

Localizing Non-Overlapping Surveillance Cameras under the L-Infinity Norm*

Branislav Micusik

Roman Pflugfelder

Safety and Security Department
AIT Austrian Institute of Technology

Abstract

This paper presents a new approach to the problem of camera localization with non-overlapping camera views, particularly relevant for video surveillance systems. We show how to recast localization as quasi-convex optimization under the L-Infinity norm. Thereby we add the problem of reconstructing camera centers and 3D points for non-overlapping cameras with known internal parameters and known rotations to the class of known geometric problems solvable with Second Order Cone Programming. The 3D points are never seen by more than one camera, which makes the localization problem ill-posed. Therefore, the proposed approach employs temporal consistency of the 3D points to supply the missing constraints. Our formulation allows a global optimal solution to be found with a clear physical meaning of the cost function being minimized.

1. Introduction

An important goal of automated video surveillance systems is event recognition in large environments, e.g. airports or train stations, which in the last years have seen a growth in the number of cameras installed. To make event recognition on such a large scale feasible, camera synchronization and knowledge about the spatial camera arrangement is clearly a pre-requisite. Driven by cost, cameras are sparsely mounted so that visual gaps are oftentimes present without any image point correspondences across the camera views, see Fig. 1.

Due to a lack of measurement correspondences, one is confronted with missing constraints for estimating the multi-view geometry [8], which makes neither manual nor automated state-of-the-art calibration methods applicable. Seeing the 3D points as points of a smooth object's trajectory, one can assume temporal consistency, which supplies the missing constraints to bridge the visual gaps and to solve the localization problem [3].

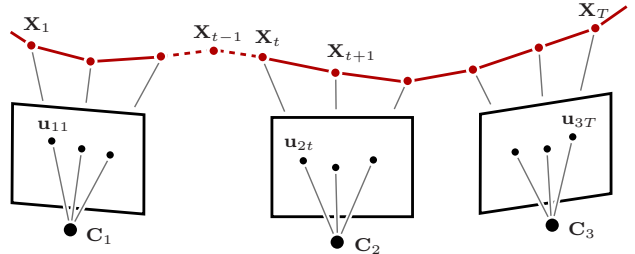


Figure 1. Multi-view geometry of non-overlapping surveillance cameras. The paper addresses the localization of the camera centers C_i , and 3D points X_t , given the measurements u_{it} .

We aim at reconstructing camera centers and trajectories in non-overlapping views from a set of measurements given by (i) synchronized, and calibrated cameras with known rotations and (ii) known trajectory correspondences within and among the views. This reconstruction problem is practically very interesting as the rotations can be obtained e.g. from vanishing points, pure translation, or other sensors. However, measuring the camera centers is difficult to realize, therefore, their automatic localization is much more relevant.

Assuming independent Gaussian noise, a least-squares formulation on reprojection error and temporal consistency under the L_2 norm provides the maximum likelihood estimate which is statistically the optimal solution for this problem. However, the least-squares formulation leads to a non-linear optimization problem with many local optima. In general, to overcome this problem, linear algorithms minimizing an algebraic error are used to initialize the non-linear bundle adjustment. These linear algorithms are also efficiently implementable, allowing them to be wrapped into robust frameworks, e.g. RANSAC. We show that a recently proposed linear algorithm [13] augmenting the Direct Reference Plane (DRP) method [15] for the problem being solved here has clear shortcomings in certain situations. Moreover, the DRP often delivers a physically infeasible solution in our problem.

The contribution of this paper is in an alternative problem formulation under the L_∞ norm via the Second Order Cone Programming (SOCP). In contrast to the DRP, the pro-

*This research received funding from Wiener Wissenschafts-, Forschungs- und Technologiefonds - WWTF, Project No ICT08-030.

posed formulation uses a softer constraint on temporal consistency, expressed as a bounded change in the trajectory’s local velocity. Additionally, SOCP treats cheirality of the trajectory points. SOCP has captured attention in recent years, as many important geometric problems can be formulated under the L_∞ norm by an iterative, globally optimal binary search through smaller quasi-convex minimizations of the physically meaningful geometric errors [7, 10]. This paper extends the class of the SOCP solvable problems under the L_∞ norm by the problem of localization of calibrated non-overlapping cameras given their rotations and trajectory observations.

We provide an extensive comparison of the SOCP method under the L_∞ norm and the DRP under the L_2 norm, and explore their limitations and advantages. We point out the issues of the DRP method in non-overlapping scenario which has not been recognized yet. In all our experiments, the proposed SOCP method provides a significantly better localization than the DRP method. If the maximum likelihood is needed, optimizing the reprojection errors and temporal consistency in 3D by bundle adjustment can still be applied as post-processing step, but we show that SOCP would bring a far better initialization than DRP. However, our results indicate that the accuracy obtained by SOCP is sufficient for video surveillance applications without further bundle adjustment.

