

# A New Set of Quartic Trivariate Polynomial Equations for Stratified Camera Self-calibration under Zero-Skew and Constant Parameters Assumptions

Adlane Habed<sup>1</sup>, Kassem Al Ismaeil<sup>2</sup>, and David Fofi<sup>1</sup>

<sup>1</sup> Le2i UMR CNRS 6306

University of Bourgogne, Auxerre/Le Creusot, France

{Adlane.Habed,David.Fofi}@u-bourgogne.fr

<sup>2</sup> Interdisciplinary Centre for Security, Reliability and Trust

University of Luxembourg, Luxembourg

Kassem.Alismaeil@uni.lu

**Abstract.** This paper deals with the problem of self-calibrating a moving camera with constant parameters. We propose a new set of quartic trivariate polynomial equations in the unknown coordinates of the plane at infinity derived under the no-skew assumption. Our new equations allow to further enforce the constancy of the principal point across all images while retrieving the plane at infinity. Six such polynomials, four of which are independent, are obtained for each triplet of images. The proposed equations can be solved along with the so-called modulus constraints and allow to improve the performance of existing methods.

**Keywords:** Camera Self-calibration, Multiple View Geometry.

## 1 Introduction

In a typical multi-view uncalibrated Structure-from-Motion pipeline, the three-dimensional (3D) projective structure of an unknown static scene is upgraded to a metric one through the self-calibration of the imaging camera. Camera self-calibration consists in retrieving the camera's intrinsic parameters solely from point correspondences across images taken from distinct viewpoints without resorting to the presence of a calibration pattern in the scene. Most existing methods cope with the absence of such pattern by relying on the omnipresence of the so-called Absolute Conic (AC). The AC is a special virtual conic lying on the plane at infinity whose image (IAC) and, consequently, the dual of its image (DIAC) are solely dependent upon the parameters of the imaging camera. Once either the IAC or DIAC is retrieved, the intrinsic parameters can easily be obtained. However, locating either conics has turned out to be difficult and challenging due to the nonlinear nature of the problem on the one hand and, on the other hand, to the numerous critical motion sequences (CMS) [1] implying various levels of reconstruction ambiguities and dooming self-calibration to failure.

In order to help resolving such ambiguities, special care must be taken to: (a) enforce the uniqueness of the AC's supporting plane (i.e. the plane at infinity) across the image sequence; (b) exploit knowledge about the camera, particularly on the shape of its pixels, such as zero-skew and possibly known aspect ratio.

The uniqueness of the AC's supporting plane cannot be enforced through the direct and sole use of epipolar geometry as proposed in the early works [2] on camera self-calibration. Instead, the prior recovery of a projective (or quasi-affine [3]) structure of the scene, that is consistent with all cameras, is required. Based on such structure, both direct and stratified self-calibration methods have been proposed. The direct methods, which simultaneously estimate the plane at infinity and the camera parameters, rely on special virtual quadrics residing in 3-space and encoding either the IAC or DIAC along with the position of the plane at infinity itself. One such geometrical object that has been employed in this regard is the Absolute Dual Quadric (DAQ) [4,5]. The DAQ is a quadric of planes tangent to the AC and projecting onto DIACs across images. Another object consisting of all space lines tangent to the AC, the so-called Absolute Line Quadric (ALQ) [6,7], and projecting onto IACs across images, was also introduced in the context of camera self-calibration. Unlike the direct DAQ or ALQ-based approaches, the stratified self-calibration proceeds by first recovering the affine structure of the scene before upgrading it to a metric one. An affine reconstruction is obtained by identifying the plane at infinity. The IAC and/or DIAC, and therefore the parameters of the camera, can then be obtained by solving inherently linear equations. No tight constraints on the position of the plane at infinity exist in the case of a moving camera with varying parameters. In such case, iterative search schemes, alternating between plane and parameter estimation either relying on chirality constraints [8,9] or, more recently, the finiteness of the camera [10], have been employed. These methods can be considered as direct rather than stratified. However, when the parameters of the camera are constant, the plane at infinity can be located through a constraint on the moduli of the eigenvalues of its inter-image homography matrices [11]. This so-called "modulus constraint" leads to quartic trivariate polynomial equations that have been solved either through homotopy continuation or by means of a globally optimal method supported by the chirality constraints [12].

Because the resulting self-calibration equations explicitly involve the intrinsic parameters, knowledge about the camera has only been exploited when using direct methods. In this respect, the DAQ and ALQ-based self-calibration equations do not equally deal with the incorporation of assumptions such as zero-skew and/or known aspect ratio. For instance, because of a coupling involving the remaining unknown parameters in the DIAC, there is no simple way to exploit such assumptions when the DAQ-based self-calibration equations are used unless further knowledge on the principal point is taken into account. Indeed, it was shown in [13] that the problem can be linearized under the additional assumption of an image-centered principal point. However, several works have pointed out that the principal point may be significantly far from the center of the image and it is advised not to fix it while attempting to recover the remaining unknowns.

The self-calibration equations involving the ALQ (and hence the IAC) are better suited than their DAQ counterparts for incorporating the no-skew constraint and/or the known aspect ratio. These equations can indeed be linearized without making any assumption with respect to the position of the principal point. However, unlike their nonlinear counterparts, the linearized equations, whether based on the ALQ or the DAQ, may yield artificial CMSs [14] as they fail to enforce the positiveness and rank constraints [15] of their respective virtual objects. Furthermore, the methods based on such linearized equations are suitable for cameras with varying parameters. If the remaining unknown parameters are kept constant, enforcing their constancy again results in nonlinear equations.

