Autocalibrage de Caméras Robuste et quasi-Optimal

Lijia Gao[1] Adlane Habed[1] Sandrine Voros[2] Christophe Doignon[1]

[1] ICube (UMR CNRS 7357), Université de Strasbourg
[2] TIMC-IMAG (UMR CNRS 5525), Université Grenoble Alpes *

l.gao@unistra.fr, habed@unistra.fr, c.doignon@unistra.fr
sandrine.voros@univ-grenoble-alpes.fr

Résumé

Nous proposons un algorithme d'autocalibrage de caméras robuste et globalement optimal, fondé sur l'estimation de la "Absolute Quadratic Complex" (AQC). Dans ce travail, nous considérons un ensemble de caméras perspectives à paramètres intrinsèques variables mais avec des rapports d'aspect connus et une absence d'asymétries (ou inclinaison) des axes, des hypothèses très raisonnables avec les caméras récentes. Le problème d'estimation est alors formulé sous la forme d'un problème de minimisation impliquant des polynômes rationnels. Ce problème est résolu ici via la stratégie "Branch-and-Prune" (BnP) et une programmation semi-définie. L'objectif atteint par l'AQC estimée est assuré de se situer dans une tolérance définie par l'utilisateur (ou un critère d'optimalité) par rapport à l'objectif fourni par la solution globalement optimale. Notre algorithme est déterministe et présente une complexité de mémoire constante. De plus, les résultats des expériences menées avec des données simulées et réelles révèlent que l'algorithme est numériquement très stable et robuste face à des niveaux importants de bruit dans les images. Plus important encore, contrairement à toute autre méthode d'autocalibrage, lorsque des images avec des données aberrantes sont présentes, notre méthode fournit encore de bons résultats.

Mots Clef

Vision et géométrie, Auto-étalonnage, Optimisation.

Abstract

We propose a robust and globally optimal camera autocalibration algorithm based on the so-called Absolute Quadratic Complex (AQC) formulation of the problem. Our working assumption is that of a set of perspective cameras, with varying intrinsic parameters, some exhibiting no skew and known aspect ratio (inliers) and some not (outliers). The problem of estimating the AQC is formulated as a median-residual minimization problem involving rational polynomials. This problem is solved via a Branch-and-Prune strategy and semidefinite programming. The objective reached by the estimated AQC is guaranteed to be within a user-defined tolerance (or optimality criterion) from the objective provided by the globally optimal solution. Our algorithm is deterministic and exhibits a constant memory complexity. Furthermore, the results of the extensive experiments conducted with simulated and real data reveal that the algorithm is numerically very stable and robust against significant levels of image noise. Most importantly, unlike any other autocalibration method, when outlier images are present, our method yields equally good results.

Keywords

Computer vision geometry, Autocalibration, Optimization.

1 Introduction

The problem of recovering the metric structure of an unknown scene from images captured by an uncalibrated camera has abundantly been addressed in the literature. The problem has long stumbled upon the challenging task of retrieving the camera's intrinsic parameters from image feature correspondences, i.e. camera autocalibration.

Existing camera autocalibration methods assume the images are captured by cameras sharing some common internal geometric properties implying either fully constant [10, 16, 25, 21, 5, 4, 2, 7, 8] or partially variable [11, 22, 17, 3, 23, 26, 6] intrinsic parameters. Such assumption is particularly viable when the images are captured by the same moving camera. In such case, the camera's internal geometry can be kept entirely unchanged by simply refraining from focusing and zooming. When the camera is allowed to change focus and zoom, pixel's aspect ratio and the image's skew factor, on which *a priori* information may be available, have been shown to remain stable. In general, autocalibration is carried out by fitting, via (preferably) a geometrically meaningful objective [17], the internal geometry of the camera associated with each image to that of a natural perspective camera model, i.e.

*This work was supported by French state funds managed by the ANR within the Investissements d'Avenir program (Labex CAMI) under reference ANR-11-LABX-0004.

a camera with square pixels (no skew and unit aspect ratio) and an image-centered principal point. Note that cameras with rectangular pixels, i.e. no skew and non-unit aspect ratio, can easily be reduced to unit aspect provided the aspect ratio is known. In the absence of prior knowledge about the aspect ratio, the natural perspective camera model is widely employed. As it turns out, using the natural perspective camera model works very well in practice as it faithfully describes a large number of consumer cameras.

However, when collecting images of the same scene from different "unknown" sources (e.g. images from the web), the natural camera model assumption may not be valid for all images in the set at hand. This may be the case either because the imaging camera is at fault, as it does not satisfy the assumptions of a natural camera, or because the image itself has been altered through, most commonly, resizing or cropping. For instance, the principal point's position is altered in the case of asymmetrically cropped images [13] (Figure 6.b) and an uneven resizing affects the aspect ratio (Fig. 6.c). Autocalibration generally revolves around using the so-called Absolute Conic (AC) as a surrogate calibration object. The AC is a virtual conic on the plane at infinity whose image (IAC) and dual image (DIAC) happen to convey the viewing camera's intrinsic parameters. Constraints, such as partial or full constancy of the camera's parameters are translated into constraints on either the IAC or DIAC. For instance, the DIAC in each view was shown to emanate from a subset of planes on the so-called Dual Absolute Quadric (DAQ) [25, 10] : a special degenerate quadric of planes tangent to the AC. Nearly a decade later, it was in turn shown that the IAC in each view resulted from lines on the so-called Absolute Quadratic Complex (AQC) [23], also referred to as the Absolute Line Quadric [26] : a degenerate quadric formed by all lines intersecting the AC. Estimating either the DAQ or AQC allows one to retrieve the sought after camera parameters along with its supporting plane which they encode. Together, the plane at infinity and the camera parameters are ultimately used to lift the projective reconstruction up to a metric frame. Autocalibration via the estimation of either the DAQ or AQC is a challenging nonlinear problem.

