Image and Vision Computing 29 (2011) 653–665

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

Maximum likelihood autocalibration☆

Stuart B. Heinrich [a,*], Wesley E. Snyder [a], Jan-Michael Frahm [b]

[a] Department of Electrical and Computer Engineering, NC State University, Box 7911, Raleigh, NC 27695-7911, United States
[b] Department of Computer Science, Campus Box 3175, Brooks Computer Science Building, 201 South Columbia Street, UNC-Chapel Hill, Chapel Hill, NC 27599-3175, United States

ARTICLE INFO

Article history:
Received 24 December 2010
Received in revised form 23 April 2011
Accepted 11 July 2011

Keywords:
Autocalibration
Self-calibration
Maximum likelihood
Absolute dual quadric
Structure from motion

ABSTRACT

This paper addresses the problem of autocalibration, which is a critical step in existing uncalibrated structure from motion algorithms that utilize an initialization to avoid the local minima in metric bundle adjustment. Currently, all known direct (not non-linear) solutions to the uncalibrated structure from motion problem solve for a projective reconstruction that is related to metric by some unknown homography, and hence a necessary step in obtaining a metric reconstruction is the subsequent estimation of the rectifying homography, known as autocalibration. Although autocalibration is a well-studied problem, previous approaches have relied upon heuristic objective functions, and have a reputation for instability. We propose a maximum likelihood objective and show that it can be implemented robustly and efficiently and often provides substantially greater accuracy, especially when there are fewer views or greater noise.

1. Introduction

This paper addresses the general problem of reconstructing, from a collection of corresponding image points identified in uncalibrated images, all of the camera parameters (position, orientation, focal length, etc.) and 3D coordinates of points in the scene. This problem is most often referred to as uncalibrated structure from motion (SFM).

A metric reconstruction is one that differs from the true configuration only by the choice of coordinate system; in other words, there is some unknown rotation, translation and scale [1]. It is well known that metric reconstruction is not possible from projective constraints alone [1–3] because the solution is ambiguous up to multiplication by some arbitrary homography. Thus, a reconstruction obtained from projection constraints alone is referred to as a projective reconstruction.

Given any additional constraints on the intrinsic camera parameters (e.g., that the images are not skewed, that pixel aspect ratio is known, that the center of projection is in the center of the image, or that multiple images were produced by the same physical camera), the ambiguity can be resolved. In practice, some of these constraints will always be available. However, incorporating these constraints directly into an initial estimate is difficult: an efficient solution is only possible in the simplest minimal case of two views with fully calibrated cameras [4]. A more general solution for two uncalibrated cameras was recently proposed in Hartley and Kahl [5], although the time complexity of this latter solution was prohibitively high.

In contrast, techniques for computing a projective reconstruction are much more efficient, so the usual approach is to compute an initial projective reconstruction minimally or linearly. This initial projective solution can be refined to a maximum likelihood (ML) projective reconstruction using projective bundle adjustment [1,6,7], which can be multiplied by the rectifying homography to yield an initial metric solution, and finally refined to a maximum likelihood metric reconstruction using metric bundle adjustment. The estimation of the rectifying homography is known as autocalibration (a.k.a. self-calibration).

In projective bundle adjustment, camera views are parameterized by projection matrices and the projection equation is simple, with the only nonlinearity being due to the perspective division. The basin of attraction for projective bundle adjustment is relatively large, and convergence is fast and reliable. In contrast, the presence of rotation matrices significantly complicates the projection equation in metric bundle adjustment, making the linearized update approximations less accurate. As a result, we observe a much smaller basin of attraction with less reliable convergence in metric bundle adjustment. Thus, even when an initial metric estimate is available, one may still prefer to do bundle adjustment in projective space in order to avoid local minima, and this would necessitate the use of autocalibration to map the result back into a metric space.

A plethora of approaches to autocalibration have been presented in the literature (see Section 1.1), but autocalibration has a reputation for instability, and obtaining robust results in the presence of realistic levels of measurement noise can often be difficult. This has motivated a recent trend towards approaches that use more computationally expensive global optimization methods, under the assumption that the instabilities are due to getting stuck in local minima. However, we

☆ This paper has been recommended for acceptance by Peter Sturm.
* Corresponding author. Tel.: +1 802 922 0731.
E-mail address: sbheinri@ncsu.edu (S.B. Heinrich).

0262-8856/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2011.07.003

will show that the heuristic objectives that are optimized by these global approaches are still fundamentally sensitive to noise.

In this paper, we formulate a maximum likelihood objective for autocalibration and show that it can be optimized efficiently and robustly. Using the maximum likelihood method avoids the sensitivity to noise and can give considerably more accurate results, especially for small numbers of views or high levels of relative measurement error (equivalent to more distant point clouds) where the autocalibration problem is more difficult.

We begin by summarizing previous autocalibration approaches in Section 1.1, and then derive our maximum likelihood objective in Section 2. A method for efficient optimization is described in Section 3. In Section 4 we devise a framework for objective evaluation of autocalibration performance, and identify a set of representative autocalibration algorithms to compare against. The results of our experiments are presented in Section 5, demonstrating the efficiency, robustness and quality of the proposed ML method, with some concluding remarks in Section 6.