2. Related work

Alternative approaches to localizing non-overlapping cameras have appeared in the past. Linear extrapolation of trajectories into the invisible image that hallucinates the missing correspondences has been proposed for uncalibrated [9] and calibrated cameras [5]. In the former work, correspondences are used to compute ground plane induced homographies, in the latter work, correspondences are used to localize only the camera centers, assuming known camera rotations. Attempts to use polynomials for extrapolation have recently been made [1].

A related approach [11, 12] estimates camera topology in terms of a graph where the vertices express the cameras and the edges express spatial proximity. Topology contains less information than a geometric camera model, however, topology can be a prior for camera localization [4].

Our work is mostly relevant to approaches assuming local temporal consistency along the trajectory, *e.g.*, nonlinear optimization on the ground plane [14, 16] with obvious problems of setting starting points. A recently proposed augmented DRP method [13] also falls under this approach but has clear shortcomings, discussed in more details in Sec. 3.1, hence, the remainder of this paper formulates then the problem differently in the framework based on the L_∞ norm in Sec. 3.2. Sec. 3.3 studies the main advantages of our formulation confirmed by the experiments in Sec. 4.

3. Concept

Let us assume a 3D point \mathbf{X} is observed by N cameras at T constant time instances on its trajectory, see Fig. 1. The formulation holds for arbitrary number of trajectories but, for simplicity of notation, we consider only one trajectory. The standard camera perspective equation [8] reads as

$$\lambda_{it} \mathbf{u}_{it} = K_i [R_i \mathbf{t}_i] (\mathbf{X}_t^\top \mathbf{1})^\top, \quad (1)$$

where we assume known calibration matrices K_i and rotations R_i .

3.1. DRP with least squares

By re-arranging Eq. (1) one can write

$$\lambda_{it} \mathbf{d}_{it} = \mathbf{X}_t - \mathbf{C}_i, \quad (2)$$

with $\mathbf{C}_i = -R_i^\top \mathbf{t}_i$ as the camera center and $\mathbf{d}_{it} = R_i^\top K_i^{-1} \mathbf{u}_{it}$ as the ray passing through \mathbf{C}_i and the image measurement \mathbf{u}_{it} . Enforcing same directions of both vectors,

$$\|\mathbf{d}_{it} \times (\mathbf{X}_t - \mathbf{C}_i)\|_2 = E_{dat}(i, t) \rightarrow 0, \quad (3)$$

eliminates λ_{it} and defines an error called the data term. This equation leads to the triangulation problem, *i.e.* observing 3D points by at least two cameras, one can estimate their locations and simultaneously localize the cameras by minimizing the error $E_{dat}(i, t)$ in Eq. (3) in a closed form [15].

In the non-overlapping case the problem is ill-posed with not enough constraints. However, the smoothness temporal consistency of the 3D points can complement for that by enforcing the following

$$\|\mathbf{X}_{t-1} - 2\mathbf{X}_t + \mathbf{X}_{t+1}\|_2 = E_{sm}(t) \rightarrow 0, \quad (4)$$

called further the smoothness term. Concatenating the data and smoothness term yields to the linear system $\mathbf{Bx} = \mathbf{0}$ with \mathbf{x} containing all unknown variables [13]. When using SVD in the overconstrained case, $\|\mathbf{Bx}\|_2$ is minimized in the least square sense subject to $\|\mathbf{x}\|_2 = 1$ where

$$\|\mathbf{Bx}\|_2^2 = \sum_{t=1}^T \sum_{i=1}^N E_{dat}(i, t)^2 + \sum_{t=2}^{T-1} E_{sm}(t)^2. \quad (5)$$

Both terms in Eq. (5) have a clear geometric meaning, however not related to each other, therefore making it difficult to properly weight them. The data term corresponds to the area of the quadrilateral with \mathbf{d}_{it} and $(\mathbf{X}_t - \mathbf{C}_i)$ as its sides, see Eq. (3). Optimally, with noise free data, the vectors are collinear hence $E_{dat} = 0$. Clearly, if the reconstructed 3D point \mathbf{X}_t is closer to the camera center \mathbf{C}_i , the area is smaller. If the 3D point is seen only by one camera, there is not cue coming from the triangulation and therefore the data term forces the points to be as close as possible to

the center of the camera observing that point. The smoothness term measures deviation of the reconstructed trajectory from a line. Therefore, the optimal solution for Eq. (5) is to keep the reconstructed trajectory as much linear and as much close to the camera centers as possible. If there are no points seen at least by two cameras this effect is an important issue, not recognized in [13].

Another problem is that the cheirality is implicitly fulfilled when doing triangulation [15], but not in the case considered here when the points are seen only by one camera. It means that the DRP formulation allows points to get reconstructed behind the cameras if such a solution is cheaper for the smoothness term, as demonstrated later by our experiments. This significant problem has been overseen in [13].