In stratified camera self-calibration, based on the modulus constraints for instance, the polynomial equations used to locate the plane at infinity are derived under the sole assumption of constant parameters. Since they do not explicitly involve the camera parameters, knowledge about the camera cannot be directly exploited. As many as 64 solutions do satisfy the polynomial equations due to each triplet of images (although only 21 are relevant [16]). However, because the modulus constraint is only a necessary condition on the position of the plane at infinity, several planes, other than the plane at infinity, can persistently satisfy these constraints across several frames even when the motion is not critical. In the presence of more than one candidate plane at infinity, identifying the true plane can only be attempted a posteriori by inspecting the quasi-constancy and the shape of the pixels (skew and aspect ratio) of the linearly calculated intrinsic parameters. Note that such inspection only works when the motion of the camera is not critical since, in order to objectively assess the shape of the pixels, knowledge neither about the skew nor about the aspect ratio can be taken into account when recovering the parameters. However, when the camera motion is critical, the sought parameters cannot be estimated unless additional knowledge is considered. Assuming the plane at infinity has been uniquely identified, setting the skew to zero often suffices to recover the remaining camera parameters. When several planes are competing though and the motion is critical, setting the skew to zero with each candidate plane and estimating the remaining parameters makes it impossible to inspect the previously fixed parameter. This is a serious limitation as the skew factor is often the most reliable parameter to inspect. Another difficulty with such a posteriori inspection of the parameters lies in verifying the quasi-constancy of the parameters across the sequence as this is already enforced during the calculation of the parameters. In order to cope with these issues, knowledge about the camera must be exploited and the constancy of its parameters further enforced as early as when locating the plane at infinity, i.e. before estimating the camera parameters.

The work presented in this paper specifically deals with the problems of exploiting camera knowledge and enforcing parameters constancy while recovering the plane at infinity. The method we propose is along the lines of Pollefeys et al.'s method [11] based on the modulus constraint in the case of moving camera with constant parameters. Unlike their method, however, we additionally assume that the camera has zero-skew and exploit this practically valid assumption in

the recovery of the plane at infinity. Note that the latter assumption leads to particularly strong constraints on the camera since the class of transformations preserving the absence of skew was found in [13,17] to be the group of similarity transformations. We do not make any assumption on the more restrictive known aspect ratio of the pixels which would otherwise require a prior calibration of the camera. The main contribution of this paper is a new set of quartic trivariate polynomial equations in the unknown plane at infinity alone derived from triplets of images and hence similar to those obtained directly from the modulus constraint. Unlike the equations that are due to the modulus constraints however, our new equations can only be derived under the zero-skew assumption and further enforce the constancy of the coordinates of the principal point across triplets of images. Six such polynomials, four of which are independent, are obtained from each triplet. Along with the three quartic polynomials which are due to the modulus constraints in three views, they provide a total of nine polynomial equations in the unknown plane at infinity. The resulting equations can be solved using any of the numerical methods proposed for solving the modulus constraints equations. In the present work, we chose to solve our new equations along with those due to the modulus constraints using homotopy continuation. Such methods allow to calculate all possible solutions of a square system of polynomial equations. This choice was made to assess the benefit of incorporating these additional equations in terms of the number of candidates plane at infinity. Our experiments show that, in most cases, and even with only 3 images, these equations alone are sufficient to single out a unique solution and the performance of the stratified method is significantly improved. We also show that incorporating these new equations in the refinement of the position of the plane at infinity, through nonlinear least-squares minimization, allows to improve both the accuracy of the result and its robustness to noise.

## 2 Background and Related Work

In this paper we consider  $n$  images of a static scene captured by an uncalibrated moving camera with constant intrinsic parameters. At each imaging position  $i = 1 \dots n$ , the camera is given by its corresponding  $3 \times 4$  projection matrix  $\mathbf{M}_i$ . All projection matrices are assumed to be calculated in some arbitrary but common projective frame using point correspondences across images [9]. A 3D point, expressed in this frame by its homogeneous coordinate vector  $\mathbf{P}$ , projects onto the 2D point with coordinates  $\mathbf{p}_i \simeq \mathbf{M}_i \mathbf{P}$  in  $i^{th}$  image. The symbol  $\simeq$  is used throughout to denote equality up to an unknown non-zero scalar. We also denote by  $\mathbf{H}_{\Pi ij}$  the  $3 \times 3$  matrix representing a 2D homography induced by some plane  $\Pi$  between two images  $i$  and  $j$ . The 3-vector  $\mathbf{e}_{ij}$  represents the homogeneous coordinate vector of the projection of the optical center of camera  $j$  on the image plane of camera  $i$ . We assume the projective world reference frame attached to the first camera and hence the projection matrices of the form:

$$\mathbf{M}_i = [ \mathbf{H}_{\Pi 1i} \mid \mathbf{e}_{i1} ] \text{ for } i = 1 \dots n \quad (1)$$

where  $\mathbf{H}_{\Pi 11}$  simplifies to the identity matrix and  $\mathbf{e}_{11}$  to the null vector. The projections  $\mathbf{e}_{i1}$  are calculated up to compatible scales such that all projection matrices remain consistent with the same 3D structure. Furthermore, the homography matrices appearing in (1) and relating the reference and  $i^{th}$  views are all assumed to be induced by a unique, but arbitrarily chosen, plane  $\Pi$ .

Pollefeys et al.'s derivation of the modulus constraints on the plane at infinity [11] relies on two key ingredients. First, the  $3 \times 3$  matrix  $\mathbf{H}_{\infty 1i}$  of the homography induced by the plane at infinity between images 1 and  $i$  is given by

$$\mathbf{H}_{\infty 1i} \simeq \mathbf{H}_{\Pi 1i} + \mathbf{e}_{i1} \mathbf{a}_{\infty}^T \quad \text{for } i = 1 \dots n \tag{2}$$

for some appropriate, but generally unknown, 3-vector  $\mathbf{a}_{\infty}$ . Second, assuming a camera with constant parameters, the homography matrix  $\mathbf{H}_{\infty ij}$  between any two images is given, with the exact scale, by

$$\mathbf{H}_{\infty ij} = \mathbf{A} \mathbf{R}_{ij} \mathbf{A}^{-1} \tag{3}$$

which involves the camera's intrinsic parameter matrix  $\mathbf{A}$ , its inverse, and the rotation matrix  $\mathbf{R}_{ij}$  relating these two views.  $\mathbf{A}$  is parameterized as follows:

