Image and Vision Computing 20 (2002) 415–426

www.elsevier.com/locate/imavis

Critical motion sequences for the self-calibration of cameras and stereo systems with variable focal length

Peter Sturm*

INRIA Rhône-Alpes, 655 Avenue de l'Europe, 38330 Montbonnot St Martin, France

Received 16 October 2000; accepted 18 December 2001

Abstract

We consider the self-calibration problem for a moving camera whose intrinsic parameters are known except for the focal length which may vary freely across different views. The conditions, under which the determination of the focal length's values for an image sequence is not possible, are derived. These depend only on the camera's motion. We give a complete catalogue of the so-called *critical motion sequences*. This is then used to derive the critical motion sequences for stereo systems with variable focal lengths. © 2002 Published by Elsevier Science B.V.

1. Introduction

One of the major goals of computer vision is the recovery of spatial information about the environment. Classical approaches assume that the cameras are *calibrated* beforehand but a great interest in *uncalibrated* vision and on-line calibration has arisen during the last decade. A key result is that even with completely uncalibrated cameras, spatial information—*projective structure*—can be obtained: the scene can be reconstructed up to an unknown projective transformation [6,9]. Furthermore, a moving camera can *self-calibrate*, i.e. the calibration parameters can be estimated solely from feature correspondences between several images [11,13]. This allows the projective ambiguity in the reconstruction to be reduced to a Euclidean one (up to a similarity transformation) and we speak of *uncalibrated Euclidean reconstruction*.

It is known that several types of camera motion prevent self-calibration, i.e. the calibration parameters cannot be determined uniquely. Accordingly, Euclidean structure cannot be obtained, although reconstruction at some level between projective and Euclidean is generally possible. These degeneracies are inherent, i.e. they cannot be resolved by any algorithm without additional knowledge. Sequences of camera motions that imply such degeneracies will be referred to as *critical motion sequences*. By 'sequences' we mean that not only the motion between two successive views but that over a complete image sequence is critical.

For the basic self-calibration scenario, a moving camera with *fixed* calibration, we derived the critical motion sequences in [17,18]. In this paper, we study the case of a moving camera with *variable* and unknown focal length, but whose other intrinsic parameters are known. A practical self-calibration algorithm was proposed by Azarbayejani and Pentland [1]. Algorithms and closed-form solutions for the two-view case are given, e.g. in Refs. [3–5,8,12]. Newsam et al. derived the critical motions for the two-view case [12]. In this paper, we derive a complete characterization of critical motion *sequences* for any number of views and the critical motions for stereo systems. This paper is an extended version of [19].

The paper is organized as follows. In Section 2 we provide some theoretical background for our approach. The problem of deriving critical motion sequences is formulated in Section 3. The critical motion sequences are derived in Section 4. A summary of the derivations is given in Section 5 and comments are made in Section 6. The critical motions for stereo systems are derived in Section 7 and conclusions are drawn in Section 8.

2. Background

The definitions in this section are mainly taken from Refs. [2,16]. Some of the results for general quadrics are presented only for central conics.

2.1. Notation

We refer to the *plane at infinity* as the *ideal plane* and

* Tel.: +33-4-76-61-52-32; fax: +33-4-76-61-54-54.
E-mail address: peter.sturm@inrialpes.fr (P. Sturm).

0262-8856/02/$ - see front matter © 2002 Published by Elsevier Science B.V.
PII: S0262-8856(02)00012-4

denote it by Π_∞. $\mathscr{P}^n$ is the n-dimensional projective space and $\sim$ means equality up to a scalar factor accounting for the use of homogeneous coordinates. We use the abbreviation PVC for proper virtual conics (see description later).

2.2. Pinhole camera model

We use perspective projection to model cameras. A projection may be represented by a 3×4 projection matrix P that maps points of 3-space to points in 2-space: $q \sim PQ$. We consider only the case of perfect perspective projection, i.e. the projection center does not lie on Π_∞.

With regard to physical cameras, the projection matrix may be decomposed into a *calibration matrix K* and a *pose matrix*. The pose matrix represents the position and orientation of the camera with respect to some world coordinate frame. In general, we distinguish five intrinsic parameters for the perspective projection model: the (effective) focal length f, the aspect ratio τ, the principal point (u_0, v_0) and a skew factor accounting for non rectangular pixels. The skew factor is usually very close to 0 and we ignore it in the following. The calibration matrix may be written as:

$$K = \begin{pmatrix} \tau f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

We decompose the projection matrix as follows:

$$P = K \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} R & -Rt \\ 0^{\mathrm{T}} & 1 \end{pmatrix}}_{P_c}$$

The matrix P_c is the *canonical projection* and we call its destination the *metric image plane*. The canonical projection depends only on the camera's extrinsic parameters—a rotation matrix R representing its orientation and a 3-vector t representing its position. The calibration matrix K describes an invertible affine transformation from the metric image plane to pixel coordinates.

2.3. Quadrics and conics

A *quadric* in $\mathscr{P}^n$ is a set of points satisfying a quadratic equation in their homogeneous coordinates. Each quadric can be represented by a symmetric $(n+1) \times (n+1)$ matrix. A *proper quadric* is a quadric whose matrix has a non zero determinant. *Conics* are planar quadrics; we will not distinguish between a conic and its matrix. A conic in $\mathscr{P}^3$ or *3D conic* is defined by its *supporting plane* and the conic's equation in that plane.

2.4. Virtual quadrics

A *virtual quadric* is a quadric with no real point. All proper virtual conics (PVC) are central [2] and hence can be transformed to *Euclidean normal form* by a Euclidean transformation (principal axis transformation). The Euclidean normal form of a virtual conic is a diagonal matrix of the conic's eigenvalues, which all have the same sign.

2.5. Cones

By *cones* we mean rank-3 quadrics in $\mathscr{P}^3$ with vertex not on Π_∞. A cone is uniquely defined by its vertex and any (conic) section by a plane not containing the vertex. Cones are used in this paper through the notion of the *projection cone* of a 3D conic, i.e. the cone formed by the projection rays of the perspective projection of the conic. The Euclidean normal form of a cone is a diagonal matrix $\mathrm{diag}(\lambda_1, \lambda_2, \lambda_3, 0)$ with non zero λ_i. If the λ_i are all distinct then the cone is an *elliptic cone*. If exactly two of the λ_i are equal the cone is *circular* (or *right*). For an *isotropic cone*, all three λ_i are equal. Each isotropic cone contains the absolute conic (see description later).

