C} \\ \hat{y} & \mathbf{b}^T \end{bmatrix}\right) = \max\left\{\sigma_5\left(\begin{bmatrix} \mathbf{a}_2 & \mathbf{C} \end{bmatrix}\right), \sigma_5\left(\begin{bmatrix} \mathbf{C} \\ \mathbf{b}^T \end{bmatrix}\right)\right\}. \tag{8}$$

Then, we plug in the solution $\hat{y}$ and solve another problem like (2), for the entry x, *i.e.*, we compute $\hat{x}$ such that

$$\sigma_5\left(\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{C} \\ \hat{x} & \hat{y} & \mathbf{b}^T \end{bmatrix}\right) = \max\left\{\sigma_5\left(\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{C}\end{bmatrix}\right), \sigma_5\left(\begin{bmatrix} \mathbf{a}_2 & \mathbf{C} \\ \hat{y} & \mathbf{b}^T \end{bmatrix}\right)\right\}.$$

Replacing (8) into the r.h.s. of the last equality, we get

$$\max\left\{\sigma_5\left(\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{C}\end{bmatrix}\right), \sigma_5\left(\begin{bmatrix} \mathbf{a}_2 & \mathbf{C}\end{bmatrix}\right), \sigma_5\left(\begin{bmatrix} \mathbf{C} \\ \mathbf{b}^T \end{bmatrix}\right)\right\},$$

which can be further simplified, because any singular value of a matrix is greater or equal than the corresponding one of any of its sub-matrices, *i.e.*, $\sigma_5\left(\begin{bmatrix} \mathbf{a}_2 & \mathbf{C}\end{bmatrix}\right) \leq \sigma_5\left(\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{C}\end{bmatrix}\right)$, leading to

$$\sigma_5\left(\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{C} \\ \hat{x} & \hat{y} & \mathbf{b}^T \end{bmatrix}\right) = \max\left\{\sigma_5\left(\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{C}\end{bmatrix}\right), \sigma_5\left(\begin{bmatrix} \mathbf{C} \\ \mathbf{b}^T \end{bmatrix}\right)\right\}.$$

Comparing this last equality with bound (7), we conclude that the pair $(\hat{x}, \hat{y})$, computed according to our SPOC sequential procedure, completes the matrix with the minimum possible value for the singular value σ_5, *i.e.*, it is the globally optimal completion. Naturally, the induction used in this example can be extended to show, in a similar way, that the global optimality of SPOC holds in general, see [11] for details.

4. EXPERIMENTS

We start by illustrating the capabilities of our SPOC method by processing the "Dinosaur" image sequence, see Fig. 2. This sequence, available from [12], has been used by several authors when addressing the recovery of rigid SFM under occlusion.
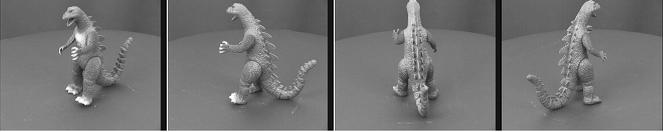

Fig. 2. Sample images from the "Dinosaur" sequence that illustrate the self-occlusion effect (obtained from [12]).

In the left plot of Fig. 3, we represent the image trajectories of 275 feature points in the dinosaur surface. Naturally, due to self-occlusion, several points are not seen in all the images available, thus the trajectories result highly incomplete. The image in the right side of Fig. 3 represents the pattern of known entries of the observation matrix collecting those trajectories. Note that we have (selected and) ordered the feature points in such a way that this pattern is a Young diagram.

We then used SPOC to compute the optimal completion of the observation matrix, *i.e.*, the completion that is closer to the set of possible trajectories of points in rigid body. The plots of Fig. 4 represent the result obtained. On the left, we show the extrapolated data, *i.e.*, the parts of the trajectories that were not visible in Fig. 3 due to occlusion. On the right plot, we represent the entire trajectories recovered from the completed observation matrix. Naturally, recovering SFM from these complete trajectories, which is trivial, *e.g.*, using the factorization method [1], leads to more accurate results than using any suboptimal strategy to process the incomplete set of trajectories observed in Fig. 3. The following example, for which we know the ground truth, demonstrates this point.
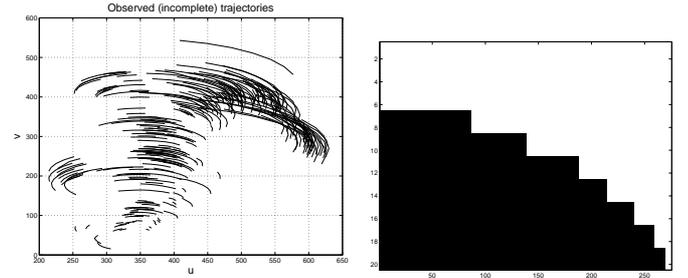