However, assuming the camera has zero-skew, the principal point coordinates are decoupled on the IAC whereas DIAC, there exists coupled terms. Adding known aspect ratio assumption, the focal lengths are stand-alone entities on the IAC whereas DIAC, conjugating principal point coordinates. Hence, these constraints can be enforced more efficiently on the IAC rather than DIAC. Despite that, estimating the DAQ has received significantly more attention in the literature than that of estimating the AQC. Those are the motivations for us to exploit the IAC-related AQC formulation to solve the problem in hand.

More recent methods take account of the global optimality of the solution, that is, the convergence towards the correct solution is guaranteed to a certain degree. In [5], Fusiello *et al*let@tokeneonedotdeveloped a method with epipolar geometry constraints under a Branch-and-Bound (BnB) interval minimization, the minimization is conducted by interval analysis where the progress takes lengthened time to converge. In [2], Bocquillon *et al*let@tokeneonedotproposed a BnB method based on DAQ but limited to the estimation of solely focal length as the only unknown intrinsic parameter. Differ to the aforementioned methods, Chandraker *et al*let@tokeneonedot([3]) considered the varying intrinsic parameters case, the method presented the constraints on DAQ as non-convex polynomials solved by a hierarchy of convex LMI (Linear Matrix Inequality) relaxations (Lasserre ,[14]) which relies on a pre-defined relaxation order. In theory, the method can reach global optimality when relaxation order is adequate, though in practice, such a relaxation order can not be determined a priori and setting a high relaxation order leads the optimization to be more likely numerically untraceable. Lastly, the method relies on a good scaling (normalization) of the polynomials.

In this paper, we address the autocalibration problem for a set of cameras where outliers on intrinsic characteristics may present. Based on the known aspect ratio and no-skew hypotheses, We devised a method which minimizes the median-residual of the rational polynomials from AQC formulation. The solution is certified up to an user-defined optimality criterion by bounded Sum-of-Square (SoS) convexity test and converge through BnP paradigm. Moreover, with polytopical polynomial presentation of the bounded problem, we minimized the bound on intrinsic parameters sequentially simultaneously with SoS convexity test.

The proposed method is compared to the State-of-the-Art globally optimal method in [3]. The proposed method yields competitive results with significantly better performance in high level of image noises and/or with ouitliers in the data.

Firstly, we define the necessary notations and lay down the background of the AQC formulation. We then introduce our proposed method in the following section. The experimental results are analyzed next and finally, we summarize the conclusion.

2 Background

In this paper, we consider a scene embedded in a projective 3-space and observed by n cameras whose respective intrinsic parameters are encapsulated in upper-triangular matrices $\mathbf{K}^i$ of the form

$$\mathbf{K} = \begin{bmatrix} \tau f & \gamma & u \\ 0 & f & v \\ 0 & 0 & 1 \end{bmatrix} \tag{1}$$

where f, τ, γ, and (u, v) respectively denote the focal length, skew factor, aspect ratio, and principal point's pixel location of the considered camera. A scene point X_j is represented by a homogeneous 4-vector $\mathbf{X}_j$. Its projection y^i_j on an arbitrary image i is represented by its homogeneous

coordinates $\mathbf{y}_j^i \sim \mathbf{M}^i\mathbf{X}_j$, where $\mathbf{M}^i$ is a 3×4 projection matrix. Similarly, a line L_k in the scene is represented by its Plücker coordinate 6-vector $\mathbf{L}_k$ and projects onto image i onto a line l_k^j with coordinates

$$\mathbf{l}_k^i \sim \mathbf{P}^i\mathbf{L}_k,\ i = 1 \dots n \tag{2}$$

where each $\mathbf{P}^i$ is a 3×6 line-projection matrix. If X_1 and X_2 are two distinct points, then the line L joining them projects onto the line $\mathbf{M}^i\mathbf{X}_1 \wedge \mathbf{M}^i\mathbf{X}_2$ joining their pixel projections in image i. Line $\mathbf{M}^i\mathbf{X}_1 \wedge \mathbf{M}^i\mathbf{X}_2$ can hence be factored into $\mathbf{P}^i\mathbf{L}$. Given $\mathbf{M}^i \sim [\ \tilde{\mathbf{M}}^i \mid \mathbf{m}^i\ ]$, where $\tilde{\mathbf{M}}^i$ is a 3×3 matrix and $\mathbf{m}^i$ a 3-vector, and for some appropriate choice of $\mathbf{L}$, $\mathbf{P}^i$ can be given by

$$\mathbf{P}^i \sim [\ \det(\tilde{\mathbf{M}}^i)\tilde{\mathbf{M}}^{i-\mathsf{T}} \mid -[\ \mathbf{m}^i\ ]_\wedge\tilde{\mathbf{M}}^i\ ],\ i = 1 \dots n. \tag{3}$$

The Absolute Quadratic Complex [23], also referred to as the Absolute Line Quadric [26], is a set of lines intersecting the Absolute Conic. The AQC, represented by a 6×6 rank 3 positive semi-definite matrix $\mathbf{\Omega}$, and the image of the AC on each image i, represented by a 3×3 positive definite matrix $\boldsymbol{\omega}^{\boldsymbol{i}} \sim \mathbf{K}^{i-\mathsf{T}}\mathbf{K}^{i-1}$, are related by

$$\mathbf{P}^i\mathbf{\Omega}\mathbf{P}^{i\mathsf{T}} \sim \boldsymbol{\omega}^i. \tag{4}$$

Assuming the world reference frame is arbitrarily attached to a camera with $\boldsymbol{\omega}$ as image of AC, $\mathbf{\Omega}$ factors as

$$\mathbf{\Omega} \sim \begin{bmatrix} \mathbf{I} \\ [\ \mathbf{n}\ ]_\wedge^\mathsf{T} \end{bmatrix}\boldsymbol{\omega}\begin{bmatrix} \mathbf{I} & [\ \mathbf{n}\ ]_\wedge \end{bmatrix} \tag{5}$$

where $\mathbf{I}$ is the 3×3 identity matrix and the 3-vector $\mathbf{n}$ is such that $(\mathbf{n}^\mathsf{T}\ 1)^\mathsf{T}$ represents the coordinates of the plane at infinity. Note that if $\mathbf{\Omega}$ is retrieved, $\mathbf{n}$ and $\mathbf{K}$ (the intrinsic parameters matrix of the reference camera) can be extracted and the projective scene and cameras can be upgraded to a metric frame [9].