3.2. SOCP with L_∞

To overcome the aforementioned issues we propose another solution. We rewrite Eq. (1) to

$$\lambda_{it} \mathbf{u}_{it} = [\mathbf{M}_i \mathbf{p}_i] (\mathbf{X}_t^\top \mathbf{1})^\top, \quad (6)$$

where $\mathbf{M}_i = \mathbf{K}_i \mathbf{R}_i$ and $\mathbf{p}_i = \mathbf{K}_i \mathbf{t}_i$. Let $\mathbf{m}^1, \mathbf{m}^2, \mathbf{m}^3$ corresponds to the rows of the matrix \mathbf{M} . Reprojection error of a 3D point \mathbf{X}_t in i th camera then reads as

$$\gamma_{it} = \left\| \frac{\mathbf{m}_i^{1\top} \mathbf{X}_t + p_i^1}{\mathbf{m}_i^{3\top} \mathbf{X}_t + p_i^3} - u_{it}^1; \frac{\mathbf{m}_i^{2\top} \mathbf{X}_t + p_i^2}{\mathbf{m}_i^{3\top} \mathbf{X}_t + p_i^3} - u_{it}^2 \right\|_2 \quad (7)$$

which becomes

$$\left\| \frac{\mathbf{m}_i^{1\top} \mathbf{X}_t + p_i^1 - u_{it}^1 (\mathbf{m}_i^{3\top} \mathbf{X}_t + p_i^3)}{\mathbf{m}_i^{2\top} \mathbf{X}_t + p_i^2 - u_{it}^2 (\mathbf{m}_i^{3\top} \mathbf{X}_t + p_i^3)} \right\|_2 = \gamma_{it} |\mathbf{m}_i^{3\top} \mathbf{X}_t + p_i^3|. \quad (8)$$

We can avoid the absolute value on the right hand side of Eq. (8), since the cheirality condition $\mathbf{m}_i^{3\top} \mathbf{X}_t + p_i^3 > 0$ holds for all \mathbf{X}_t . This comes from the fact that the points must physically lie in front of the cameras observing them. Then, one can rewrite Eq. (8) to

$$\|\mathbf{A}_{it} \mathbf{x} + \mathbf{b}_{it}\|_2 = \gamma_{it} (\mathbf{c}_{it}^\top \mathbf{x} + d_{it}) \quad (9)$$

with known $\mathbf{A}_{it}, \mathbf{b}_{it}, \mathbf{c}_{it}$ and d_{it} , and a vector \mathbf{x} of unknown variables

$$\mathbf{x} = (\mathbf{X}_2^\top \cdots \mathbf{X}_T^\top \tilde{\mathbf{p}}_1^\top \mathbf{p}_2^\top \cdots \mathbf{p}_N^\top)^\top \quad (10)$$

and unknown γ_{it} . We set $\mathbf{X}_1 = (0 \ 0 \ 0)^\top$ and $\mathbf{p}_1 = (\tilde{\mathbf{p}}_1^\top \mathbf{1})^\top$ to fix the scale of the resulting reconstruction. Note that each projection defines a conical surface by Eq. (9).

Treating the projective depth λ_{it} correctly, and not eliminating it as done by DRP in Eq. (3), brings a non-linearity into the problem. The reprojection error for one point in Eq. (7) is a non-linear but quasi-convex function [10]. However, when summing squares of reprojections errors for all projected points, *i.e.* many quasi-convex functions, as done

in the standard bundle adjustment, one gets a multi-modal residual function with many local minima. To prevent such a sum, and keep the problem quasi-convex with the guaranty of a global minimum, the use of L_∞ norm was proposed [7]. The reprojection error in Eq. (7) under the L_∞ norm is defined as $\gamma = \max_{i,t} \gamma_{it}$ and the problem of searching for unknown variables reads as

$$\mathbf{x}^* = \underset{\mathbf{x}}{\operatorname{argmin}} \gamma = \underset{\mathbf{x}}{\operatorname{argmin}} \max_{i,t} \gamma_{it}. \quad (11)$$

It can be transformed to the equivalent problem, called the SOCP feasibility problem (SOCPF) [2],

$$\begin{aligned} & \text{find} && \mathbf{x} \\ & \text{s.t.} && \|\mathbf{A}_{it} \mathbf{x} + \mathbf{b}_{it}\|_2 \leq \gamma (\mathbf{c}_{it}^\top \mathbf{x} + d_{it}) \end{aligned} \quad (12)$$

given $\gamma \geq 0$. We do not know γ , however, for each guess of it, we solve a convex optimization problem, the Second Order Cone Program, with a guaranteed global optimum. *E.g.* a bisectioning algorithm can be used to search for γ [10].