A circular cone is invariant to arbitrary rotation about a single line passing through its vertex. This line is called the cone's *axis*. An isotropic cone is invariant to any rotation about its vertex.

2.6. Absolute quadric and absolute conic

The *absolute quadric* of $\mathscr{P}^n$ is defined by the equations $x_1^2 + \ldots + x_n^2 = x_{n+1} = 0$. The *absolute conic* Ω is the absolute quadric of P^3. Ω is a proper virtual conic in the ideal plane whose position uniquely defines the Euclidean structure of 3-space. The calibration of a camera is equivalent to determining the image ω of Ω, respectively, its dual ω^* [7,11]. From the relation $\omega^* \sim KK^{\mathrm{T}}$, the calibration matrix K can uniquely be recovered by Cholesky decomposition [15].

3. Problem formulation

We consider a sequence of n views, generally taken from different positions and with different orientations. The focal lengths for the views may all be different and the other intrinsic parameters (aspect ratio and principal point) are known (they need not be equal for all the views). The problem at hand is to perform focal length self-calibration, i.e. to determine the n different values for the focal length, which allows in general to obtain a Euclidean reconstruction of the scene. In the following, we describe this problem in geometrical terms, in analogy to Ref. [18].

First, *calibration* of a camera is equivalent to the determination of the image of the absolute conic, as 'produced' by that camera. *Self-calibration* means the same but with the conotation that information used to calibrate does not stem from, e.g. known metric 3D structure. *Euclidean reconstruction* is equivalent to the determination of the position of the absolute conic in 3D. The problem of Euclidean reconstruction is slightly more general than that of self-calibration: degeneracy of self-calibration implies degeneracy of Euclidean reconstruction while the reciproque is not always true (e.g. self-calibration of a camera rotating about its optical center is in general possible while any level of 3D reconstruction is impossible, including Euclidean reconstruction). The derivations that follow refer to degeneracies of Euclidean reconstruction.

To determine the position of the absolute conic, some constraints are needed. We will describe these constraints in the following paragraph but first we give a straightforward informal definition for degeneracy of Euclidean reconstruction: the Euclidean reconstruction problem is degenerate exactly if there is a conic in 3D not identical with the absolute conic that satisfies the mentioned constraints [17,18]. All such conics will be called *potential absolute conics*.

We now describe the constraints that may be used to determine the absolute conic. First, the absolute conic must be a proper virtual conic. Second, the image of the absolute conic by any perfect perspective projection is also a proper virtual conic. Third, the knowledge of some intrinsic parameters constrains the *projections* of the absolute conic in a given set of views, which in turn gives us constraints on the absolute conic itself. For the scenario considered here we make these constraints explicit in the following.

The image of the absolute conic, as 'produced' by a camera with calibration matrix K is given by:

$$\omega \sim K^{-\mathrm{T}}K^{-1} \sim \begin{pmatrix} 1 & 0 & -u_0 \\ 0 & \tau^2 & -\tau^2 v_0 \\ -u_0 & -\tau^2 v_0 & u_0^2 + \tau^2 v_0^2 + \tau^2 f^2 \end{pmatrix} \quad (1)$$

Since f may vary and we know the other intrinsic parameters, there is, for each view, exactly one family of possible images of the absolute conic. Consider now a conic Φ in 3D and its projection Φ in one view. For ϕ being a potential absolute conic its projection ϕ must be of the form Eq. (1) for some non zero real value a possibly different from the true f (remember that we suppose that the other intrinsic parameters are known):

$$\phi \sim \begin{pmatrix} 1 & 0 & -u_0 \\ 0 & \tau^2 & -\tau^2 v_0 \\ -u_0 & -\tau^2 v_0 & u_0^2 + \tau^2 v_0^2 + \tau^2 a^2 \end{pmatrix}$$

It is easy to show that a conic has this form exactly if, in the metric image plane, the conic is a virtual circle, centered in the origin. To see this, we map ϕ from pixel coordinates to the metric image plane using the true calibration matrix:

$$\phi_m \sim K^{\mathrm{T}}\phi K \sim \begin{pmatrix} f^2 & 0 & 0 \\ 0 & f^2 & 0 \\ 0 & 0 & a^2 \end{pmatrix}$$

For any non zero real value of a, this represents a proper virtual circle whose center is the coordinate origin. It is important to note that this statement is independent of the actual true value f of the focal length. Since all other intrinsic parameters are known the only important parameters for the consideration of degeneracy of Euclidean reconstruction are the extrinsic parameters of the views in a given sequence.

We now summarize the discussion.

Proposition 1. *Consider a sequence of n views with known aspect ratio and principal point but unknown and possibly different values for the focal length. Let P_{ci} be the canonical projection for view $i, i = 1, \ldots, n$.*

Euclidean reconstruction is degenerate exactly if there is at least one 3D conic Φ not identical with the absolute conic such that:

- *Φ is a proper virtual conic;*
- *the ϕ_i, $i = 1, \ldots, n$, where ϕ_i is the projection of Φ by P_{ci} are proper virtual circles centered in the origin. This is equivalent to the ϕ_i being represented by diagonal matrices whose diagonal elements are all non zero real values of the same sign the first two elements being equal.*

Definition 1. Any Φ as defined in Proposition 1 is called a *potential absolute conic*.

Consider a sequence of n views. Let (R_i, t_i) be the extrinsic parameters of view $i, i = 1, \ldots, n$. If Euclidean reconstruction is degenerate for the sequence of views, we say that $\{(R_i, t_i) | i = 1, \ldots, n\}$ is a *critical motion sequence* for Euclidean reconstruction.

The aim of the following section is to derive all generic critical motion sequences, i.e. all configurations where there is no unique solution to Euclidean reconstruction.

4. Derivation of the critical motion sequences

In this section the critical motion sequences are derived based on Proposition 1 and Definition 1. We proceed in a constructive manner: given a generic proper virtual 3D conic Φ, we determine all possible extrinsic parameters that form a critical motion sequence with respect to Φ, i.e. for which Φ is a potential absolute conic. The derivations are divided into two parts considering potential absolute conics Φ which lie/do not lie on Π_∞. The results are summarized in Section 5.

4.1. Potential absolute conics on Π_∞

Let Φ be a PVC on the ideal plane. Its canonical projection ϕ by a camera with extrinsic parameters (R, t) is given by:

$$\phi \sim R\Phi R^{\mathrm{T}} \quad (2)$$

Like it is the case for all geometric entities on the ideal plane the projection depends only on the orientation of the camera not on its position.