The AQC formulation (4) is particularly well suited for autocalibration when the cameras exhibit no skew and unit aspect ration. For instance, the image of AC of a camera with $\gamma = 0$ and $\tau = 1$ is of the simplified form

$$\boldsymbol{\omega} = \begin{bmatrix} 1 & 0 & -u \\ 0 & 1 & -v \\ -u & -v & \beta \end{bmatrix} \tag{6}$$

where $\beta = f^2 + u^2 + v^2$. Denoting by $\omega_{k\ell}^i$ the entry at the k^{th} row and ℓ^{th} column of $\mathbf{P}^i\mathbf{\Omega}\mathbf{P}^{i\mathsf{T}}$, the skew and aspect ratio assumptions yield the following constraints

$$\omega_{11}^i - \omega_{22}^i = 0 \text{ and } \omega_{12}^i = 0,\ i = 1 \dots n. \tag{7}$$

The entries of $\mathbf{\Omega}$ may be obtained through a linear least-squares solution using (7) along with the trace property $\mathrm{trace}(\mathbf{\Omega}) = 0$ from at least 10 images. Additionally, when an image-centered principal point is considered, a proper shifting of the image's origin to the image center yields two additional constraints on $\mathbf{\Omega}$:

$$\omega_{13}^i = 0 \text{ and } \omega_{23}^i = 0,\ i = 1 \dots n. \tag{8}$$

In such case, the number of required images for a linear estimate of $\mathbf{\Omega}$ drops to 6. In all cases, the rank of $\mathbf{\Omega}$ may be enforced *a posteriori*. However, the positive semi-defniteness of $\mathbf{\Omega}$ is not guaranteed. After an estimate of $\mathbf{\Omega}$ is obtained linearly, $\mathbf{n}$ and the intrinsic parameters of the reference camera can be refined through nonlinear least-squares.

3 Robust autocalibration

In the camera autocalibration problem we address, we assume that half or more of the cameras, including the reference one, satisfy the pinhole camera model with no skew and unit aspect ratio. It is customary to seek the solution that, in addition to the skew and aspect ratio constraints, also minimizes the distance of the principal point to the image center. Up to half the cameras could violate such assumptions, whether on the skew, the aspect ratio or on the principal point. Such cameras would be referred to as outliers.

Assuming each image's reference frame is at its center, each camera would provide, for any given $\mathbf{n}$ and $\boldsymbol{\omega}$, the vector of residuals

$$r_i = [\frac{\omega_{11}^i - \omega_{22}^i}{\omega_{11}^i}\ \frac{\omega_{12}^i}{\omega_{33}^i}\ \frac{\omega_{13}^i}{\omega_{11}^i}\ \frac{\omega_{23}^i}{\omega_{22}^i}]. \tag{9}$$

A robust method for solving such problem with outliers is by minimizing the Least-Median-Squares (LMS) residuals :

$$\begin{aligned} \delta^* = \min_{\mathbf{n} \in \mathcal{B}_\mathbf{n},\, \boldsymbol{\omega} \mathcal{B}_{\boldsymbol{\omega}}} \quad & \underset{i=1\dots n}{\mathrm{median}}\ ||r_i||_\infty^2 \\ \text{s.t.} \quad & \boldsymbol{\omega} \succeq 0 \end{aligned} \tag{10}$$

where $||.||_\infty$ is the infinity norm. The sets $\mathcal{B}_{\boldsymbol{\omega}}$ and $\mathcal{B}_\mathbf{n}$ respectively denote the bounding boxes defined by $[\underline{\beta}, \overline{\beta}] \times [\underline{u}, \overline{u}] \times [\underline{v}, \overline{v}] \subset \mathbb{R}^3$ and $[\underline{n}_1, \overline{n}_1] \times [\underline{n}_2, \overline{n}_2] \times [\underline{n}_3, \overline{n}_3] \subset \mathbb{R}^3$. $\underline{n}_i, \overline{n}_i$ denote the lower and upper bound of i-th entry of $\mathbf{n}$ respectively, and $\underline{n}_i \leq \overline{n}_i$. Note that each entry of the vector of residuals is a rational polynomial of the form

$$\frac{f_i(\mathbf{x})}{g_i(\mathbf{x})} \tag{11}$$

where $\mathbf{x} = (\boldsymbol{\omega}, \mathbf{n})$. Since the polynomials are cubic in $(\boldsymbol{\omega}, \mathbf{n})$, δ^* is optimized through a non-linear optimization and not guaranteed in anyways, an optimal solution.

3.1 Globally ϵ-optimal and pruning principle

First, for a δ^* ($\beta_\mathbf{x}^* \subset \beta_\mathbf{x}$) to be a better solution comparing to all the solutions in an arbitrary bound sub to the initial bound ($\beta_\mathbf{x}' \subset \beta_\mathbf{x}$), the following conditions must be satisfied : for 50% of the images, there must exist one polynomial per image has the costs of all its possible solutions higher in absolute value comparing to δ^*, that is, depending on the sign of the polynomial, bigger or smaller than δ^* or $-\delta^*$ respectively, such a polynomial inequality can

be formulated as :

$$\varsigma \frac{f_i(\mathbf{x})}{g_i(\mathbf{x})} > \delta^*,\ \varsigma \in \{-1, +1\} \rightarrow c_i(\mathbf{x}) > 0, \tag{12}$$

the problem then can be generalized as :

$$\begin{aligned} &\sum_{j=1}^{n} E_j(\mathbf{x}) \geq \frac{j}{2}, \quad \forall x \in \beta'_{\mathbf{x}},\ i = 1 \dots 8 \\ &E_j(\mathbf{x}) = 1 \iff \exists\, c_i(\mathbf{x}) > 0, \\ &E_j(\mathbf{x}) = 0 \iff \nexists\, c_i(\mathbf{x}) > 0. \end{aligned} \tag{13}$$

If true, $\beta'_{\mathbf{x}}$ can never reach δ^*, thus, pruned ; If false, in the BnP paradigm, $\beta'_{\mathbf{x}}$ is branched into two child nodes where the biggest interval among all variables in $\mathbf{n}$ is split at the mean.