Each image measurement adds one conical constraint in Eq. (12). It means that, to get the SOCPF problem solvable, each 3D point must be seen at least by two cameras. Our non-overlapping setup yields therefore to an ill-posed problem. To add the missing constraints we enforce the following requirement on the 3D trajectory

$$\|\mathbf{X}_{t-1} - 2\mathbf{X}_t + \mathbf{X}_{t+1}\|_2 \leq \alpha. \quad (13)$$

It defines a second order cone and therefore it can be rewritten in the form of the SOCP constraint in Eq. (12). We investigate the properties of such a constraint with practical implications in the following section.

By adding temporal smoothness term as a cone into the SOCP, we introduce another unknown parameter α . However, since we search for the minimal L_∞ norm of the reprojection error we can decrease the α until the problem in Eq. (12) is feasible and when it allows lower γ . We therefore iterate for α , and for each fixed guess of α we search for γ , until the minimum of maximal reprojection error is reached, Eq. (11). For both parameters we use a bisectioning algorithm [10]. Both, α and γ , have a clear geometric meaning. The parameter γ controls the maximal reprojection error in all the images while α controls the maximal deviation of the trajectory in 3D from being linear.

3.3. The smoothness term as constraint

According to Eq. (7) a 3D point can be swept along its corresponding projective ray while the reprojection error remains the same. It means that there is a family of solutions for the 3D points seen only by one camera. The only constraint which can prevent the random shifts along the projective rays and stabilizes the solution is the smoothness term. We added this term into the SOCPF formulation in

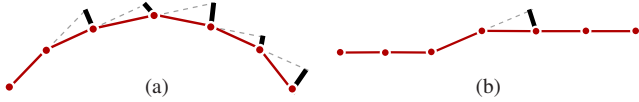


Figure 2. Smoothness term penalty for 3D points of two different trajectories. The thick black lines correspond to the cost paid for each triplet of the 3D points. Under the L_∞ norm, (a), (b) have the same cost unlike for the least squares where (b) is less expensive.

Eq. (12) as an additional constraint of Eq. (13) in contrast to the DRP. The smoothness in the SOCP is defined as an upper bound while the DRP minimizes the smoothness term together with the data term in least square sense in Eq. (5). It has important practical consequence with significant influence on camera localization and reconstructed trajectory.

The least square fitting done by DRP favors linearity of the trajectory. When the trajectory is pairwise linear with turns or loops, clearly, the smoothness term strongly penalizes it trying to find the most linear, although incorrect, fit. Formulating the smoothness by the L_∞ norm allows much higher freedom. There is a family of trajectories within the given bound that have the same cost under the L_∞ norm. They may have different shape but all are equally likely as shown in Fig. 2 and give the same camera centers. We are primarily interested in estimating camera centers therefore the reconstructed trajectory does not necessarily have to correspond to the real one (as opposed to the classical SfM). Possibly, Eq. (13) can be encoded into the SOCP as a criterion function through slack variables and minimized for a fixed γ similarly to DRP. However, we have observed that enforcing the smoothness as a constraint is for handling real non straight trajectories much better than its direct minimization.

In case of smooth, but not linear trajectory with many loops, *e.g.*, when a laser pointer or person is tracked as it freely moves across the field of views of all the cameras, the smoothness as a bound is a better choice. On the other hand, if we knew that an object moves with a constant velocity along a line with small deviations, but no turns, then the least square smoothness by the DRP would be a better approximation. It can suppress the noise in image measurements or tiny off-line shifts like in Fig. 2b better. We demonstrate this later in Sec. 4.

Outliers are problematic for both methods. The SOCP fits the noisiest point or the outlier and then allows too high freedom to other points. However, in surveillance the feature points are projections of a smoothly moving 3D point, therefore a projected trajectory in the image must be smooth as well. Outlier detection reduces then to a robust 2D spline fitting. The outliers can thus be detected in an early stage and simply removed before running SOCP since missing measurements are treated by completion by the imposed smoothness term. The outliers along projective rays do not play role unless they are seen by more than one camera,

	RMS C [m]	SETUP A	SETUP B	SETUP C
<i>SOCP</i> without 3D noise	0.58 (0.43)	0.13 (0.07)	0.80 (0.27)	
<i>SOCP</i> with 3D noise	1.36 (1.05)	0.14 (0.05)	1.67 (0.20)	
<i>DRP</i> without 3D noise	4.19 (0.04)	0.30 (0.01)	6.40 (0.34)	
<i>DRP</i> with 3D noise	4.31 (0.04)	0.47 (0.02)	10.21 (0.37)	

Table 1. Synthetic experiment. Mean and standard deviation (in brackets) of RMS error of localizing the camera centers C_i w.r.t. the ground truth at the noise level $\sigma = 0.5\text{pxl}$.

then [17] is applicable.

4. Experiments

We carried out experiments with both synthetic and real data to show performance of the proposed SOCP solution and give a comparison to the DRP method [13] for typical setups. For solving the SOCP we use the publicly available toolbox SeDuMi [18].