$$\mathbf{A} = \begin{bmatrix} f & \gamma & u_0 \\ 0 & \tau f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4}$$

where  $f$  is the focal length (in pixels) along the horizontal axis of the image,  $\tau$  is the aspect ratio,  $\gamma$  is the skew factor and  $(u_0, v_0)$  are the pixel coordinates of the principal point.  $\mathbf{H}_{\infty ij}$  being "similar" to a rotation matrix (see (3)), its eigenvalues must all have unit moduli. Because the matrix on the right-hand side of (2) is scaled, its eigenvalues are not unimodular but nevertheless have the particularity of having equal moduli, hence the "modulus constraint". Given the characteristic equation

$$\det(\mathbf{H}_{\Pi 1j} + \mathbf{e}_{j1} \mathbf{a}_{\infty}^T - \lambda(\mathbf{H}_{\Pi 1i} + \mathbf{e}_{i1} \mathbf{a}_{\infty}^T)) = c_{3ij} \lambda^3 + c_{2ij} \lambda^2 + c_{1ij} \lambda + c_{0ij} = 0, \tag{5}$$

the following condition

$$\mathcal{M}_{ij} \equiv c_{3ij} c_{1ij}^3 - c_{2ij}^3 c_{0ij} = 0 \text{ for all distinct } i, j = 1 \dots n \tag{6}$$

on  $\mathbf{a}_{\infty}$  has been found to be necessary for enforcing the equal moduli of the eigenvalues of each homography matrix  $\mathbf{H}_{\infty ij} \simeq (\mathbf{H}_{1j} + \mathbf{e}_{j1} \mathbf{a}_{\infty}^T)(\mathbf{H}_{1i} + \mathbf{e}_{i1} \mathbf{a}_{\infty}^T)^{-1}$  relating any two distinct views. For clarity of the exposition, the symbol  $\equiv$  is used here (and throughout the paper) to rename the expression on its right-hand side by the notation on its left-hand side. The coefficients  $c_{kij}$ ,  $k = 0, 1, 2$  and  $3$  of the characteristic equation (5) being affine functions of the sought entries of  $\mathbf{a}_{\infty}$ , each resulting equation (6) is a polynomial of degree 4 in 3 unknowns. Three such equations can be obtained as soon as three images are available. A system of polynomial equations can then be solved using a continuation method to recover all possible solutions (i.e. at most 64 solutions in this case) from which

complex ones can be discarded. More than three images are generally used in order to isolate a unique solution which can be refined through the nonlinear least-squares minimization of an appropriate cost function based on (6) and all pairs of images. Once the plane at infinity recovered and all inter-image homographies calculated, either the IAC or DIAC can be linearly retrieved thus providing an initial estimate of the intrinsic parameters. These parameters along with the plane at infinity can further be refined through nonlinear least-squares. Readers unfamiliar with the details of this approach may refer to [11].

### 3 Homography Matrix Inversion

It is well known that the relationship between inter-image homographies (2) does not specifically characterize the sole homography that is induced by the plane at infinity. On the contrary, the homography induced by any plane  $\Phi$  between the reference and some  $i^{th}$  image can be obtained through such relationship, i.e.

$$\mathbf{H}_{\Phi i} \simeq \mathbf{H}_{\Pi i} + \mathbf{e}_{i1} \mathbf{a}_{\Phi}^T \quad \text{for } i = 1 \dots n \tag{7}$$

for some given 3-vector  $\mathbf{a}_{\Phi}$ . While Pollefeys et al.'s derivation of the modulus constraints relies on this property of inter-image homographies, the new equations proposed in this paper make use of another - equally important - property which seems to have been overlooked in the literature. For instance, the property that all homography matrices  $\mathbf{H}_{\Phi ij}$  (including their respective inverse homographies  $\mathbf{H}_{\Phi ji}$ ), regardless of which views are involved, can also be affine functions of the same vector  $\mathbf{a}_{\Phi}$  appearing in (7). For the inverse homographies relating some image  $i$  to the reference one, this relationship, which will be demonstrated farther below in this section, is given by:

$$\mathbf{H}_{\Phi i1} \simeq \mathbf{H}_{\Pi i}^* + [\mathbf{a}_{\Phi}]_{\times} \mathbf{H}_{\Pi i}^T [\mathbf{e}_{i1}]_{\times}^T \tag{8}$$

where  $[\cdot]_{\times}$  denotes the  $3 \times 3$  skew-symmetric matrix constructed from its 3-vector argument and associated with the cross-product. The matrix  $\mathbf{H}_{\Pi i}^*$  refers to the adjoint matrix of  $\mathbf{H}_{\Pi i}$  defined as the transpose of the matrix of co-factors.  $\mathbf{H}_{\Pi i}^*$  can thus be expressed symbolically as well as obtained numerically through its relationship to the inverse of  $\mathbf{H}_{\Pi i}$ , i.e.

$$\mathbf{H}_{\Pi i}^* = \det(\mathbf{H}_{\Pi i}) \mathbf{H}_{\Pi i}^{-1} \tag{9}$$

If ever employed, expressing homography matrix inversion through (8) is rather uncommon. To express such inversion, one may naturally want to consider using the more "popular" relationship (7) leading to  $\mathbf{H}_{\Phi i1} \simeq \mathbf{H}_{\Pi i}^{-1} + \mathbf{e}_{i1} \mathbf{b}_{\Phi}^T$ . This, however, not only introduces a new vector  $\mathbf{b}_{\Phi} \neq \mathbf{a}_{\Phi}$  for the considered image  $i$ , but also a different vector with every new image. This obviously renders exploiting such inverse homography expression rather complicated. As it turns out, expressing both  $\mathbf{H}_{\Phi i1}$  and  $\mathbf{H}_{\Phi 1i}$  in terms of a unique vector  $\mathbf{a}_{\Phi}$  is especially advantageous. In particular, the entries of a homography induced by  $\Phi$  between any