Note that if $\forall \beta'_{\mathbf{x}} \subset \beta_{\mathbf{x}}$, Eq.(13) is satisfied, meaning, all nodes can be pruned by δ^*, then naturally, δ^* is the globally optimal solution.

Moreover, for Eq.(12) to be true for all $\beta'_{\mathbf{x}}$, eventually, $\beta'_{\mathbf{x}}$ has to be infinitely small. Therefore, to converge in finite time in practice, a user-defined optimality criterion (ϵ) is added :

$$\varsigma \frac{f_i(\mathbf{x})}{g_i(\mathbf{x})} + \epsilon > \delta^* \tag{14}$$

Hence, δ^* is globally ϵ-optimal.

3.2 Branch-and-Prune

To summarize the BnP paradigm :

1. Initialize $\beta_{\mathbf{x}}$ and initialize δ^* with $\beta_{\mathbf{x}}$'s solution (Eq.(11)), add $\beta_{\mathbf{x}}$ into the BnP queue ;
2. Take the first node in the queue, i.e, $\beta'_{\mathbf{x}}$. Process it with Eq.(13). If pruned, skip to Step.4 ;
3. If branched, estimate δ' with Eq.(11) in the child nodes, if $\delta' < \delta^*$, we update $\delta^* = \delta'$, add the nodes to the back of the queue.
4. Return to step 2 or stop if the queue is empty and δ^* is the globally ϵ-optimal solution.

Note that in a convectional BnP paradigm, pruning is done by comparing optimistic and pessimistic costs whereas in our case, we only have one cost, pruning is done as described in previous section.

3.3 Linearization via Gram matrix

As summarized in Eq.(12), pruning is depending on determining the positiveness of a polynomial in a given bound. Relating $c_i(\mathbf{x}) > 0$ to any rational polynomials described in Eq.(9), one can find that $c_i(\mathbf{x})$ is quadratic in $\mathbf{n}$ and $(\boldsymbol{\omega}, \mathbf{n})$ is cubic. The problem is $\mathcal{NP}$-hard and to test its positiveness, $(\boldsymbol{\omega}, \mathbf{n})$ must be decoupled and linearized, at the same time, subject to the bound on $\mathbf{x}$.

To do so, let us first assume $\boldsymbol{\omega}$ is fixed at a vertex, for instance, $\boldsymbol{\omega}_1 = \boldsymbol{\omega}(\beta \mapsto \overline{\beta}, u \mapsto \overline{u}, v \mapsto \overline{v})$. One may add a SoS underestimator term :

$$\begin{aligned} &c_1(\boldsymbol{\omega}_1, \mathbf{n}) - \sum_{j=1}^{3} \alpha_{1j}(\overline{n}_j - n_j)(n_j - \underline{n}_j) \geq 0, \\ &s.t. \quad \alpha_{1j} \geq 0,\ \boldsymbol{\alpha}_1 \in \{\alpha_{1j}\},\ j = 1 \dots 3, \end{aligned} \tag{15}$$

then convert such a quadratic polynomial into its Gram matrix formulation $\boldsymbol{G}(\boldsymbol{\omega}_1, \boldsymbol{\alpha}_1)$ (or simplified to $\boldsymbol{G}_1$) as follows :

$$[\mathbf{n} \quad 1]^{\mathsf{T}} \boldsymbol{G}(\boldsymbol{\omega}_1, \boldsymbol{\alpha}_1) [\mathbf{n} \quad 1] \geq 0. \tag{16}$$

By checking the PSD of $\boldsymbol{G}_1$: $\boldsymbol{G}_1 \succeq 0$ is equivalent to proving the positiveness of Eq.(15), note that the bound of $\mathbf{n}$ has been embedded in $\boldsymbol{G}_1$ as parameters. For more details regarding Gram matrix, [19] can be referred to.

3.4 Polytopic polynomials and bound optimization

Following the previous section, we are able to verify the positiveness of a rational polynomial $c_i(\boldsymbol{\omega}_i, \mathbf{n})$ at an arbitrary vertex $\boldsymbol{\omega_i}$ in the given $\boldsymbol{n}$ bound, by checking the feasibility of its equivalent LMI problem : $\boldsymbol{G}_i \succeq 0$. This is one step apart to the goal : verify the positiveness of $c_i(\boldsymbol{x}) > 0$ which is bounded not only by $\boldsymbol{n}$, but also $\boldsymbol{\omega}$. For that, according to the polytope theory ([20]), one can verify the positiveness of a set of polytopic polynomials representing the vertices of the bound. In our case, a set of LMIs representing the vertices of $\boldsymbol{\omega}$:

$$\{\boldsymbol{G}_i \succeq 0\}_{i=1}^{N}. \tag{17}$$

N is the number of vertices. In this case, $N = 8$, which is the amount of the combinations of the lower and upper $\boldsymbol{\omega}$ bound. Further, we can also define the polytope $\mathcal{G}$ of the polynomial, again, through its Gram matrices :

$$\mathcal{G} = \left\{ \boldsymbol{G} = \sum_{i=1}^{N} \theta_i \boldsymbol{G}_i,\ \sum_{i=1}^{N} \theta_i = 1,\ \theta_i \geq 0 \right\} \tag{18}$$