We now determine all orientations R for which Φ is a proper virtual circle centered in the origin. This implies that ϕ is a diagonal matrix of the form: $\phi \sim \mathrm{diag}(b, b, 1)$. If we choose the free scale factor for Φ such that $\det \Phi = b^2 = \det \phi$, the $\sim$ in Eq. (2) can be replaced by an equality sign:

$$\begin{pmatrix} b & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & 1 \end{pmatrix} = R\Phi R^{\mathrm{T}} \tag{3}$$

Eq. (3) implies that Φ has a double eigenvalue b and the single eigenvalue 1. The case $b = 1$ is of no interest here because this would mean that Φ is the *true* absolute conic (the absolute conic is the only conic on the ideal plane with a triple eigenvalue).

From Eq. (3), we derive:

$$R^{\mathrm{T}} \begin{pmatrix} b & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \Phi R^{\mathrm{T}} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

and further:

$$R^{\mathrm{T}} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \Phi R^{\mathrm{T}} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Hence the vector

$$v_R = R^{\mathrm{T}} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{4}$$

is an eigenvector of Φ to the eigenvalue 1. Since 1 is a single eigenvalue, all its associated eigenvectors are equal up to scale. This means that for all rotation matrices R in a critical motion sequence the vectors v_R must be equal up to scale.

It is easy to show that $(v_R^{\mathrm{T}}, 0)^{\mathrm{T}}$ is nothing else than the ideal point of the optical axis for a camera with orientation R. All v_R being equal up to scale is thus equivalent to the optical axes of all views being parallel. This is thus a necessary condition for critical motion sequences with respect to a conic on the ideal plane.

We now show that this is also a sufficient condition. Remember that eigenvectors of symmetric matrices that are associated to different eigenvalues are mutually orthogonal [2]. Thus, the eigenspace of Φ for the double eigenvalue b consists of all vectors orthogonal to v_R. Let r_i^{T} be the row vector representing the ith row of the rotation matrix R. From Eq (4) we have $r_3 = v_R$. Since R is an orthogonal matrix, we have $r_1 \perp v_R$ and $r_2 \perp v_R$, which means that r_1 and r_2 are eigenvectors of Φ associated to the eigenvalue b. We thus obtain:

$$R\Phi R^{\mathrm{T}} = \begin{pmatrix} r_1^{\mathrm{T}} \\ r_2^{\mathrm{T}} \\ r_3^{\mathrm{T}} \end{pmatrix} \Phi \begin{pmatrix} r_1 & r_2 & r_3 \end{pmatrix} = \begin{pmatrix} r_1^{\mathrm{T}} \\ r_2^{\mathrm{T}} \\ r_3^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} br_1 & br_2 & r_3 \end{pmatrix} = \begin{pmatrix} b & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

In conclusion, the projection of Φ is a virtual centered circle exactly if the ideal point of the camera's optical axis is $(v^{\mathrm{T}}, 0)^{\mathrm{T}}$ where v is the eigenvector of Φ associated to its single eigenvalue. Hence, a motion sequence is critical with respect to a PVC on the ideal plane if and only if the optical axes of all the views are parallel.

4.2. Potential absolute conics not on Π_∞

Contrary to conics on the ideal plane the projection of conics not on Π_∞ depends on both camera position and orientation. First, we deal with position then with orientation.

4.2.1. Position

Let Φ be a PVC *not* on the ideal plane. Consider a view with optical center at position t. Let ϕ be the canonical projection of Φ. Let Λ be the projection cone of Φ (cf. Section 2.5).

One condition for Φ being a potential absolute conic is that ϕ is a circle, centered in the origin of the metric image plane. Note that the origin of the metric image plane is the camera's principal point, i.e. the intersection of the optical axis with the image plane. Since the optical axis is perpendicular to the image plane the projection cone Λ must be circular and its axis is the camera's optical axis. The vertex of the projection cone being the optical center t, we obtain constraints on the possible camera positions in a critical motion sequence: for a potential absolute conic Φ all possible camera positions are the vertices of circular cones that contain Φ. These are summarized in the following (proofs are given in Appendix A).

If Φ is a virtual circle then the locus of possible camera positions in a critical motion sequence is the line L perpendicular to the circle's supporting plane and passing through the circle's center (Fig. 1(a)). If Φ is a virtual ellipse then the locus of camera positions is the union of a real ellipse Ψ_{e} and a real hyperbola Ψ_{h} (Fig. 1(b)). The supporting planes of the three conics Φ, Ψ_{e} and Ψ_{h} are mutually perpendicular (there are further relations between these conics, cf. Appendix A).

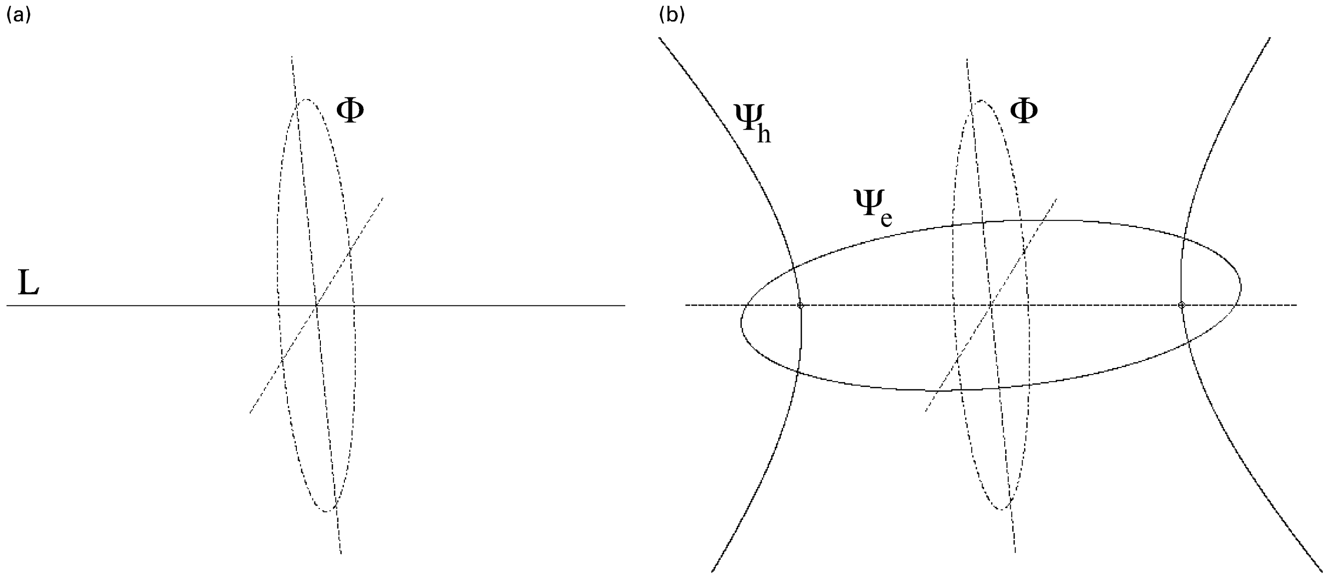

Fig. 1. Locus of camera positions in a motion sequence critical with respect to a conic Φ not on the ideal plane. The conic Φ is shown in dotted style to illustrate that it is virtual and can in fact not be drawn. The figures are further explained in the text. (a) The case of the PVC Φ being a circle. (b) The case of the PVC Φ being an ellipse.