two images  $i$  and  $j$ , along with its inverse homography, can also be expressed as affine functions of this vector. For instance, because  $\mathbf{H}_{\Phi ij} \simeq \mathbf{H}_{\Phi 1j} \mathbf{H}_{\Phi i1}$  and the fact that  $[\mathbf{a}_{\Phi}]_{\times} \mathbf{a}_{\Phi} = \mathbf{0}$ , using (7) and (8) leads, on the one hand, to the expression

$$\mathbf{H}_{\Phi ij} \simeq \mathbf{H}_{\Pi 1j} (\mathbf{H}_{\Pi 1i}^* + [\mathbf{a}_{\Phi}]_{\times} \mathbf{H}_{\Pi 1i}^T [\mathbf{e}_{i1}]_{\times}^T) + \mathbf{e}_{j1} \mathbf{a}_{\Phi}^T \mathbf{H}_{\Pi 1i}^* \text{ for all } i \neq j \quad (10)$$

and, on the other hand, to its inverse expression

$$\mathbf{H}_{\Phi ji} \simeq \mathbf{H}_{\Pi 1i} (\mathbf{H}_{\Pi 1j}^* + [\mathbf{a}_{\Phi}]_{\times} \mathbf{H}_{\Pi 1j}^T [\mathbf{e}_{j1}]_{\times}^T) + \mathbf{e}_{i1} \mathbf{a}_{\Phi}^T \mathbf{H}_{\Pi 1j}^* \text{ for all } i \neq j. \quad (11)$$

Note that (10) and (11) are general enough to include the case where either  $i$  or  $j$  can be the reference image. In such case, it is enough to consider  $\mathbf{H}_{\Pi 11}$  as the identity matrix and  $\mathbf{e}_{11}$  as the null vector in which case (10) and (11) simplify to (7) and (8). Also, because these relationships have been obtained for an arbitrary plane  $\Phi$ , they also hold when  $\Phi$  coincides with the plane at infinity.

The rest of this section is dedicated to the derivation of the property (8) relating inverse homographies. Any image point  $p$  verifying  $\mathbf{a}_{\Phi}^T \mathbf{p} = 0$  is equally mapped by both homographies  $\mathbf{H}_{\Pi 1i} \mathbf{p} \simeq \mathbf{H}_{\Phi 1i} \mathbf{p}$  onto the same point in the next image. This shows that  $\mathbf{a}_{\Phi}$  represents the coordinate vector of a line in the first image whose 3D counterpart is the line, say  $L$ , at the intersection of  $\Phi$  and the reference plane  $\Pi$ . Consider now a plane  $\Psi$ , other than  $\Phi$  and the reference frame  $\Pi$ , containing  $L$ , hence linearly dependent upon  $\Pi$  and  $\Phi$ . The homography,  $\mathbf{H}_{\Psi 1i} \simeq \mathbf{H}_{\Pi 1i} + \mathbf{e}_{i1} \mathbf{a}_{\Psi}^T$  induced by  $\Psi$ , is such that  $\mathbf{a}_{\Psi} \simeq \mathbf{a}_{\Phi}$ . A direct consequence is that the homography of any of the three planes can be linearly obtained through the homographies induced by the two other planes under the assumption of the three planes being linearly dependent. Note that this result is independent of the choice of the reference image. One can therefore also deduce that the inverse homographies induced by these three planes are also linearly dependent. Hence, for some given image  $i$ , a non-zero scalar  $\mu$  exists such that  $\mathbf{H}_{\Phi i1} \simeq \mathbf{H}_{\Pi i1} + \mu \mathbf{H}_{\Psi i1}$ . Since the only requirement on the plane  $\Psi$  is that it contains the space line  $L$ , we are free to choose this plane as passing through the optical center of the first camera and intersecting image 1 exactly onto the line with coordinates  $\mathbf{a}_{\Phi}$ . In such case, the homography matrix of this plane is of the form (see for instance [9], page 334)  $\mathbf{H}_{\Psi i1} \simeq [\mathbf{a}_{\Phi}]_{\times} \mathbf{H}_{\Pi 1i}^T [\mathbf{e}_{i1}]_{\times}^T$ . Furthermore, because  $\mathbf{H}_{\Pi 1i} \simeq \mathbf{H}_{\Pi 1i}^*$ , one may deduce that  $\mathbf{H}_{\Phi i1}$  is of the form  $\mathbf{H}_{\Phi i1} \simeq \mathbf{H}_{\Pi 1i}^* + \mu [\mathbf{a}_{\Phi}]_{\times} \mathbf{H}_{\Pi 1i}^T [\mathbf{e}_{i1}]_{\times}^T$  for some  $\mu$ . It can be verified through expansion, typically using a symbolic processor like *Maple*, that the product  $(\mathbf{H}_{\Pi 1i} + \mathbf{e}_{i1} \mathbf{a}_{\Phi}^T)(\mathbf{H}_{\Pi 1i}^* + \mu [\mathbf{a}_{\Phi}]_{\times} \mathbf{H}_{\Pi 1i}^T [\mathbf{e}_{i1}]_{\times}^T)$  simplifies, up to a scale, to identity for the value  $\mu = 1$ . Hence, the expression of  $\mathbf{H}_{\Pi i1}$  indeed corresponds to (8).

## 4 A New Set of Quartic Polynomial Equations

The new set of affine self-calibration equations presented in this paper is based on the following observation: the difference between the matrix of the inter-image homography induced by the plane at infinity between two images and that of

its inverse results in a matrix that is "similar" to a skew-symmetric matrix. That is, given the exact expression of  $\mathbf{H}_{\infty ij}$  as in (3) and that of its inverse  $\mathbf{H}_{\infty ji} = \mathbf{A}\mathbf{R}_{ij}^T\mathbf{A}^{-1}$ , we have

$$\mathbf{H}_{\infty ij} - \mathbf{H}_{\infty ji} = \mathbf{A}[\mathbf{r}_{ij}]_{\times}\mathbf{A}^{-1} \text{ where } [\mathbf{r}_{ij}]_{\times} = \mathbf{R}_{ij} - \mathbf{R}_{ij}^T. \tag{12}$$