Note that θ_i only affects $\boldsymbol{\omega}_i$ terms in $\boldsymbol{G}_i$. This is to show that $\mathcal{G}$ can represent all values in bound by the linear combinations of the vertices. This means that if one of the entry of $\boldsymbol{\omega}$ is subtitled as a variable, it can be minimize linearly. This property is used to minimize the bound on $\boldsymbol{\omega}$. For instance, to minimize the bound on β, at the same time, verify the positiveness of concerning polynomial :

$$\begin{aligned} \min_{\beta, \theta_i, \boldsymbol{\alpha}_i} \quad & \beta, \quad i = 1 \dots N, \\ s.t. \quad & \boldsymbol{G}(\beta) \succeq 0, \\ & \textstyle\sum_{i=1}^{N} \theta_i = 1,\ \theta_i \geq 0. \end{aligned} \tag{19}$$

$N = 4$ since β is a variable and (u, v) are at their vertices. It is obvious to maximize the β bound similarly to (19) and adopt it to update (u, v) bounds. Note that if the PSD of $\boldsymbol{G}(\beta)$ is not satisfied while optimizing the bounds, the polynomial $c_i(\boldsymbol{x})$ is pruned. Therefore, the bound minimization is done at Step.2 in Section 3.2 simultaneously with pruning.

4 Experiments

We have tested the proposed method in synthetic simulations and with real image datasets. Our method is compared to the State-of-Art method from Chandraker *et*

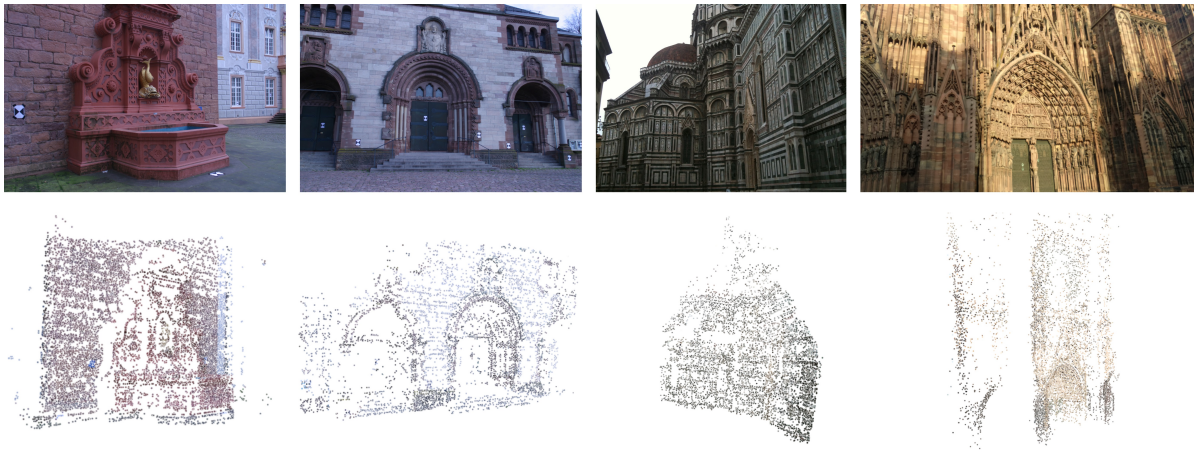
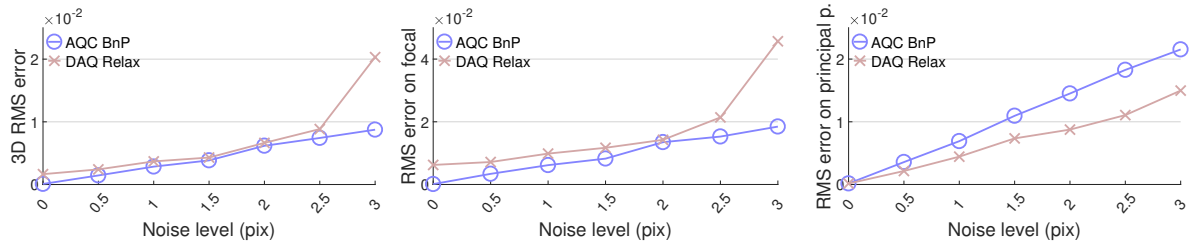


FIGURE 1 – Image samples from datasets (top) and their 3D reconstructions with outlier images (bottom). Left to right : 'Fountain', 'Herz-jesu', 'Florence', 'Notre dame'.



(a) 3D RMS (b) Focal length (c) Principal p.

FIGURE 2 – Results from setups of 10 images and varying noise level (pixel noise)

*al*let@tokeneonedot[3]. The two methods are further referred to as 'AQC BnP' and 'DAQ Relax' respectively.

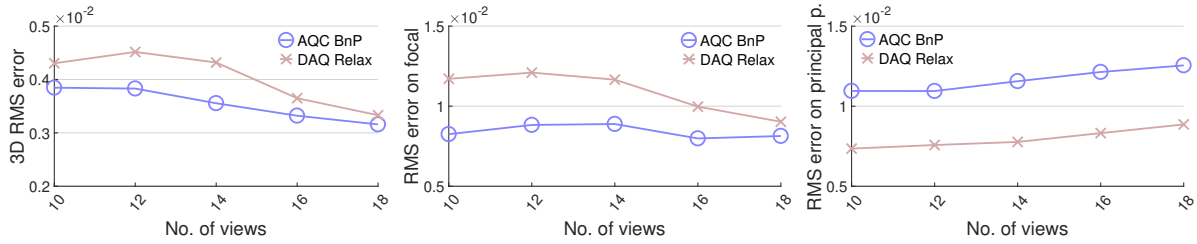
In all experiments, the image points are normalized beforehand. The projective scene reconstruction (recovering point projection matrices $\mathbf{M}_i$) is then retrieved using [18] and further refined by Bundle Adjustment (from Vincent SfM toolbox [24]). The implementation and experiments are carried out with MATLAB 2017a. The LMI constraints are implemented with Yalmip [15] parser and solved by MOSEK [1]. The hardware specifications are Intel i7@2.80GHz CPU and 16GB RAM. For DAQ Relax, the degree of polynomial relaxation is set to 2 as suggested in the original paper [3]. For AQC BnP, the optimality criterion ϵ is set to $5e^{-2}$. The local non-linear optimization is implemented with the Matlab function so-called 'fminisearchcon' where the non-linear non-smooth problem is solved by a simplex method ([12]) and non-linear constraints are converted to penalty functions and added to the optimization. Reader may refer to Matlab documentation for more details.