4.1. Synthetic data

We simulate a scenario of three cameras positioned as shown in Fig. 3. We assume the camera resolutions are of 640×480 pxl, known calibration matrices and camera rotations. The trajectory of a moving object is an elliptical spiral to capture non-constant motion curvature and velocity. In each camera we have a measurement at equal time instances, approximately every 1 m on the 3D trajectory. Measurement matrix in Fig. 3c shows that there is no point correspondence between the cameras at all.

To investigate the behavior of both methods, we add Gaussian noise with standard deviation σ to the image measurements, repeat that 50 times at each σ and plot the error bars. Then we repeat the same but add uniform noise with standard deviation of 0.2 m to the 3D points of the ground truth trajectory, referred as 3D noise, to simulate an inaccurate camera synchronization or smooth motion violations. Note that the noise can shift each point by 20% from its initial position.

We investigate three cases. *SETUP A*, where we use all the three cameras, *SETUP B*, where only two cameras looking in opposite directions are considered, and *SETUP C*, the most complicated case, consisting of two cameras looking in the same direction but having a 4 m gap in which the trajectory is invisible. For each setup and noise level we evaluate the camera localization error which is computed as follows. First, the gauge of the estimate and the ground truth is fixed by aligning the first camera and by scaling the points and camera centers such that the second cameras have equal x coordinate. The localization error is then computed as RMS error on distances between estimated and ground truth camera centers.

We show an example of one localization result per setup, Fig. 3a, noise influence characteristics in Fig. 3b and sum-

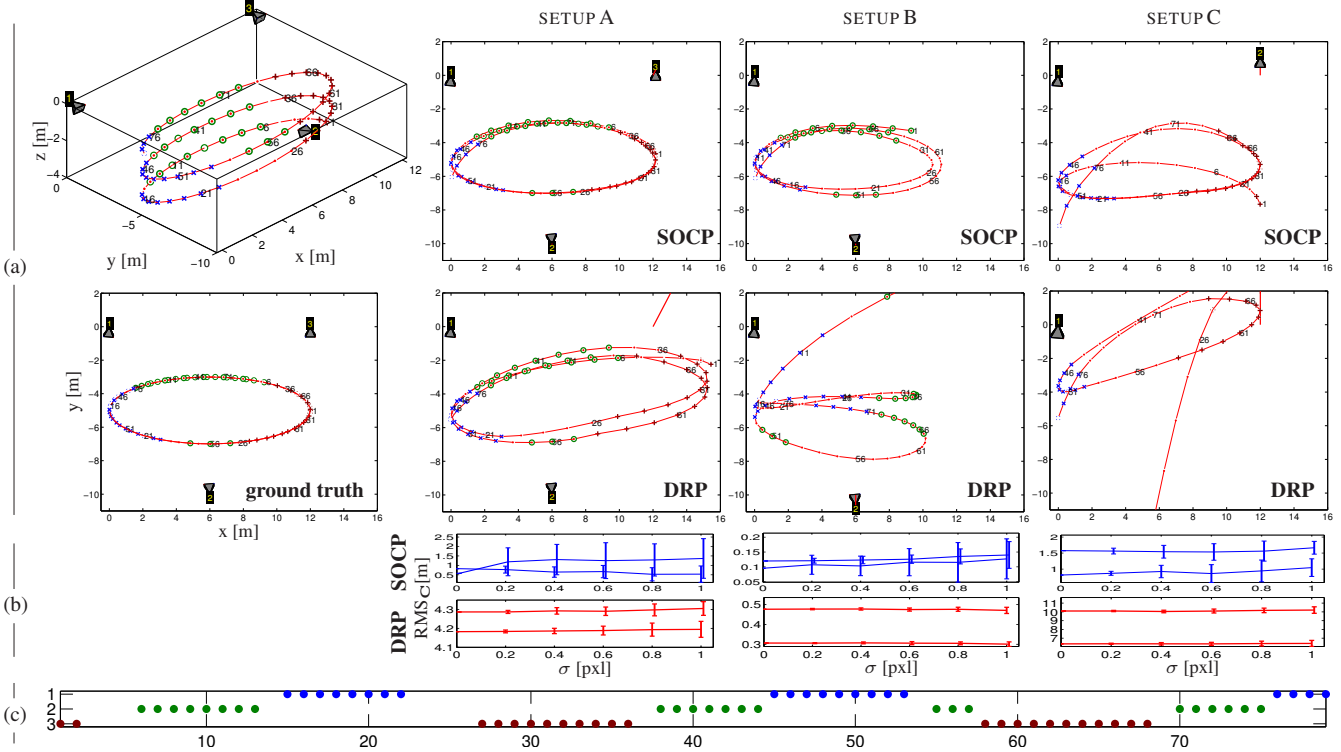


Figure 3. Synthetic experiment. (a) 1st column: Setup with cameras looking at an elliptical spiral trajectory and its top view. 2nd, 3rd, and 4th columns: An example of camera localization with Gaussian noise of $\sigma = 0.2$ for three different setups with SOCP and DRP method. (b) RMS error on camera localization vs. noise level for all three setups. The curves from the top correspond to SOCP with, SOCP w/out, DRP with, and DRP w/out the 3D noise. (c) Measurement matrix where rows correspond to camera indices and columns to time instances.