Because (12) is true for all pairs of views and it involves both a homography and its inverse, it appears particularly judicious to resort at this point to the right-hand sides of (10) and (11) to express  $\mathbf{H}_{\infty ij}$  and  $\mathbf{H}_{\infty ji}$ . This indeed allows to express all homography matrices using a unique vector  $\mathbf{a}_{\infty}$  for all pairs of images. We will refer to these matrix expressions as  $\tilde{\mathbf{H}}_{\infty ij}$  and  $\tilde{\mathbf{H}}_{\infty ji}$  such that

$$\tilde{\mathbf{H}}_{\infty ij} \equiv \mathbf{H}_{\Pi 1j}(\mathbf{H}_{\Pi 1i}^* + [\mathbf{a}_{\infty}]_{\times}\mathbf{H}_{\Pi 1i}^T[\mathbf{e}_{i1}]_{\times}^T) + \mathbf{e}_{j1}\mathbf{a}_{\infty}^T\mathbf{H}_{\Pi 1i}^* \text{ for all } i \neq j \tag{13}$$

and

$$\tilde{\mathbf{H}}_{\infty ji} \equiv \mathbf{H}_{\Pi 1i}(\mathbf{H}_{\Pi 1j}^* + [\mathbf{a}_{\infty}]_{\times}\mathbf{H}_{\Pi 1j}^T[\mathbf{e}_{j1}]_{\times}^T) + \mathbf{e}_{i1}\mathbf{a}_{\infty}^T\mathbf{H}_{\Pi 1j}^* \text{ for all } i \neq j. \tag{14}$$

Unlike  $\mathbf{H}_{\infty ij}$  and  $\mathbf{H}_{\infty ji}$  which are exact expressions,  $\tilde{\mathbf{H}}_{\infty ij}$  and  $\tilde{\mathbf{H}}_{\infty ji}$  represent these homographies up to unknown scales turning (12) into

$$\lambda_{ij}\tilde{\mathbf{H}}_{\infty ij} - \tilde{\mathbf{H}}_{\infty ji} \simeq \mathbf{A}[\mathbf{r}_{ij}]_{\times}\mathbf{A}^{-1}. \tag{15}$$

Although the scalar  $\lambda_{ij}$  is unknown, it can in turn be expressed in terms of the sought vector  $\mathbf{a}_{\infty}$ . This can be done by noticing that  $trace(\lambda_{ij}\tilde{\mathbf{H}}_{\infty ij} - \tilde{\mathbf{H}}_{\infty ji}) = trace(\mathbf{A}[\mathbf{r}_{ij}]_{\times}\mathbf{A}^{-1}) = 0$ ; a consequence of which  $\lambda_{ij}trace(\tilde{\mathbf{H}}_{\infty ij}) = trace(\tilde{\mathbf{H}}_{\infty ji})$ . This allows to eliminate  $\lambda_{ij}$  from the left-hand side of (15). It can now be stated that the entries of the following matrix expression,

$$\mathbf{Q}_{\infty ij} \equiv trace(\tilde{\mathbf{H}}_{\infty ji})\tilde{\mathbf{H}}_{\infty ij} - trace(\tilde{\mathbf{H}}_{\infty ij})\tilde{\mathbf{H}}_{\infty ji} \tag{16}$$

are quadratic functions of  $\mathbf{a}_{\infty}$ , and that  $\mathbf{Q}_{\infty ij}$  is similar to  $[\mathbf{r}_{ij}]_{\times}$  since

$$\mathbf{Q}_{\infty ji} \simeq \mathbf{A}[\mathbf{r}_{ij}]_{\times}\mathbf{A}^{-1}. \tag{17}$$

Under the zero-skew assumption, the parameter  $\gamma$  in the intrinsic parameters matrix  $\mathbf{A}$  (given by (4)) reduces to nullity. This assumption induces a particular structure to the matrix  $\mathbf{A}[\mathbf{r}_{ij}]_{\times}\mathbf{A}^{-1}$  and hence to  $\mathbf{Q}_{\infty ij}$ . That is,

$$\mathbf{Q}_{\infty ij} \simeq \begin{bmatrix} -\frac{r_2 u_0}{f} & -\frac{r_3}{\tau} + \frac{u_0 r_1}{\tau f} & f\left(\frac{r_3 v_0}{\tau f} + r_2\right) + u_0\left(\frac{r_2 u_0}{f} - \frac{r_1 v_0}{\tau f}\right) \\ \tau r_3 - \frac{v_0 r_2}{f} & \frac{r_1 v_0}{\tau f} & \tau f\left(-\frac{r_3 u_0}{f} - r_1\right) + v_0\left(\frac{r_2 u_0}{f} - \frac{r_1 v_0}{\tau f}\right) \\ -\frac{r_2}{f} & \frac{r_1}{\tau f} & \frac{r_2 u_0}{f} - \frac{r_1 v_0}{\tau f} \end{bmatrix} \tag{18}$$

where  $\mathbf{r}_{ij} = (r_1, r_2, r_3)^T$ . This allows one to extract the exact expressions of the coordinates  $(u_0, v_0)$  of the principal point in terms of the entries of  $\mathbf{Q}_{\infty ij}$ :

$$u_0 = \frac{(\mathbf{Q}_{\infty ij})_{11}}{(\mathbf{Q}_{\infty ij})_{31}} \text{ and } v_0 = \frac{(\mathbf{Q}_{\infty ij})_{22}}{(\mathbf{Q}_{\infty ij})_{32}} \tag{19}$$



where the subscripts  $(\mathbf{Q}_{\infty ij})_{rc}$  refers to the entry at the  $r^{th}$  row and  $c^{th}$  column of  $\mathbf{Q}_{\infty ij}$ . Given any two pairs of images  $(i, j)$  and  $(k, \ell)$ , we have