The results are compared in 3 quantified measures : the normalized Root-Mean-Square (RMS) error of 3D point reconstruction (Δ_{3D}), the normalized intrinsic parameters RMS error (Δf and Δuv) and the computing time. The 3D point are normalized to a unit sphere and Δ_{3D} the RMS error of the euclidean distance from result points to corresponding ground truth points. Other RMS errors represent the normalized error from result to the ground truth value.

4.1 Simulations

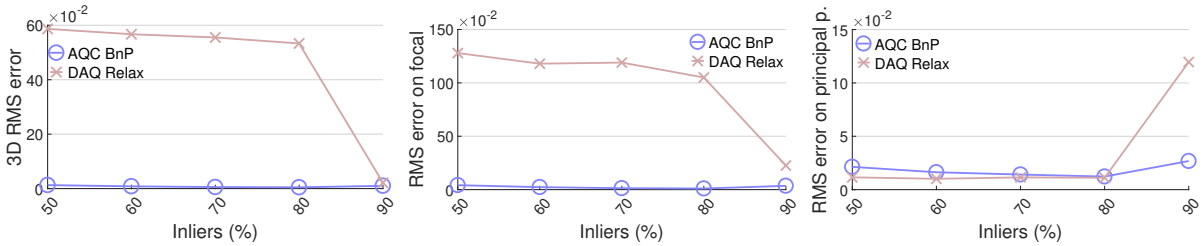
All synthetic images are in size 512×512 pixels. The camera parameters are : $f = 800, u = 256, v = 256, \tau = 1, \gamma = 0$ unless noticed otherwise. The image sequences are generated with various lengths ($n \in \{10, 12, 14, 16, 18\}$), and zero-mean Gaussian pixel noise (σ) is added to image point coordinates in 7 levels, the standard deviations are : $\sigma \in \{0, 0.5, 1, 1.5, 2, 2.5, 3\}$. AQC BnP is initialized with the bounds as $u, v \in \{0, 512\}$ which represents the full image plane disparity, and $f \in \{0, 5120\}$. A combination of n and σ makes one experimental setup. The RMS error per setup is calculated from 100 samples. For an arbitrary sample, the data is generated as follows : A set of 500 3D feature points is generated with normal distribution. 2D image points are acquired from 3D points projection with random motions.

Fixed number of images : In Figure 2, the RMS errors ($\Delta_{3D}, \Delta f, \Delta uv$) increase as σ increases. As observed, the quality of 3D reconstruction is more correlated to the quality of f estimation rather than to (u, v) as the Δ_{3D} follows closely to Δf proportionally. Therefore, although DAQ



(a) 3D RMS (b) Focal length (c) Principal p.

FIGURE 3 – Results from setups of 1.5 pixels of noise and varying number of images



(a) 3D RMS (b) Focal length (c) Principal p.

FIGURE 4 – Results from setups with 10 images, 1.5 pixels of noise and varied numbers of inliers where $\tau = 1.1$.

Relax has better principal point estimation, AQC BnP has a better reconstruction (Δ_{3D}) throughout. Moreover, the errors increased linearly for AQC BnP whereas for DAQ Relax, the increment is significantly larger at $\sigma = 3$, this indicates the advantage of AQC BnP in terms of robustness against high-level image noises.

Fixed pixel noise level : In Figure 3, the RMS errors decrease as n increases. This is expected as the more images, the more information/constraints added to the cost. The difference between two methods is decreasing while n increases, this indicates a better minimal-view performance on AQC Relax.

With outliers : Since AQC BnP is solving a median-residual optimization problem, inherently, the method is robust to so-called 'outlier' images, where the assumed constraints on intrinsics (τ, γ) varies, meaning the unit-aspect ratio and/or zero-skew condition has been violated. For example, in Fig. 4, the outliers, instead of unit-aspect ratio ($\tau = 1$), $tau = 1.1$. It is obvious the impact of outliers on the result for DAQ Relax whereas AQC BnP remain invariant to outliers and yields good performance throughout. Note with 90% inliners, or 1 outlier in the set, DAQ Relax's reconstruction improved significantly, compromising the principal point (P.p) estimation. On the other hand, if the aspect ratio is increased to $\tau = 1.25$ for the same setup, as shown in Table 1, DAQ Relax is unable to cope further with the outlier where AQC Relax is effected as well on certain samples, but it still gives high percentage of quality reconstructions among all samples. This shows the high sensitivity of DAQ Relax while AQC BnP shows its robustness on know-aspect ratio violation in the outliers.

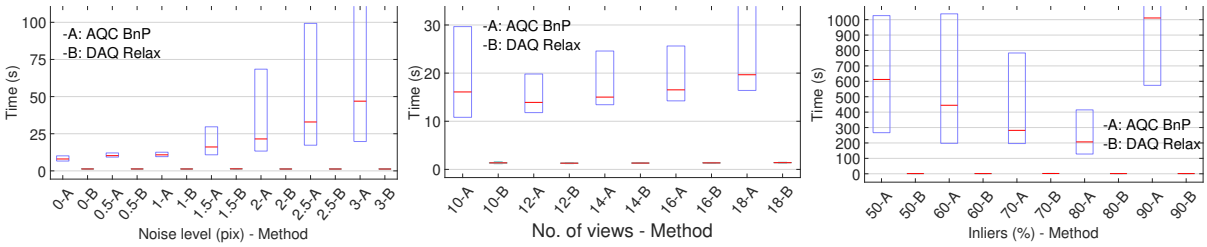
The impacts on (u, v, γ) violations are not as significant as on f against both methods, although it is worth mentioning that AQC BnP has a bit better performance (Tab. 1).