4.2.2. Orientation

The results of the previous paragraph are necessary conditions that hold for camera *positions* in a critical motion sequence. To obtain sufficient conditions, we now consider the *orientation* of cameras. First note that rotations about the optical axis are not important here: if the projection of Φ is a circle centered in the origin (the principal point) then any rotation about the optical axis will preserve this property. Hence, the only part of camera orientation that matters is the *direction of the optical axis*.

For the camera positions derived in the previous paragraph, we have to determine the directions of the optical axis for which Φ is projected onto a circle centered in the origin. The proofs of the following statements are given in Appendix A.

If Φ is a virtual circle (cf. Fig. 1(a)) then the optical axes of all views have to coincide with L for the camera configuration to be critical. The only exception is that at two camera positions on L the optical axis might be oriented arbitrarily. These positions are symmetric with respect to the supporting plane of Φ; their distance d from that plane is related to the radius r of Φ by $d = \mathrm{Im}(r)$ (the imaginary part of the radius, which is complex due to Φ being a virtual circle). For more details, see Section B.1.

To summarize, critical motion sequences with respect to a *circle* Φ consist of collinear optical centers and optical axes passing through all optical centers except that at two positions the optical axis may be oriented arbitrarily (several views might be taken from these two positions by a camera rotating about its optical center).

If Φ is a virtual ellipse (cf. Fig. 1(b)) and the optical center lies on the ellipse Ψ_e (respectively the hyperbola Ψ_h) then the optical axis has to be the tangent of Ψ_e (respectively Ψ_h) at the optical center in order for the camera configuration to be critical.

5. Summary of critical motion sequences for a moving camera

The following camera positions/orientations constitute critical motion sequences for Euclidean reconstruction:

Case 1: Arbitrary position of optical centers but parallel optical axes. This means that camera motions are pure translations possibly combined with an arbitrary rotation about the optical axis and a reversal of the gaze direction.

Case 2: Collinear optical centers. The optical axes at two positions may be oriented arbitrarily, all others coincide with the line joining the optical centers. This means that camera motions are pure forward translations with two exceptions where the translation may be followed by an arbitrary rotation about the optical center.

Case 3: The optical centers lie on an ellipse/hyperbola pair as shown in Fig. 1(b). At each position, the optical axis is tangent to the ellipse/hyperbola. A necessary condition derived from this is: the views may be partitioned into at most two sets for which the centers and optical axes are all coplanar. In addition, these two sets define planes which are perpendicular to each other.

If we consider only one of the two conics Ψ_e or Ψ_h we can describe the critical motion sequences as follows: a camera moving on a trajectory that may be described as (arc of) a conic and always gazing in the direction of motion, i.e. the current tangent direction. An example might be a camera mounted on a vehicle gazing in driving direction.

Newsam et al. derived degenerate configurations for two-view self-calibration [12]. Their results are of course contained in the earlier list.

Our results were reported in Ref. [19]. Kahl considered the problem in Ref. [10] but his results are not complete. As for case 2, he only obtains two subcases:

- two optical centers and arbitrary orientation;
- collinear optical centers and all optical axes aligned with the optical centers.

Whereas the 'union' of these cases (case 2 described earlier) also describes critical motion sequences.

5.1. Degree of ambiguity

We now discuss the resulting degree of ambiguity in the Euclidean reconstruction or the ego-motion estimation of the camera for the earlier cases of critical motion sequences.

5.1.1. Case 1

We only consider the case of unaligned optical centers. Aligned optical centers are discussed in Section 5.1.2.

All potential absolute conics lie on the ideal plane. Let $(Q^T, 0)^T$ be the ideal point of the optical axes in the critical motion sequence. It can be shown that the potential absolute conics form exactly the following 1-degree-of-freedom family:

$$\Phi(\lambda) \sim I + \lambda QQ^T$$

i.e. the family of conics spanned by the absolute conic (represented by the identity matrix I) and the degenerate conic QQ^T.

Since all potential absolute conics lie on the ideal plane the ideal plane can be recovered uniquely, which means that affine reconstruction is possible [18,10]. This implies, e.g. that relative camera displacements in the gazing direction can be estimated. However, the 1-dof-ambiguity for Euclidean reconstruction does not allow measuring angles correctly. For example, the direction of translation between different viewing positions (with respect to, e.g. the gazing direction) cannot be determined or analogously the direction of a detected obstacle's location.

5.1.2. Case 2

Different subcases have to be discussed. First, if all the optical axes are aligned (i.e. are identical with the line L shown in Fig. 1(a)) then the motion sequence is also critical according to case 1. There is a 2-dof-family of potential absolute conics: the conics described in Section 5.1.1 and all circles whose centers lie on L and whose supporting planes are orthogonal to L. Compared to case 1, affine reconstruction is not possible here. This would usually cause a wrong estimation of relative displacements, e.g. the time of impact with respect to an obstacle might be wrongly estimated.

If all the optical axes are aligned with one exception (i.e. at one viewing position t along the line L the camera gazes at other directions than along L) then there remain a 1-dof-family of potential absolute conics. These are the (true) absolute conic and one virtual circle per plane that is orthogonal to L (the circles' centers lie on L). These conics are the intersections of the planes orthogonal to L and the isotropic cone with vertex t. Affine reconstruction is not possible and as before relative displacements and angles cannot be estimated correctly.

The third subcase occurs when the camera gazes in other directions than along L at exactly two viewing positions. Only two potential absolute conics remain, which means that there are only two different solutions for Euclidean reconstruction. The potential absolute conics are the two intersections of the isotropic cones with vertices at the two exceptional viewing positions (one intersection is the true absolute conic of course, the second one lies on the equidistance plane of the two viewing positions). The wrong solution can often be ruled out in practice by imposing that the reconstructed scene lies in front of all the views.