$$\frac{(\mathbf{Q}_{\infty ij})_{11}}{(\mathbf{Q}_{\infty ij})_{31}} = \frac{(\mathbf{Q}_{\infty k\ell})_{11}}{(\mathbf{Q}_{\infty k\ell})_{31}} \text{ and } \frac{(\mathbf{Q}_{\infty ij})_{22}}{(\mathbf{Q}_{\infty ij})_{32}} = \frac{(\mathbf{Q}_{\infty k\ell})_{22}}{(\mathbf{Q}_{\infty k\ell})_{32}} \tag{20}$$

which enforce the constancy of the coordinates  $(u_0, v_0)$  of the principal point across the considered images. This in turn leads to the quartic polynomials

$$\begin{aligned} \mathcal{U}_{ijkl} &\equiv (\mathbf{Q}_{\infty ij})_{11}(\mathbf{Q}_{\infty k\ell})_{31} - (\mathbf{Q}_{\infty k\ell})_{11}(\mathbf{Q}_{\infty ij})_{31} = 0 \\ \mathcal{V}_{ijkl} &\equiv (\mathbf{Q}_{\infty ij})_{22}(\mathbf{Q}_{\infty k\ell})_{32} - (\mathbf{Q}_{\infty k\ell})_{22}(\mathbf{Q}_{\infty ij})_{32} = 0. \end{aligned} \tag{21}$$

Equations (21) are our new quartic polynomials that can be solved along with the modulus constraints (6) through homotopy continuation to find all possible solutions. They can also be used within an appropriate cost function to refine the resulting solution. Unlike [12], we do not propose in this paper a new method for solving such quartics but rather follow Pollefeys and al.’s algorithm. The globally optimal method [12] may be adapted to take into account our new equations. Typically, when using homotopy continuation, the images can be considered by triplets each of which providing three pairs of images. Therefore, six new polynomial equations are obtained from (21), four of which are independent, and three equations are directly obtained from the modulus constraints (6). For each triplet of images, our new equations along with the ones due to the modulus constraints are divided into 3 groups of 3 equations. Each such group is solved by continuation and the common solutions are isolated. More than three images may be required to isolate a unique solution which can then be refined through nonlinear least-squares. The following cost function can be minimized:

$$\mathcal{C}(\mathbf{a}_{\infty}) = \sum_{ij} \mathcal{M}_{ij}^2 + w \sum_{ij} \sum_{k\ell} (\mathcal{U}_{ijkl}^2 + \mathcal{V}_{ijkl}^2) \tag{22}$$

for a fixed value of the weight  $w$ . It is advised however to use a normalized version of this cost function by replacing  $\tilde{\mathbf{H}}_{\infty ij}$  by  $\tilde{\mathbf{H}}_{\infty ij}/\det(\tilde{\mathbf{H}}_{\infty ij})^{1/3}$  in all  $\mathcal{M}_{ij}$ ,  $\mathcal{U}_{ijkl}$  and  $\mathcal{V}_{ijkl}$ . In order to incorporate our new constraints, a judicious choice of  $w \neq 0$  must be made. While each  $\mathcal{M}_{ij}$  involves two images, each individual  $\mathcal{U}_{ijkl}$  or  $\mathcal{V}_{ijkl}$  involves 3 and possibly 4 images. Given  $n$  images, the modulus constraints would provide  $w_1 = n(n - 1)/2$  equations. Considering the images per triplets, our new constraints provide  $6w_2$  equations where  $w_2 = n!/(3!(n - 3)!)$  is the number of triplets. It is clear that the number of triplets grows much faster than the number of pairs of images. Our new equations may thus introduce a great deal of redundancy which, while good to stabilize the final solution, tends to make the contribution of the  $\mathcal{M}_{ij}$  terms insignificant (with the increasing number of images) in the total cost  $\mathcal{C}(\mathbf{a}_{\infty})$ . In order to reflect the fact that each triplet of images provides 4 independent equations out of the 6 new equations (21) and that the modulus constraints (6) provides only 3, we have chosen  $w$  such that  $w = \frac{2w_1}{9w_2}$  in our experiments. Once  $\mathbf{a}_{\infty}$  is obtained, an estimate of the intrinsic parameters can be calculated by solving an overdetermined linear system of equations through SVD. This estimate can be further refined along with  $\mathbf{a}_{\infty}$  using nonlinear least-squares minimization (refer to [11] for details).

## 5 Experiments

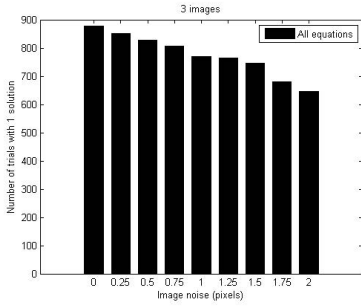
We have tested the method described in Section 4 through several experiments including both synthetic and real images. The results we have obtained are very satisfactory. In all our experiments, the projective projection matrices have been calculated using [18]. The polynomial systems of equations have been solved using the PhcPack solver [19] and data normalization has been used throughout.

### 5.1 Simulations

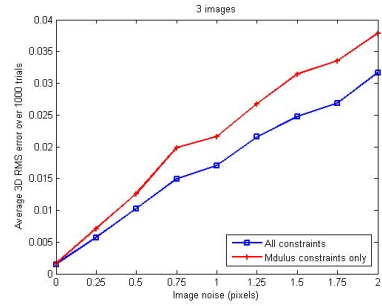
Each simulated scene consisted of 100 randomly generated points scattered within the unitary sphere. For each scene a number of cameras, roughly oriented toward the scene, were placed at random with a mean distance of 2 units from the center of the sphere and 0.3 standard deviation. The scene points have then been projected onto  $512 \times 512$  images with intrinsic parameters  $f = 800$ ,  $\gamma = 0$ ,  $\tau = 1$  and  $u_0 = v_0 = 256$ . New scene and cameras have been generated for various levels of noise added to the pixel coordinates. The added noise was obtained from zero-mean Gaussian distribution with a standard deviation in the range  $0 \dots 2$  pixels (with a step of 0.25). We have run 1000 independent trials for each level of noise and sequence length.