mary in Tab. 1. We see that the SOCP method significantly outperforms the DRP method and always provides better estimates in all the cases. In SETUP A, and C the DRP method gives implausible results while SOCP is still within accuracy of 80 cm. In SETUP B, DRP performs well, it provides an accuracy comparable to what reported in [13], about 30 cm for similar configuration. SOCP performs better also here, around 13 cm. The next real experiments corresponding to the same setup achieved comparable accuracy, validating the outcome of the other two setups.

4.2. Real data

This section evaluates the methods on the image sequences of a room scene presented in [13] with two real calibrated, synchronized cameras, arranged accordingly to SETUP B, whereas baseline and the borders of the fields of view coincide. We established the ground truth in a metric coordinate frame by using provided data and marking manually the epipoles which are mutually visible on the image border in both camera views.

Head-shoulder experiment: This experiment shows a walking person whose head-shoulder part was automatically located by [6] with some missing detections and localization error within some pixels. Fig. 4 shows the results

of the DRP and SOCP method, with measurement matrix indicating missing observations, and a short subsequence. The localization error for the SOCP is 8 cm and for DRP 33 cm, which is significantly higher. Trajectory reconstructions in Fig. 4 are both almost planar which agrees with the person’s movement.

To compare DRP against SOCP in case of straight trajectories without any turns in one of the camera views, we selected close to the image borders particular detections forming two straight tracks. The localization error increased slightly to 54 cm for the DRP, but increased significantly to 74 cm for the SOCP. This confirms our expectation that DRP finds tighter solutions than SOCP in case of real straight trajectories. The turns and the time duration of the person’s invisibility capture valuable depth information and improve the reconstruction of camera centers and trajectory for both methods. However, we see that the SOCP gains from loops much more than the DRP.

LED light experiment: This second experiment shows a bright, moving LED light, detectable by subpixel accuracy [19], carried by a walking person under shaded room conditions. The LED light movements are non-planar, containing abrupt changes in direction and accelerations. The estimated camera and trajectory locations are shown in