Finally, we refer the interested reader to the appendix, where we explain the geometric relationships between all previous autocalibration constraints in Appendix B, and show that all previous constraints are enforced by our ML method. In Appendix A we provide insight into the fundamental instability and limitations behind the heuristic maximum a priori objectives (not to be confused with maximum likelihood), which are considered the current state of the art.

1.1. Background

The first known method of autocalibration was based on the Kruppa equations [8–11], now understood to be an algebraic representation of the correspondence of epipolar lines tangent to the dual image of the absolute conic (DIAC).

It was shown in [12] that an equivalent constraint to the Kruppa equations is that the essential matrix between any view pair must have two equal non-zero singular values, called the rigidity constraint. This is the fundamental principle behind several autocalibration approaches that theoretically work for two views when focal length is the only unknown [3,13–16], although they are highly sensitive to noise.

When more than two views are considered, autocalibration via the Kruppa equations requires finding the simultaneous solutions to many quadratic equations, which has not been regarded as a promising approach [1], but has been attempted using homotopy continuation [17], nonlinear methods [18,19], and more recently using globally convergent interval analysis [20]. Because the Kruppa equations do not enforce all of the calibration constraints that are now understood, such as the common support plane for the plane at infinity, these methods are subject to singularities that can lead to instabilities.

In [21], the modulus constraint on the plane at infinity was introduced, which is complementary to the Kruppa equations because it enforces constraints on the common plane at infinity without enforcing constraints on the DIAC. A unifying framework for these entities was presented with the absolute dual quadric (ADQ) [22], a fixed entity in space that encodes for both the plane at infinity and absolute dual conic (ADC) and projects to the DIACs. The ADQ is useful because all autocalibration constraints can be translated onto it.

The ADQ can be estimated using linear and nonlinear least squares [1,22–25], sometimes weighted according to prior assumptions as in [26]. Unfortunately, both of these variations are often unstable in practice [14]. It has been commented [27] that the main reason for instability of the linear method is that the rank and positive-semidefinite constraints of the ADQ are not enforced. However, we believe that the greater issue with the linear method is that the constraint equations do not directly correspond to the parameters they are intended to constrain in the presence of noise.

The nonlinear method has no singularities and enforces all known constraints, but still does not have any geometric meaning [1] and frequently produces unstable results in practice. We speculate from the recent trend towards more global approaches that minimize essentially the same cost function that the instability of the nonlinear method has been largely attributed to difficulties in obtaining a good initialization.

For example, in Hartley's stratified approach [28], chirality constraints [29] are used to solve for a finite bounding volume for the plane at infinity and then this space is explored with a brute force search. From each candidate location, the infinite homography constraint is used to linearly estimate the ADC from any desired calibration constraints, the best plane is taken as the one that minimizes the least squares residual, and finally the result is improved nonlinearly. Unfortunately, this brute force search can be slow, and we have observed that the minimum is often so pointlike that the basin of attraction is not reliably found using any reasonably spaced discretization. Additionally, it has been pointed out [30] that a single outlier can cause the chirality constraints to have no solution, or to not contain the correct solution.

More recently, the issue of discretization has been addressed by globally convergent methods. For example, interval analysis (IA) with branch and bound was used to minimize a heuristic based on the essential matrix constraint in Fusiello et al. [20]. Unfortunately, the method was not very efficient, having computation times of about 1.5 h for a problem with 40 views. IA was used again in [27], but the parameterization that was used only works for constant focal length and does not evenly distribute error. Computation times were improved in this latter method, but were still on the order of a minute for 20 views, which is too slow for many applications.

Under the constraint of zero skew (which can always be assumed in practice) and known principal point (which can be guessed but is often not known exactly), semidefinite programming was used to globally minimize a heuristic cost function in Agrawal [31], which was extended with a brute force search for principal point in Agrawal [32]. These methods enforced the internal ADQ constraints, but neglected the constraints on aspect ratio and always assumed that principal point is constant, which makes them applicable to video but not photo collections.

Convex relaxation was used with branch and bound to identify the plane at infinity that globally minimizes a heuristic cost associated with the modulus constraint in the recent stratified approach of Chandraker et al. [33,34],but the heuristic is not ideal because it does not consider constraints on the DIAC in the search for the plane at infinity. Similar techniques were used to estimate the ADQ directly using all known constraints in Chandraker et al. [35], which makes it perhaps the most generally applicable global approach.

In general, the globally convergent approaches are very difficult to implement and not very efficient. As an alternative, the dual stratified approach of estimating the plane at infinity from known calibration matrix, first proposed in Bougnoux [14], has recently been revived with a closed form solution from a view pair in Gherardi and Fusiello [36]. The advantage of the dual stratified approach is that prior knowledge may be used to restrict the search into a very narrow plausible region, rather than exhaustively searching through all of parameter space for the plane at infinity. This leads to an algorithm that is simple, fast and robust. However, it lacks in precision, and still minimizes a heuristic objective that is not geometrically meaningful. As a result, attempting to further minimize the heuristic using nonlinear methods can result in divergence.

The fundamental limitation of all previous algorithms is that the objective being minimized is a heuristic with no particular geometric meaning, and these heuristics do not always work as well as one would hope. This becomes especially apparent for projective reconstructions with greater noise (or equivalently, more distant geometry) and for short reconstructions which are commonly on