The first objective of these experiments was to assess the benefit of using the new equations along with those due to the modulus constraints against using the latter equations alone. We have particularly considered the case of short sequences, typically 3 and 4 images. For each trial, we have solved the available polynomial equations due to triplets of images through homotopy continuation. We report in Fig. 1(a) the number of trials with 3 images for which a unique solution was found to simultaneously satisfy our new equations (21) and those due to the modulus constraints (6). Note that these results have been obtained without resorting to the calculation and examination of the intrinsic parameters. In such case, the modulus constraints alone hardly provide a unique solution (at most in 2 trials out of 1000) unless the camera parameters are examined. The same experiments were conducted with 4 images (see Fig. 2(a)) and the number of trials leading to a unique solution without further examination has been seen to peak when all polynomial equations are considered even with high levels of noise. It can also be seen that, with 4 images, isolating a unique solution when using the modulus constraints alone becomes increasingly difficult with noise.

The second objective of these experiments was to initialize two cost functions using the solution obtained when solving our new set of equations along with the modulus constraints: one cost function in which all the constraints are considered, hence (22) with  $w \neq 0$  as given in Section 4, while the second function only involves the modulus constraints, i.e. (22) with  $w = 0$ . In each case, the cost function was minimized using the Levenberg-Marquardt algorithm. These experiments have allowed us, on the one hand, to verify that the results obtained through homotopy continuation are indeed correct and, on the other hand, to compare the quality of the reconstruction obtained by each method. The refined solutions were then used to linearly calculate the intrinsic parameters of the

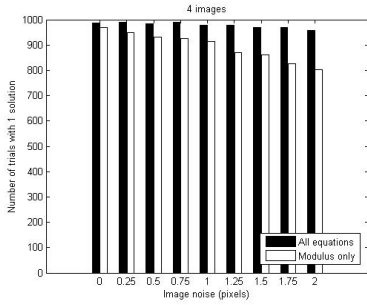


(a) Trials leading to a 1 solution

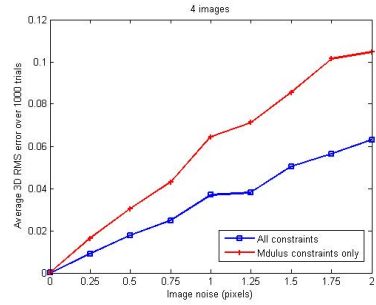


(b) Avg. 3D RMS error vs. image noise

Fig. 1. Experiments with 3 images

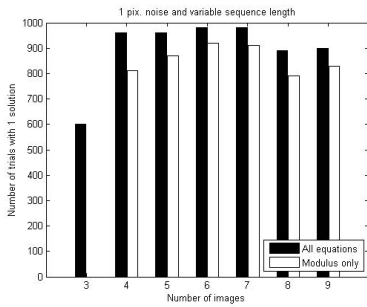


(a) Trials leading to 1 solution

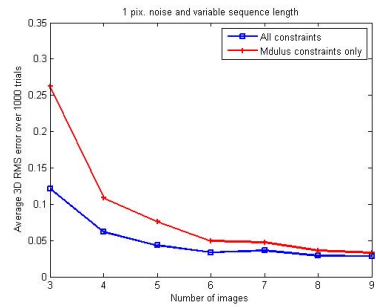


(b) Avg. 3D RMS error vs. image noise

Fig. 2. Experiments with 4 images

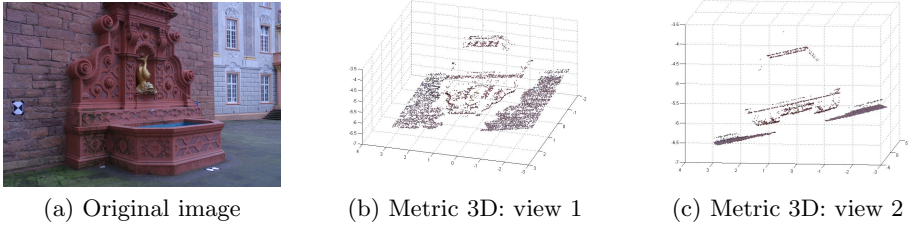


(a) Trials leading to 1 solution

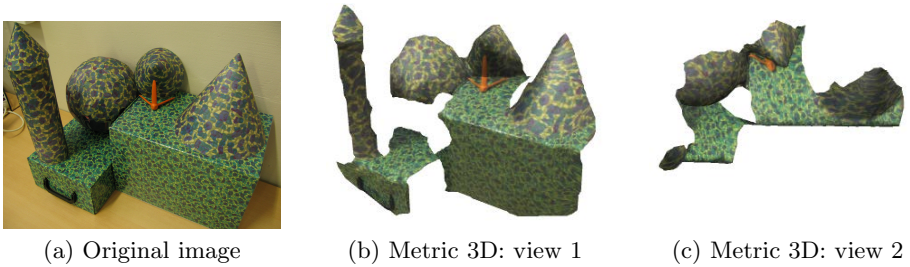


(b) Avg. 3D RMS error vs. image noise

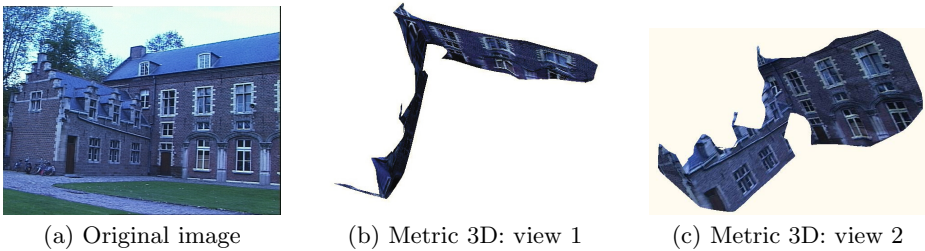
Fig. 3. Experiments with 1 pixel of noise



**Fig. 4.** The fountain sequence: 3 images and 6800 matched points



**Fig. 5.** The cone sequence: from 4 images and 1302 matched



**Fig. 6.** The castle sequence: 5 images and 212 matched points

camera. Note that the results we report here are not followed by any refinement of the camera parameters. Using these parameters, the metric structure of the scene is recovered and aligned with the original point cloud through the best (in the least-squares sense) metric transformation. The average 3D RMS errors over all the trials are given in Fig. 1(b) for 3 images and Fig. 2(b) for 4 images. Because optimization has been identically initialized for both cost functions, the advantage of incorporating the new constraints is made clear by these 3D results.