5.1.3. Case 3

This case is difficult to explain when the motion sequence comprises three or fewer viewing positions (please contact the author for results). For more than three viewing positions, however, it is easy to see that only two potential absolute conics are possible: the true absolute conic and the virtual ellipse shown in Fig. 1(b). Thus, although case 3 (motion on conic arcs) seems to be relevant in practice the existence of only two solutions for Euclidean reconstruction is reassuring. As noted by Pollefeys the wrong solution can often be ruled out because it would lead to unrealistic estimates of the focal length [14].

6. Comments

In this section, we discuss a few special cases.

6.1. Two cameras in general position

From case 2 in the Section 5.1.2, it follows that Euclidean reconstruction is always degenerate from only two views independently of their position and orientation. In fact, it is known that in this case there is a two fold ambiguity for the absolute conic: let ω_1 and ω_2 be the projections of the absolute conic in the two views. The projection cones of ω_1 and ω_2 intersect of course in the absolute conic but in general also in a second conic Φ hence the ambiguity. However, self-calibration can in general be achieved since

Φ has the same projections as the absolute conic (it lies on the same projection cones). This illustrates that Euclidean reconstruction and self-calibration are not exactly equivalent problems. As mentioned before the ambiguous solution for Euclidean reconstruction can often be ruled out in practice by imposing that the reconstructed scene lies in front of both cameras.

6.2. Camera rotating about its optical center

A camera rotating about its optical center while possibly changing its focal length can always be calibrated from two views whose optical axes do not coincide. This is briefly explained in the following. The two views are critical for Euclidean reconstruction according to case 2 (cases 1 and 3 allow only one optical axis per camera position in a critical motion sequence). The position t of the two views is one of the two exceptional points. The cone defined by Φ and t contains the absolute conic (it is an isotropic cone, cf. Section B.1). Hence, the projections of Φ are identical with the image of the absolute conic, which means that (self-) calibration has a unique solution.

6.3. Fixation

Consider cameras fixating a finite point (i.e. the point lies on the optical axes of the cameras). For two cameras, the configuration is always critical whereas for more than two cameras (with different optical centers) it is always non critical.

The first statement is easy to understand. There is a one-dimensional family of possibilities to realize case 3, i.e. a one-dimensional family of ‘motion conics’ Ψ: the conic has to contain the optical centers and has the two optical axes as tangents. This gives four constraints on the five degrees of freedom for a conic which means that there remains one degree of freedom for Ψ and thus for the potential absolute conic Φ.

Consider now an additional camera that fixates the same finite point as the two others. Suppose the configuration were still critical. From Section A.2.1 we know that optical axes for optical centers on Ψ_e and Ψ_h are mutually skew hence it follows that the three optical centers have to lie on either Ψ_e or Ψ_h and the optical axes are tangents to the same conic. This would imply that the fixated point lies on three different tangents to the same conic, which is not possible [16]. Hence, the assumption that the configuration is critical, is contradicted.

7. Derivation of the critical motion sequences for stereo systems

The results presented in the previous sections allow us to study degeneracies of Euclidean reconstruction for stereo systems. Here, a stereo system consists of two cameras with coplanar optical axes and symmetric but possibly variable vergence angles α (cf. Fig. 2). The distance between the two cameras is fixed. The focal lengths of the two cameras are not constrained to be equal and they may vary freely between different images. We define the *axis of a stereo system* as the line perpendicular to the baseline and passing through its midpoint.

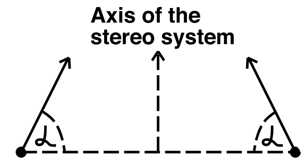

Fig. 2. The type of stereo system discussed in this paper.

A single pair of images taken by such a stereo system is of course critical (cf. Section 6.3). In the following, we reveal the conditions for two stereo pairs to be critical. They are critical if the set of individual views constitute a critical motion sequence as described in the previous sections. We consider several cases:

- coplanar stereo pairs with identical vergence angles;
- coplanar stereo pairs with variable vergence angles;
- non coplanar stereo pairs.

By coplanar stereo pairs we mean that all the optical centers and optical axes are located in the same plane.

Before examining the different cases, we give an introductory remark concerning case 3 of Section 5.

7.1. Concerning case 3

For a given stereo system, we want to establish constraints on the possible locations of ‘motion conics’ Ψ. Since the optical axes of the two cameras in the stereo system are coplanar the two optical centers have to lie on either the ellipse Ψ_e or the hyperbola Ψ_h but cannot be distributed on both (cf. Section A.2.1). It is now easy to show that due to symmetry of the vergence angles, Ψ must be symmetric with respect to the axis of the stereo system. Since Ψ cannot be a circle (cf. Section A.2.1) it has exactly two symmetry lines (which are perpendicular to each other). This means that the axis of the stereo system coincides with one of these two symmetry lines of Ψ.

7.2. Coplanar stereo pairs with identical vergence angles

For the trivial case of a stationary stereo system any number of stereo pairs are critical. Non-trivial cases are discussed in the following.

7.2.1. Parallel optical axes

If the vergence angles are of 90°, i.e. if the two optical axes are parallel then each stereo pair is critical according to cases 1 and 3. According to case 1 the combination of two stereo pairs is only critical if the stereo system undergoes a pure translation or a rotation by 180° in the plane of motion,

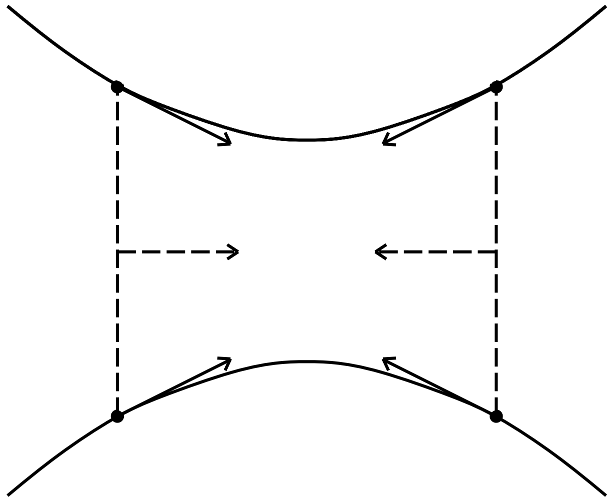

Fig. 3. Two stereo systems with opposite gazing direction but identical axes constitute a critical motion sequence according to case 3. The hyperbola shown is the 'motion conic' Ψ.

i.e. a reversal of gaze direction. Case 3 is dealt with as in the following paragraph.