Setup	Method	Δ_{3D}	Δf	Δuv
τ = 1.25	AQC	0.0799	0.0266	$2.41e^{-2}$
	DAQ	0.572	1.228	$1.06e^{-2}$
P.p = 1.6 (u_o, v_o)	AQC	$4.4e^{-3}$	$9.1e^{-3}$	$1.35e^{-2}$
	DAQ	$6.8e^{-3}$	$52.2e^{-3}$	$2.19e^{-2}$
Skew = 85°	AQC	$4.3e^{-3}$	$9.6e^{-3}$	$1.01e^{-2}$
	DAQ	$4.9e^{-3}$	$12.5e^{-3}$	$0.52e^{-2}$

TABLE 1 – Results from setups with 10 images with one outlier image, 1.5 pixels of noise. The outlier varies depending on the setup as shown in the table.

Time analysis : Regarding computing time (Figure 5), as the complexity of DAQ Relax is invariant to n, computing time stays invariant to varying σ and n. DAQ Relax gives a fast finishing time throughout. AQC BnP gives a faster performance on lower noise level (Fig. 5.a) and stays variance of n (Fig. 5.b). The computing time increases as the error of the outlier value increases or the number of outlier images increases (Fig. 5.c) with an exception on 90% inliers case. The reason can be the same as the effect on DAQ Relax in Fig. 4. AQC BnP is dealing the ambiguity between identifying the outlier, excluding the image from median cost and a deviated P.p estimation which may improve the cost as well.

To summarize, the simulation shows that our method has the robustness advantage when outlier images are present in the data. Despite a bigger error on principle point esti-



(a) (b) (c)

FIGURE 5 – Computing time analysis : a : the setup in Fig. 2 ; b : the setup in Fig. 3 ; c : the setup in Fig. 4.

(a) (b) (c)

FIGURE 6 – An example of outlier images. a : Original image ; b : After ($\tau = 1.1$) being stretched or resized ; c : After asymmetrical cropping, shifting (u, v). Image is from the 'fountain' dataset.

mation, the 3D reconstruction performance is superior with and without outliers in comparison to DAQ Relax.

4.2 Real images

The methods are tested on 2 well-known multi-view datasets (the 'Fountain' (11 images) and 'Herz-jesu' (8 images) sequence), also, on two datasets created with images taken by a phone camera on the scene at 'Notre dame' in Strasbourg (18 images) and in 'Florence'. The image samples from these datasets are shown in Fig.1. To represent the possible outliers commonly seen on the web, (τ, u, v) are varied, the examples are shown in Fig.6.

The results and its ground truth are shown in Table 2. On 'Notre dame' and 'Florence', One may refer the quality of the solution to results in Setup 1 (No outliers). The results follows the analysis of the simulations. The AQC BnP yields a good performance in spite of the various outlier setups. On 'Fountain', large τ outlier leads to DAQ Relax failure (Setup 2) whereas shifting (u, v) (Setup 3) did little effects. On 'Herz-jesu', DAQ Relax results in large errors (Setup 2 and 3) where τ outliers present. On 'Notre dame', as τ and number of outlier images increases from setup 2 to 4, AQC BnP's results stay close to Setup 1 where the errors grow for DAQ Relax. On 'Florence', note that we have 5 images from the dataset, therefore, only 2 images are considered as inliers for the median-residual optimization as shown in Eq.(13) which is below the minimal-view case, hence, a minimal number of residual is fixed to 4 when median residual falls below, this also shows the flexibility of AQC Relax. For Setup 2, 2 additional images (7 in total) found on-line are added in for calibration. We can observe the impacts of such a least favorable case (minimal-view sequence, high percentage of outliers with high level deviation on aspect ratio, unknown source of images) on DAQ Relax in Setup 2 whereas AQC BnP's result stay invariant.

5 Conclusion

In this paper, we proposed a novel robust globally optimal autocalibration method that is based on the AQC formulation. The proposed method is robust to geometry changes (pinhole camera with no skew and unit aspect ratio) across images. Our method relies on Least-Median-Squares optimization to minimize the deviation from such camera model. The retained solution is guaranteed to be optimal and relies on a polytopic formulation of the problem and polynomial SoS theory within a Branch-and-Prune search to provide such guarantee. In both simulations and real image experiments, our method has shown not only a very good performance throughout, but also significant advantages regarding the robustness against high-level pixel noise as well as when outlier images were present.

Sequences	Setup	Method	f_1	f_2	u	v	Time (s)
Fountain	1	AQC	2785.1	-	1529.1	1023.1	66.3
		DAQ	2763.0	2762.9	1535.0	1023.9	1.2
	2	AQC	2783.5	-	1530.2	1023.5	707.3
		DAQ	1248.7	1251.9	1529.2	1021.3	1.3
	3	AQC	2759.1	-	1536.9	1024.1	280.1
		DAQ	2750.5	2752.3	1543.1	1026.9	2.0
	GT	-	2759.5	2764.2	1520.7	1006.8	-
Herz-jesu	1	AQC	2769.3	-	1533.6	1026.3	14.1
		DAQ	2776.1	2776.5	1535.8	1025.2	1.1
	2	AQC	2768.9	-	1532.8	1029.0	93.7
		DAQ	2184.9	2185.1	1524.0	1028.5	2.0
	3	AQC	2762.2	-	1533.2	1027.0	245.4
		DAQ	551.4	552.0	1528.3	1034.5	1.1
	GT	-	2759.5	2764.2	1520.7	1006.8	-
Florence	1	AQC	2081.0	-	1039.9	770.6	45.9
		DAQ	2122.1	2125.8	1024.3	771.2	1.0
	2	AQC	2029.9	-	1022.4	766.4	202.8
		DAQ	329.6	353.8	1000.4	800.2	4.1
Notre dame	1	AQC	3012.5	-	2000.0	1125.0	64.4
		DAQ	2985.7	2966.9	2002.5	1125.6	1.5
	2	AQC	2922.9	-	2027.4	1152.4	956.4
		DAQ	2883.5	2811.7	1992.2	1111.2	1.6
	3	AQC	2896.6	-	2038.9	1160.3	532.6
		DAQ	2547.5	2445.5	1940.9	1071.9	1.5
	4	AQC	2887.3	-	2007.7	1132.7	245.4
		DAQ	2451.7	2383.3	1889.4	1049.0	1.1