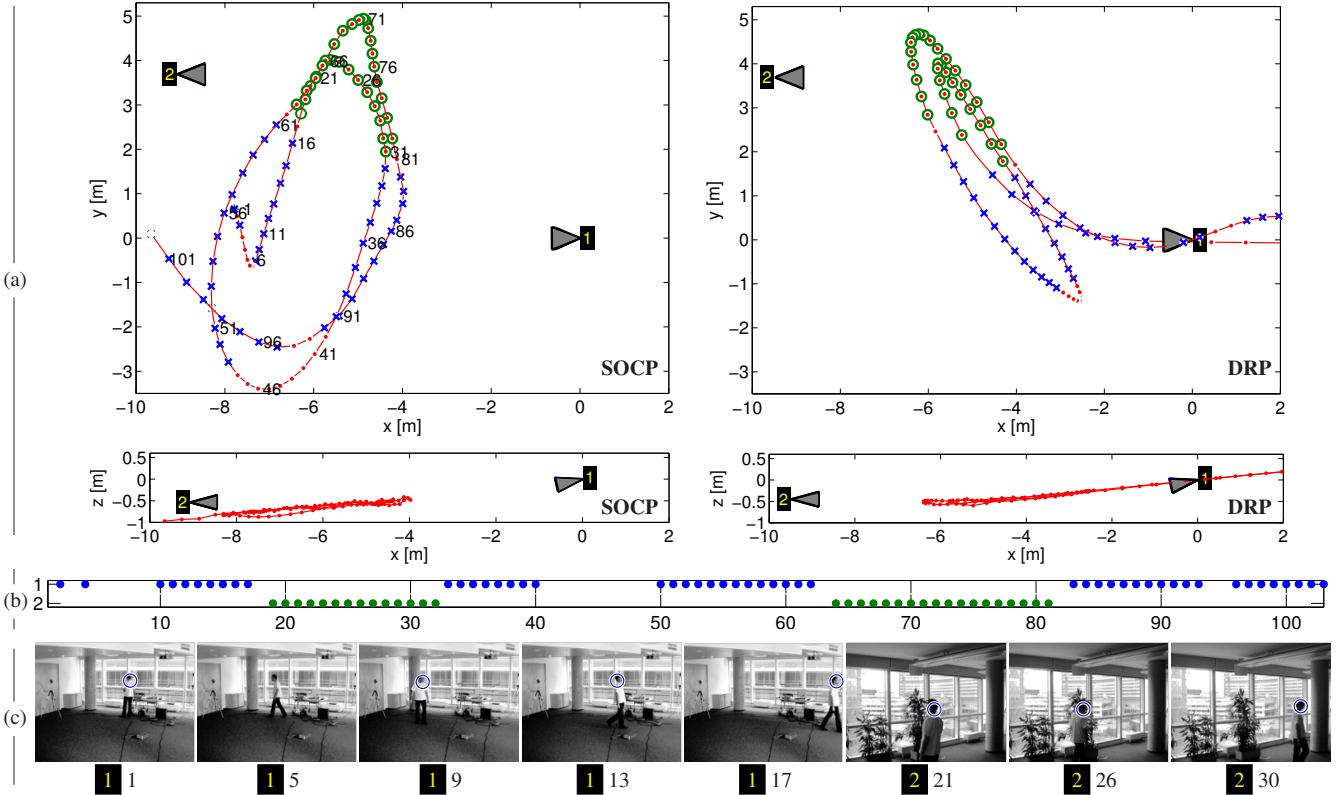


Figure 4. Head-shoulder experiment. (a) Bird's eye and x - z views of the reconstructed camera centers and trajectory. 3D points seen by the camera 1 / camera 2 are shown in blue crosses / green circles while small red dots are the unseen points. (b) Measurement matrix. (c) Image frames and measurements for a short subsequence.

Fig. 5, localization error for SOCP is 6 cm, for DRP 17 cm, which is again higher. However, both methods perform better than in the head-shoulder detector experiment thanks to the sub-pixel accurate detection of the LED light. The quality of trajectory reconstruction is astonishing under the LED light's difficult dynamics, shown by the V-shaped motion shown in Fig. 5c.

Remarks The real data experiments show that SOCP outperforms the DRP significantly, verified by visual check of the topology of the reconstructed trajectories with the video stream and the magnitude of the localization error. The DRP reconstructs some points behind the camera no.1 in both cases, which is physically impossible and demonstrates clearly the cheirality violation, pointed out as a problem of the DRP in Sec. 3.1. To handle the cheirality in the DRP correctly, one could possibly include its cost function into the linear program with inequalities on all projective depths to be positive. However, the objective function of the DRP has still clear disadvantages.

In both experiments we measured RMS reprojection and L_∞ errors. The SOCP scores with negligible values for both and DRP on average about 7 pxl for the RMS of reprojection errors and 56 pxl for the L_∞ error. These bounds are

a direct consequence of missing point correspondences and of DRP's inability to minimize the reprojection error. The SOCP provides thus a much better starting point for further possible bundle adjustment.

The results confirm our expectations in the SOCP and show the limits of DRP's problem formulation with trajectories containing turns or loops. When judging the camera localization accuracy, one must take into account the accuracy of the detectors (several pixels) and the varying motion velocity and therefore, the achieved localization error is within the expected bounds.

5. Conclusion

We proposed a method for localizing non-overlapping calibrated cameras given their rotations based on SOCP minimizing the L_∞ norm with enforced temporally consistent dynamic points. We demonstrated that SOCP provides very plausible results and accuracy for the shown experiments which are representative examples of video surveillance scenarios. We pointed out clear advantages of the SOCP formulation to the existing DRP method for typical trajectories. This result promises beneficial combination of camera localization from detection and tracking.