7.2.2. Convergent (non parallel) optical axes

The stereo pairs can only be critical according to case 3 and this only if there is an ellipse or a hyperbola Ψ containing all four optical centers and having all optical axes as tangents. From Section 7.1, we conclude that the axes of the two stereo pairs do either coincide or are perpendicular to each other. It can be shown that in case they are perpendicular, there is no possibility to place the two stereo pairs in a way that a conic Ψ as described earlier exists (since the distance between the two cameras and the vergence angles are fixed). In case the axes coincide a conic Ψ exists exactly if the two stereo pairs have opposite gaze direction as shown in Fig. 3.

7.2.3. Summary

Two coplanar stereo pairs with identical vergence angles are critical in exactly the following situations. If the optical axes of the stereo system are parallel then the stereo pairs are critical if they are related by a pure translation possibly followed by a reversal of gaze direction. If the optical axes are convergent then the stereo pairs are only critical if they gaze in opposite directions and if their axes are identical. The only case of practical importance is pure translation of a stereo system with parallel optical axes.

7.3. Coplanar stereo pairs with variable vergence angles

Due to varying vergence angles at least one stereo pair has convergent optical axes. Hence, the combination of the stereo pairs can only be critical according to case 3. As stated in Section 7.1, the conic Ψ must be symmetric with respect to the axes of the stereo pairs, which implies that these are either identical or perpendicular to each other. It can be shown that if they are identical no conic Ψ as described earlier can exist. For perpendicular axes, there is a one-dimensional family of possibilities for placing the stereo pairs and setting their vergence angles relative to each other such that the combination of the stereo pairs is critical.

In summary, if the vergence angles for the two stereo

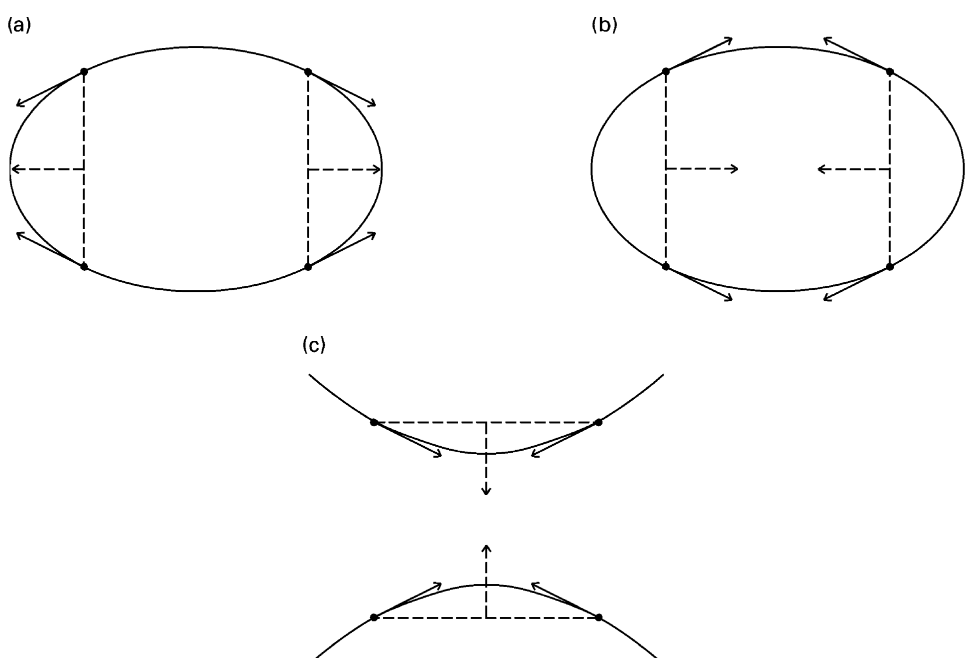

Fig. 4. Possible arrangements of stereo systems and 'motion conics' (described in the text). (a) Possible arrangement of stereo systems and motion ellipse Ψ_e, as described by case 3 in Section 5. For vergence angles $\alpha < 90°$ the stereo system gazes away from the ellipse's center. (b) Only for vergence angles $\alpha > 90°$ the stereo system gazes towards the ellipse's center. (c) Possible arrangement of stereo systems and motion hyperbola Ψ_h.

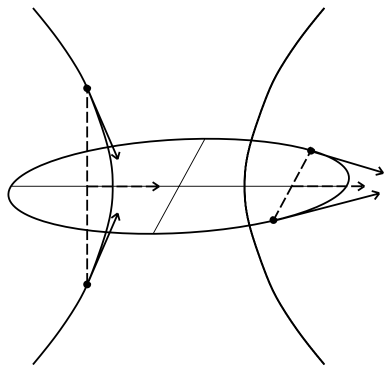

Fig. 5. One possible non coplanar critical motion sequence for stereo systems.

pairs are different the configuration can only be critical if the axes of the stereo pairs are perpendicular to each other. The exact conditions for being critical are complicated and omitted here since they do not really contribute to the understanding of this discussion.

7.4. Non coplanar stereo pairs

Non coplanar stereo pairs can only be critical according to cases 1 or 3. Case 1 is only possible if the optical axes are all parallel (vergence angle $\alpha = 90°$). Critical motions are exactly pure translations possibly followed by rotations about axes parallel to the optical axes or by reversals of the gazing direction.

As for case 3, the non coplanarity implies that one stereo pair is located on the ellipse Ψ_e and the other on the hyperbola Ψ_h. We assume that practical vergence angles are inferior (or equal to) 90°. This means that the axis of the stereo system located on the ellipse Ψ_e is directed away from the ellipse's center as shown in Fig. 4(a), i.e. a case as in Fig. 4(b) is not possible. As for the hyperbola the contrary is valid, i.e. the stereo system's axis will be directed towards the hyperbola's center (cf. Figs. 3 and 4(c)).

By comparing this discussion with Fig. 1(b), it is clear that for the only cases relevant in practice the axes of the two stereo pairs must be identical as shown in Fig. 5 (for the other potential arrangements, one stereo pair would fixate a point that is behind the other stereo pair, i.e. the common field of view would be either empty or very small making image matching and thus self-calibration inherently impossible). The relative position of the stereo pairs is thus as follows. Their axes coincide and their 'supporting planes' are orthogonal to each other.