TABLE 2 – Results from real image datasets, compared to grand truth (*GT*let@tokeneonedot) with various outlier setups.
'Fountain' (11 images) : Setup 1 : no outlier images ; Setup 2 : 2 outlier images with $\tau = 1.25$; Setup 3 : 1 outlier image with shifted (u, v) as shown in Fig. 6.c.
'Herz-jesu' (8 images) : Setup 1 : no outlier images ; Setup 2 : 1 outlier image with $\tau = 1.1$; Setup 3 : 2 outlier images with $\tau = 1.1$ and $\tau = 1/1.1$ respectively.
'Florence' : Setup 1 : no outlier images (5 images) ; Setup 2 : 3 outlier images with $\tau = 1.25$.
'Notre dame' (18 images) : Setup 1 : no outlier images ; Setup 2 : 5 outlier images with $\tau = 1.1$; Setup 3 : 5 outlier images with $\tau = 1.2$; Setup 4 : 7 outlier images with $\tau = 1.2$.

Références

[1] M. ApS. The mosek optimization toolbox for matlab manual. version 7.1 (revision 28), 2015.

[2] B. Bocquillon, A. Bartoli, P. Gurdjos, and A. Crouzil. On constant focal length self-calibration from multiple views. In *IEEE Conf. on Computer Vision and Pattern Recognition*, 2007.

[3] M. Chandraker, S. Agarwal, F. Kahl, D. Nistér, and D. Kriegman. Autocalibration via rank-constrained estimation of the absolute quadric. In *IEEE Int'l Conf. on Computer Vision and Pattern Recognition*, 2007.

[4] M. Chandraker, S. Agarwal, D. Kriegman, and S. Belongie. Globally optimal algorithms for stratified autocalibration. *International journal of computer vision*, 90(2) :236–254, 2010.

[5] A. Fusiello, A. Benedetti, M. Farenzena, and A. Busti. Globally convergent autocalibration using interval analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(12) :1633–1638, 2004.

[6] R. Gherardi and A. Fusiello. Practical autocalibration. In K. Daniilidis, P. Maragos, and N. Paragios, editors, *Computer Vision – ECCV 2010*, pages 790–801, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.

[7] A. Habed, K. A. Ismaeil, and D. Fofi. A new set of quartic trivariate polynomial equations for stratified camera self-calibration under zero-skew and constant parameters assumptions. In *European Conf. on Computer Vision*, pages 710–723. Springer Berlin Heidelberg, 2012.

[8] A. Habed, D. Paudel, C. Demonceaux, and D. Fofi. Efficient pruning lmi conditions for branch-and-prune rank and chirality-constrained estimation of the dual absolute quadric. In *IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2014.

[9] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge university press, 2003.

[10] A. Heyden and K. Astrom. Euclidean reconstruction from constant intrinsic parameters. In *Pattern Recognition, 1996., Proceedings of the 13th International Conference on*, volume 1, pages 339–343. IEEE, 1996.

[11] A. Heyden and K. Åström. Flexible calibration : Minimal cases for auto-calibration. In *Proceedings of the 7th International Conference on Computer Vision*, 1999.

[12] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright. Convergence properties of the nelder–mead simplex method in low dimensions. *SIAM Journal on optimization*, 9(1) :112–147, 1998.

[13] V. Larsson, Z. Kukelova, and Y. Zheng. Camera pose estimation with unknown principal point. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[14] J. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3) :796–817, 2001.

[15] J. Lofberg. Yalmip : A toolbox for modeling and optimization in matlab. In *Int'l Symposium on Computer Aided Control Systems Design*, pages 284–289, 2004.

[16] Q. Luong and O. Faugeras. Self-calibration of a moving camera from point correspondences and fundamental matrices. *International Journal of computer vision*, 22(3) :261–289, 1997.

[17] D. Nistér. Untwisting a projective reconstruction. *Int'l Journal of Computer Vision*, 60(2) :165–183, 2004.

[18] J. Oliensis and R. Hartley. Iterative extensions of the sturm/triggs algorithm : Convergence and nonconvergence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(12) :2217–2233, December 2007.

[19] D. Paudel, A. Habed, C. Demonceaux, and P. Vasseur. Robust and optimal sum-of-squares-based point-to-plane registration of image sets and structured scenes. In *IEEE International Conference on Computer Vision*, pages 2048–2056, 2015.

[20] D. Peaucelle, D. Arzelier, O. Bachelier, and J. Bernussou. A new robust d-stability condition for real convex polytopic uncertainty. *Systems & control letters*, 40(1) :21–30, 2000.

[21] M. Pollefeys and L. V. Gool. Stratified self-calibration with the modulus constraint. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(8) :707–724, 1999.

[22] M. Pollefeys, R. Koch, and L. V. Gool. Self-calibration and metric reconstruction inspite of varying and unknown intrinsic camera parameters. *International Journal of Computer Vision*, 32(1) :7–25, Aug 1999.

[23] J. Ponce, K. McHenry, T. Papadopoulo, M. Teillaud, and B. Triggs. On the absolute quadratic complex and its application to autocalibration. In *IEEE Int'l Conf. on Computer Vision and Pattern Recognition*, 2005.

[24] V. Rabaud. Vincent's structure from motion toolbox (vincentssfmtoolbox).

[25] B. Triggs. Autocalibration and the absolute quadric. In *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, pages 609–614. IEEE, 1997.

[26] A. Valdés, J. Ronda, and G. Gallego. The absolute line quadric and camera autocalibration. *Int'l Journal of Computer Vision*, 66(3) :283–303, 2006.