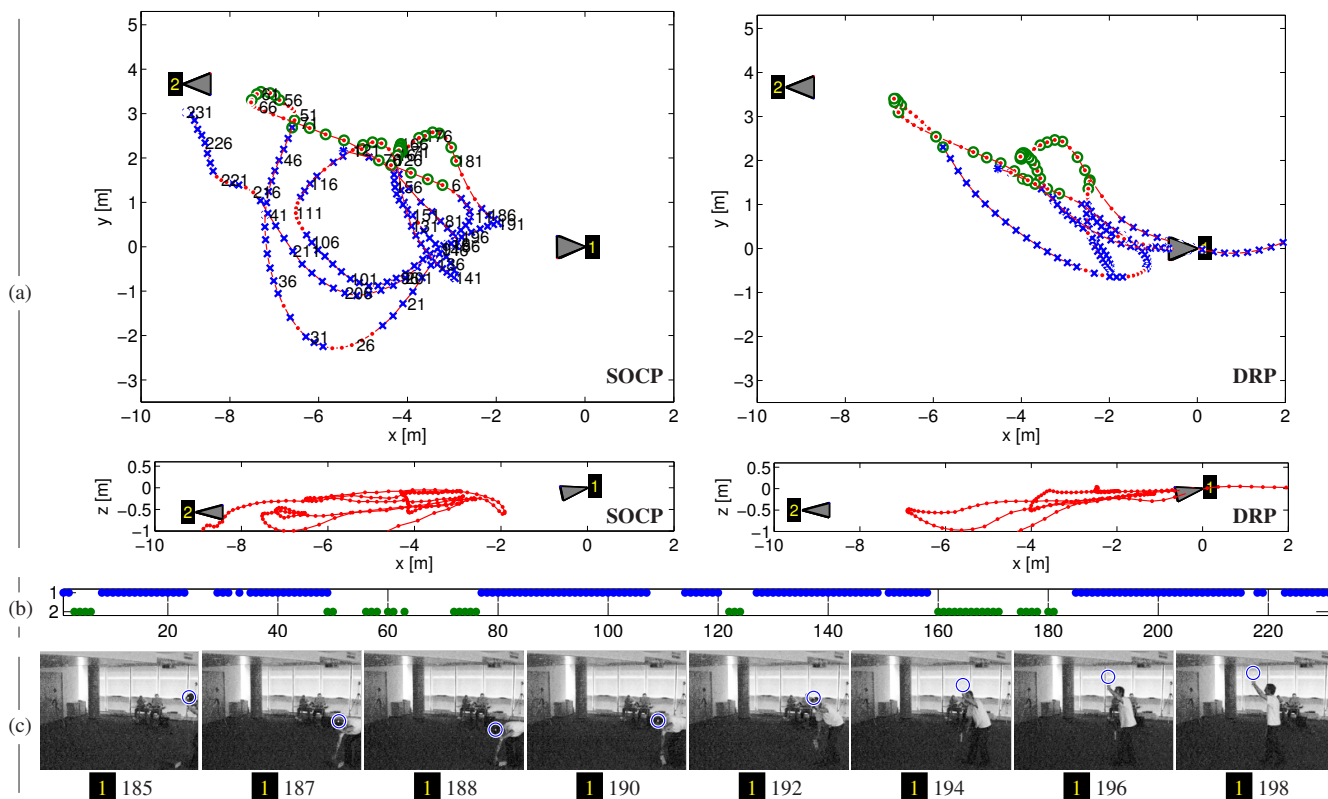


Figure 5. LED light experiment. (a) Bird's eye and x - z views of the reconstructed camera centers and trajectory. (b) Measurement matrix. (c) Image frames and measurements for a V-shaped vertical motion subsequence.

References

- [1] N. Anjum, M. Taj, and A. Cavallaro. Relative position estimation of non-overlapping cameras. In *IEEE Intern. Conf. on Acoustics, Speech and Signal Processing*, 2007.
- [2] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [3] Y. Caspi and M. Irani. Aligning non-overlapping sequences. *IJCV*, 48(1):39–51, 2002.
- [4] D. Devarajan, Z. Cheng, and R. Radke. Calibrating distributed camera networks. *Proceedings of the IEEE*, 96(10):1625–1639, 2008.
- [5] R. B. Fisher. Self-organization of randomly placed sensors. In *ECCV*, pages IV: 146–160. Springer-Verlag, 2002.
- [6] H. Grabner, H. Bischof. On-line boosting and vision. In *CVPR*, pages (I):260–267, 2006.
- [7] R. Hartley and F. Schaffalitzky. L_∞ minimization in geometric reconstruction problems. In *CVPR*, 2004.
- [8] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, second edition, 2004.
- [9] O. Javed, Z. Rasheed, O. Alatas, and M. Shah. M-knight: A real time surveillance system for multiple and non-overlapping cameras. In *IEEE Intern. conf. on Multimedia and Expo*, 2003.
- [10] F. Kahl and R. Hartley. Multiple view geometry under the L-infinity norm. *PAMI*, 30(9):1603–1617, 2008.
- [11] D. Makris, T. Ellis, and J. Black. Bridging the gaps between cameras. In *CVPR*, pages 205–210, 2004.
- [12] D. Marinakis and G. Dudek. Self-calibration of a vision-based sensor network. *Image and Vision Computing*, 27(1-2):116–130, 2009.
- [13] R. Pflugfelder and H. Bischof. Localization and trajectory reconstruction in surveillance cameras with non-overlapping views. *PAMI*, 2009.
- [14] A. Rahimi, B. Dunagan, and T. Darrell. Simultaneous calibration and tracking with a network of non-overlapping sensors. In *CVPR*, pages (I):187–194, 2004.
- [15] C. Rother and S. Carlsson. Multi view reconstruction and camera recovery. In *ICCV*, 2001.
- [16] M. Rudoy and C. E. Rohrs. Simultaneous sensor calibration and path estimation. In *IEEE Asilomar Conf. on Signals, Systems, and Computers*, 2006.
- [17] K. Sim and R. Hartley. Removing outliers using the L-infinity norm. *CVPR*, pages I:485–494, 2006.
- [18] J. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11–12:625–653, 1999.
- [19] T. Svoboda, D. Martinec, and T. Pajdla. A convenient multi-camera self-calibration for virtual environments. *PRES-ENCE: Teleoperators and Virtual Environments*, 14(4):407–422, 2005.