In conclusion, the only critical situations that might be encountered with non coplanar stereo systems are:

- a stereo system with parallel optical axes undergoing pure translations (possibly in several different directions), possibly followed by rotations about axes parallel to the optical axes and reversals of the gazing direction.
- a stereo system with parallel or convergent optical axes that is rotated by 90° about its axis followed by a translation along its axis and possibly a reversal of gazing direction.

8. Conclusions

We have derived all motion sequences that are critical for Euclidean reconstruction from image sequences with variable and unknown focal length whose other intrinsic parameters are known. The critical motion sequences are described geometrically. Our results are rather encouraging since only few cases exist that are likely to be met in practice. The most important cases are pure translation and especially pure forward motion. One also has to be aware of the fact that motion on a conic arc while gazing in the direction of motion is critical although in general there are only two ambiguous solutions for Euclidean reconstruction and self-calibration. Another important result is that an image sequence taken by a camera that fixates a finite point is always critical when two views only are used but never critical with three or more views.

Both pure translation and motion on a conic are also critical for self-calibration when all intrinsics are constant but unknown [18]. In that case, however, the degree of ambiguity in the solution is higher. Moreover, general planar motions are always critical, which is not the case for the situation dealt with in this paper.

Our results allowed us to study the critical motions of stereo systems. We did this for various situations, stereo systems undergoing planar or non planar motion and having fixed or variable vergence angles. Our results show that it should be rather easy to avoid critical motions in practice: it suffices to guarantee that the axes of the different stereo pairs in a sequence are neither parallel nor perpendicular to each other.

Acknowledgements

This work was done while the author stayed with the Computational Vision Group of the University of Reading and was supported by the EPSRC funded project GR/K89221 (Vector).

Appendix A. Locus of vertices of circular cones containing a conic section

Consider a proper virtual conic Φ. We want to determine the locus of all real points C such that the cones formed by Φ and with C as vertex are circular. Without loss of generality, we can choose simple coordinates for the problem as follows: let the supporting plane of Φ be the plane $Z=0$ and let the conic be centered in the origin and with axes aligned with the X and Y axes. Hence, the conic's matrix is

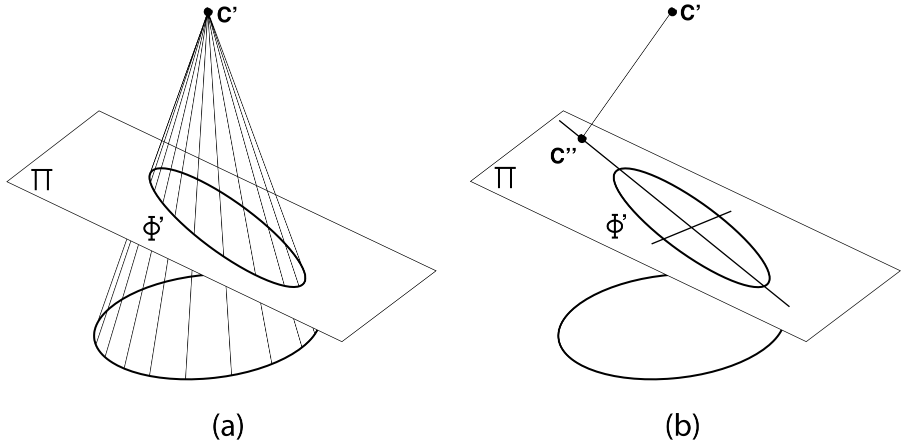

Fig. A1. For any circular cone and any plane, the orthogonal projection of the cone's vertex on that plane lies on one of the two symmetry lines of the conic section induced by the cone and the plane (see text for more details).

of the form:

$$\Phi \sim \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Since Φ is a proper virtual conic we have $a, b < 0$.

Let now $C = (X, Y, Z, 1)^{\mathrm{T}}$ be the vertex of a cone Λ that contains Φ (with $Z \neq 0$). The cone's matrix is given by:

$$\Lambda \sim \begin{pmatrix} a & 0 & -a\frac{X}{Z} & 0 \\ 0 & b & -b\frac{Y}{Z} & 0 \\ -a\frac{X}{Z} & -b\frac{Y}{Z} & \frac{aX^2 + bY^2 - 1}{Z^2} & \frac{1}{Z} \\ 0 & 0 & \frac{1}{Z} & -1 \end{pmatrix}$$

We want to establish C for which the cone Λ is circular. Cones with finite vertex are circular exactly if the conic obtained by intersection with the ideal plane has a double eigenvalue [2]. This condition is explored for the two different cases of Φ being a virtual circle or an ellipse (there are no other cases for PVC [2]).

A.1. Φ is a virtual circle

This case occurs when $a = b$. For Λ to be a circular cone, its vertex C must lie on the line passing through the center of Φ and being orthogonal to its supporting plane. Here, this is the Z-axis.

A.2. Φ is a virtual ellipse

This case occurs when $a \neq b$.

First, we show that C must lie in one of the symmetry planes of Φ, which are defined as follows. A symmetry plane of an ellipse is a plane orthogonal to the ellipse's supporting plane, which contains one of the ellipse's two symmetry lines. Besides the supporting plane itself the two symmetry planes are the only planes which conserve Φ by reflection.

Consider a circular virtual cone with vertex C'. Let Π be a plane with $C' \notin \Pi$ and let Φ' be the conic cut out from the cone by Φ (Fig. A1(a)). Let C'' be the orthogonal projection of the cone's vertex C' on Φ. It is easy to show that C'' lies on one of the two symmetry lines of Φ' (Fig. A1(b)). Since C'' is the orthogonal projection of C' on Π the plane spanned by C' and that symmetry line is a symmetry plane of Φ'. We conclude that for all conic sections of a circular cone the cones vertex lies on an associated symmetry plane.