Fig. 3. Left: trajectories of the projections of 275 feature points tracked across the sequence in Fig. 2. Due to the dinosaur self-occlusion, the trajectories are highly incomplete. Right: binary mask of the observation matrix collecting these trajectories. The pattern is a Young diagram (observed entries are represented in white and missing ones in black).
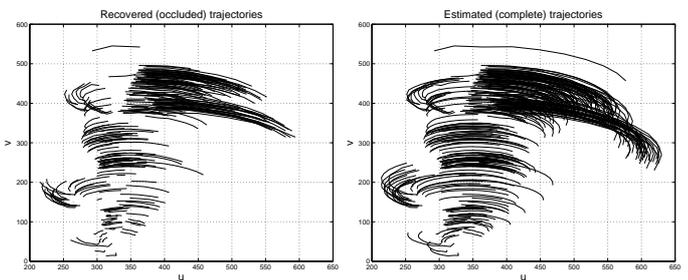

Fig. 4. Trajectories recovered by using our SPOC method to complete the observation matrix whose pattern of known entries is in Fig. 3. Left: estimates of the non-visible (occluded) parts of the trajectories. Right: complete trajectories, recovered from the ones partially observed in Fig. 3.

Fig. 5 illustrates the scenario. On the right plot, we show noisy (synthetic) trajectories, obtained by projecting in the image plane 100 points lying on the surface of the jar on the left side, for a path of 20 camera positions. To simulate occlusion, we have then removed several segments of these trajectories, obtaining an incomplete observation matrix (respecting the Young diagram pattern).

For comparison, we have processed the incomplete trajectories by using a standard approach: first, calibrate the cameras, *i.e.*, compute the camera positions across time, using the set of points whose trajectories are visible across the entire sequence; then, use the estimated camera positions to recover the entire 3D shape. The result obtained is illustrated in Fig. 6, together with the solution obtained by using our SPOC algorithm to complete the observation matrix. Naturally, since our method takes into account the object rigidity in a global way, we get more accurate estimates. This also happens for the estimates of the camera calibration parameters, see the examples plotted in Fig. 7.

Finally, we report the comparison of the re-projection errors of the two approaches above, in terms of their dependence on the observation noise level. We have performed 1000 runs, for each noise level, using partial observations of 50-frame trajectories, generated by a moving body whose shape was described by a set of 100 random points in 3D. The image coordinates were set in the interval $[-100, 100]$ and the noise standard deviation $\sigma \in [0, 2]$, in order to replicate situations that may arise when processing real-life videos.
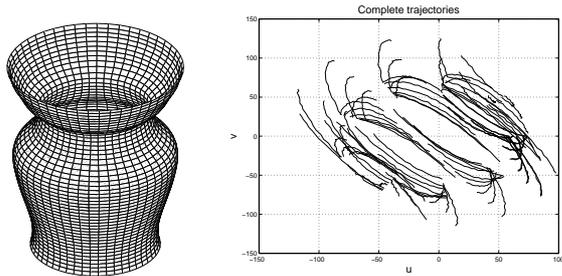
Fig. 5. Synthetic jar used for performance evaluation and noisy trajectories of projections of 100 points in the jar surface.
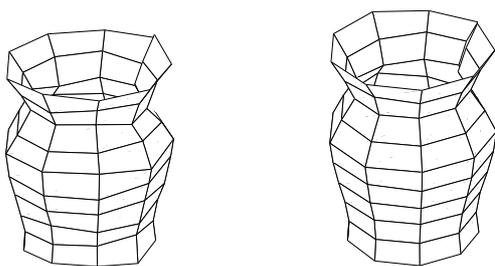

Fig. 6. Estimates of 3D shape computed from partial observations of the trajectories in Fig. 5. Left: result obtained by a standard calibration-first approach. Right: solution obtained by using SPOC to complete the trajectories in a globally optimal way.

We have repeated the experiment for several random shapes, motions, and patterns of missing data, obtaining always results similar to the one plotted in Fig. 8. Although for low levels of noise, both methods lead to similar re-projection errors, when the observation noise increases, the higher accuracy of the optimal solution (provided by SPOC) becomes evident.