Finally, we have run the same experiments with 1 pixel of noise while the number of images has been varied from 3 to 9 with a unit step. The results, similar to those conducted with 3 and 4 images, are reported in Fig. 3. These suggest that the reconstruction errors using both cost functions tend to converge

while decreasing with the number of images. The new polynomial equations, however, keep contributing more efficiently than their modulus counterparts in isolating a unique solution even for as many as 9 images.

Note that the average running time for solving 1 set of equations on a i7-2.2GHz-4Go was 1.2 seconds. Solving 3 such equations takes 3 times as much which, considering the whole Structure-from-Motion pipeline, is acceptable.

## 6 Real Images

We present here three examples with real images in which we have deliberately used short image sequences to test the effectiveness of our newly proposed method. For each scene, we have detected and matched points across images using a SIFT-based algorithm implemented in-house. The new polynomial equations along with the modulus constraints were solved using homotopy continuation. A unique solution was isolated without any further inspection of the parameters in all three examples. The solutions have been refined using (22) with all available constraints. The intrinsic parameters have been linearly calculated. The result is a 3D metric reconstruction ready for visual assessment. Only 3 images but as many as 6800 matched points have been used in the Fountain example (obtained from <http://www.robots.ox.ac.uk>) in Fig. 4. However, 4 images and 1302 corresponding points have been used in the Cone example in Fig. 5 and 5 images, but only 212 points, in the Castle sequence (obtained from <http://www.cs.unc.edu>) in Fig. 6. Only views of the reconstructed point cloud are provided in the Fountain example (Fig. 4(b)-(c)). In the two remaining examples, texture was mapped from the images onto the triangulated point clouds (Fig. 5(b)-(c) and Fig. 6(b)-(c)). It can be seen that the quality of the reconstruction is satisfactory in all three examples.

## 7 Conclusion

We have proposed a new set of quartic polynomial equations in the unknown coordinates of the plane at infinity. These new equations have been derived under the no-skew assumption of a moving camera with constant parameters. They can be efficiently used in conjunction with the modulus constraints whether to be solved through numerical optimization methods or by continuation. Our experiments show that these equations are not only useful when only few images are available but tend also to improve the reconstruction quality and the success rate of camera self-calibration even for longer image sequences.

## References

1. Sturm, P.: Critical Motion Sequences for Monocular Self-calibration and Uncalibrated Euclidean Reconstruction. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1100–1105 (1997)

2. Luong, Q.: Self-calibration of a Moving Camera from Point Correspondences and Fundamental Matrices. *International Journal of Computer Vision* 22, 261–289 (1997)
3. Nistér, D.: Untwisting a Projective Reconstruction. *International Journal of Computer Vision* 60, 165–183 (2004)
4. Triggs, B.: Autocalibration and the Absolute Quadric. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 609–614 (1997)
5. Heyden, A., Åström, K.: Euclidean Reconstruction from Constant Intrinsic Parameters. In: *International Conference on Pattern Recognition*, vol. 1, pp. 339–343. IEEE (1996)
6. Valdés, A., Ronda, J.I., Gallego, G.: The Absolute Line Quadric and Camera Autocalibration. *International Journal of Computer Vision* 66, 283–303 (2006)
7. Ponce, J., Mc Henry, K., Papadopoulos, T., Teillaud, M., Triggs, B.: On the Absolute Quadratic Complex and its Application to Autocalibration. In: *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 780–787 (2005)
8. Hartley, R., Hayman, E., Agapito, L., Reid, I.: Camera Calibration and the Search for Infinity. In: *IEEE International Conference on Computer Vision*, pp. 510–517 (1999)
9. Hartley, R.I., Zisserman, A.: *Multiple View Geometry in Computer Vision*, 2nd edn. Cambridge University Press (2004)
10. Gherardi, R., Fusiello, A.: Practical Autocalibration. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) *ECCV 2010, Part I. LNCS*, vol. 6311, pp. 790–801. Springer, Heidelberg (2010)
11. Pollefeys, M., van Gool, L.: Stratified Self-calibration with the Modulus Constraint. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21, 707–724 (1999)
12. Chandraker, M., Agarwal, S., Kriegman, D., Belongie, S.: Globally Optimal Algorithms for Stratified Autocalibration. *International Journal of Computer Vision* 90, 236–254 (2010)
13. Pollefeys, M., Koch, R., van Gool, L.: Self-calibration and Metric Reconstruction in spite of Varying and Unknown Internal Camera Parameters. *International Journal of Computer Vision* 32, 7–25 (1999)
14. Gurdjos, P., Bartoli, A., Sturm, P.F.: Is Dual Linear Self-calibration Artificially Ambiguous? In: *IEEE International Conference on Computer Vision*, pp. 88–95 (2009)
15. Chandraker, M., Agarwal, S., Kahl, F., Kriegman, D., Nister, D.: Autocalibration via Rank-Constrained Estimation of the Absolute Quadric. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8 (2007)
16. Schaffalitzky, F.: Direct Solution of Modulus Constraints. In: *Indian Conference on Computer Vision, Graphics and Image Processing*, pp. 314–321 (2000)
17. Heyden, A., Astrom, K.: Flexible Calibration: Minimal Cases for Auto-calibration. In: *IEEE International Conference on Computer Vision*, vol. 1, pp. 350–355 (1999)
18. Rothwell, C.A., Faucher, O.D., Csürka, G.: Different Paths towards Projective Reconstruction. In: *Europe-China Workshop on Geometrical Modelling and Invariants for Computer Vision*. Xidan University Press (1995)
19. Verschelde, J.: Polynomial Homotopy Continuation with PhcPack. *ACM Communications in Computer Algebra* 44, 217–220 (2011)