5. CONCLUSION

We have presented a globally optimal method to exploit the scene rigidity when recovering SFM under occlusion. Our algorithm, which we call SPOC (from SPectrally Optimal Completion), computes the best completion of the observation matrix in a finite number of steps, without search, in a non-iterative way. Our method leads
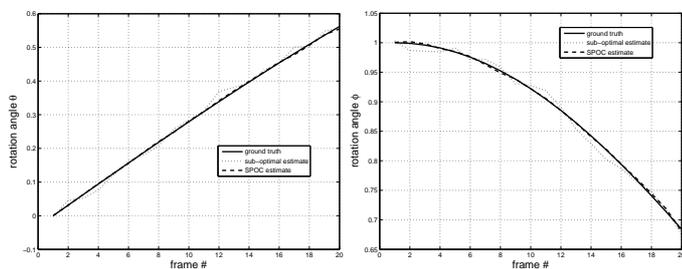

Fig. 7. Comparison of the accuracy of the recovery of 3D motion from partial observations of the trajectories in Fig. 5. The plots show the estimates of two of the parameters describing the 3D rotation of the camera across time.
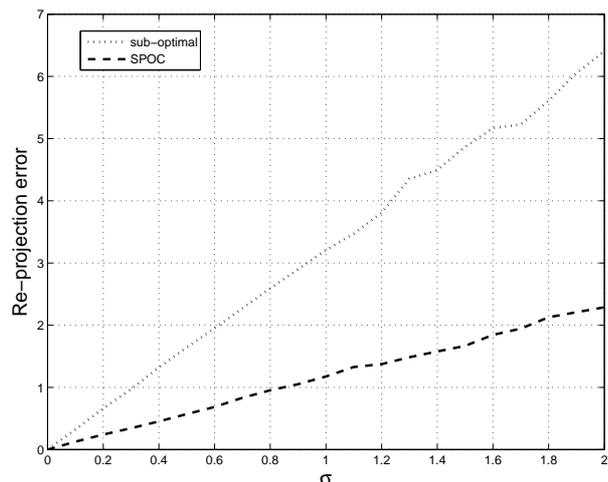

Fig. 8. Comparison of the re-projection errors of the suboptimal calibration-first approach and of our globally optimal method, as functions of the noise level.

to more accurate estimates of 3D shape and motion than standard calibration-first approaches. Although in the recent past several authors have proposed optimization algorithms to cope with missing data, these algorithms are greedy and often fall into local optima.

6. REFERENCES

[1] C. Tomasi and T. Kanade, "Shape and motion from image streams under orthography: a factorization method," *Int. Journal of Computer Vision*, vol. 9, no. 2, 1992.

[2] P. Aguiar and J. Moura, "Rank 1 weighted factorization for 3D structure recovery: Algorithms and performance analysis," *IEEE Trans. on PAMI*, vol. 25, no. 9, 2003.

[3] D. Jacobs, "Linear fitting with missing data: Applications to structure-from-motion and to characterizing intensity images," in *IEEE CVPR*, Puerto Rico, USA, 1997.

[4] R. Guerreiro and P. Aguiar, "3D structure from video streams with partially overlapping images," in *IEEE ICIP*, New York, USA, 2002.

[5] A. Buchanan and A. Fitzgibbon, "Damped newton algoritms for matrix factorization with missing data," in *IEEE CVPR*, San Diego CA, USA, 2005.

[6] T. Okatani and K. Deguchi, "On the Wiberg algorithm for matrix factorization in the presence of missing components," *Int. Journal of Computer Vision*, vol. 72, no. 3, 2007.

[7] P. Aguiar and J. Xavier and M. Stosic, "Spectrally optimal factorization of incomplete matrices," in *IEEE CVPR*, Anchorage, Alaska, USA, 2008.

[8] W. Freeman and J. Tenenbaum, "Learning bilinear models for two-factor problems in vision," in *IEEE CVPR*, Puerto Rico, USA, 1997.

[9] S. Thrun, "Affine structure from sound," in *NIPS 18*. MIT Press, Cambridge, MA, 2006.

[10] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.

[11] P. Aguiar, M. Stošić, and J. Xavier, "On singular values of partially prescribed matrices," Submmitted to *ELSEVIER Linear Algebra and its Applications*, 2007.

[12] University of Oxford – Visual Geometry Group, http://www.robots.ox.ac.uk/~vgg/data.