In our case, the symmetry planes of Φ are the planes $X = 0$ and $Y = 0$. From the earlier discussion it follows that the vertex C of the circular cone Λ must lie in one of these planes. In the following, we examine the case $X = 0$ (the other case can be treated in an analogous way). The cones matrix is thus simplified to:

$$\Lambda \sim \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & b & -b\frac{Y}{Z} & 0 \\ 0 & -b\frac{Y}{Z} & \frac{bY^2 - 1}{Z^2} & \frac{1}{Z} \\ 0 & 0 & \frac{1}{Z} & -1 \end{pmatrix}$$

As stated earlier, the cone is circular exactly if the conic obtained by intersection with the ideal plane has a double eigenvalue. This conic Φ_∞ is given by the upper left 3×3

submatrix of Λ:

$$\Phi_\infty \sim \begin{pmatrix} a & 0 & 0 \\ 0 & b & -b\frac{Y}{Z} \\ 0 & -b\frac{Y}{Z} & \frac{bY^2-1}{Z^2} \end{pmatrix}$$

The three eigenvalues of Φ_∞ are:

$$a \text{ and } \frac{b(Y^2+Z^2)-1 \pm \sqrt{b^2(Y^2+Z^2)^2+2b(Z^2-Y^2)+1}}{2Z^2}$$

Equating the second and third eigenvalues leads to subcases of equating the first eigenvalue a with the second or third eigenvalue. Equating the first eigenvalue a with the second or third eigenvalue leads to the following constraint on Y and Z (after some manipulations using MAPLE):

$$abY^2 + (b-a)(aZ^2+1) = 0 \tag{A1}$$

Eq. (A1) can be represented by the following matrix equation:

$$(Y,Z,1)\begin{pmatrix} ab & 0 & 0 \\ 0 & a(b-a) & 0 \\ 0 & 0 & b-a \end{pmatrix}\begin{pmatrix} Y \\ Z \\ 1 \end{pmatrix}$$

The matrix

$$\Psi \sim \begin{pmatrix} ab & 0 & 0 \\ 0 & a(b-a) & 0 \\ 0 & 0 & b-a \end{pmatrix}$$

represents a conic in the plane $X = 0$ whose type is:

- a real ellipse, if $b < a$. Note that Ψ cannot be a circle (this would be the case if $\Psi_{11} = \Psi_{22}$ which is equivalent to $a = b$ in contradiction to the assumption that $a \neq b$).
- a hyperbola, if $a < b$.

If we consider the case $Y = 0$, we obtain just the reciproque result.

Hence the locus of vertices of circular cones containing a virtual ellipse Φ is the union of a real ellipse Ψ_e and a real hyperbola Ψ_h (cf. Fig. 1(b)). The following relations hold between Φ, Ψ_e and Ψ_h. Their supporting planes are mutually orthogonal and each of them is a symmetry plane for the two other conics. The hyperbola Ψ_h passes through the focii of Ψ_e.

A.2.1. Further observations

We note that Ψ cannot be a circle (property used in Section 7). It would be a circle if (cf. Eq. (A2)) $\Psi_{11} = \Psi_{22}$, i.e. if $ab = a(b-a)$, hence if $a = 0$. This is in contradiction with the fact that a is an eigenvalue of the *proper* virtual conic Φ and thus non zero.

Another property that is used in Section 6 is easy to show. Namely, all real tangents of Ψ_e and of Ψ_h are mutually skew, i.e. there is no tangent of Ψ_e that has a real intersection point with any tangent of Ψ_h.

Appendix B. Axis of circular cones containing a conic section

In the following we prove the statements made in Section 4.2.2. Remember that the optical axis is the axis of the circular cone Λ (cf. Section 4.2.1). As in Appendix A we consider the two cases of Φ being a virtual circle or a virtual ellipse. We use the same simple coordinates as in Appendix A.

B.1. Φ is a virtual circle

We already know that the vertex of a circular cone Λ containing Φ is a point with coordinates $(0,0,Z,1)^T$, i.e. Λ is given by:

$$\Lambda \sim \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & -\frac{1}{Z^2} & \frac{1}{Z} \\ 0 & 0 & \frac{1}{Z} & -1 \end{pmatrix}$$

The intersection of Λ with the ideal plane is a conic given by:

$$\Phi_\infty \sim \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & -\frac{1}{Z^2} \end{pmatrix}$$

For $Z^2 \neq -1/a$, Φ_∞ has a double and a single eigenvalue, i.e. Λ is a circular cone with the Z-axis as axis. However, for $Z = \pm 1/\sqrt{-a}$ (these are real numbers since $a < 0$), Φ_∞ is the absolute conic (given by the identity matrix) which means that Λ is an isotropic cone. Isotropic cones are invariant to rotation about their vertex. Hence, the projection of an isotropic cone by a camera located at its vertex is always a centered virtual circle, regardless of the camera's orientation.

B.2. Φ is a virtual ellipse

It is easy to show that in this case, Λ cannot be an isotropic cone, i.e. at each possible camera position, there is only one possible direction for the optical axis (given by the cone's axis), such that the camera configuration is critical. In the following we prove that the axis of Λ is the tangent line to Ψ at C that lies in the supporting plane of Ψ (notation as in Appendix A).

Another definition of the axis of a circular cone as that given in Section 4.2.1, is as follows: exactly planes orthogonal to the axis cut the cone in circles, i.e. conics containing circular points.

The tangent line to Ψ at C is given by (coordinates in the plane $X = 0$):

$$T \sim \Psi \begin{pmatrix} Y \\ Z \\ 1 \end{pmatrix} \sim \begin{pmatrix} abY \\ a(b-a)Z \\ b-a \end{pmatrix}$$

The planes orthogonal to the line T are given by:

$$\Pi \sim \begin{pmatrix} 0 \\ a(a-b)Z \\ abY \\ d \end{pmatrix}$$

for real d. The two ideal intersection points of Π and Λ are:

$$Q \sim \begin{pmatrix} \pm\sqrt{a(a^2 + b^2 - 2ab - a^2bY^2)} \\ abY \\ a(b-a)Z \\ 0 \end{pmatrix}$$

They are circular points since:

$$Q_1^2 + Q_2^2 + Q_3^2 = a(b-a)(abY^2 + (b-a)(aZ^2 + 1))$$

which is, cf. Eq. (A1), equal to 0.

References

[1] A. Azarbayejani, A.P. Pentland, Recursive estimation of motion, structure and focal length, IEEE Transactions on Pattern Analysis and Machine Intelligence 17 (1995) 562–575.

[2] W. Boehm, H. Prautzsch, Geometric concepts for geometric design (A.K. Peters, 1994).

[3] S. Bougnoux, From projective to Euclidean space under any practical situation, a criticism of self-calibration, Proceedings of the International Conference on Computer Vision, Bombay, India, 1998, pp. 790–796.

[4] M.J. Brooks, L. de Agapito, D.Q. Huynh, L. Baumela, Direct methods for self-calibration of a moving stereo head. Proceedings of the European Conference on Computer Vision, Cambridge, 1996, pp. 415–426.

[5] L. de Agapito, D.Q. Huynh, M.J. Brooks, Self-calibrating a stereo head: an error analysis in the neighbourhood of degenerate configurations, Proceedings of the International Conference on Computer Vision, Bombay, India, 1998, pp. 747–753.

[6] O. Faugeras, What can be seen in three dimensions with an uncalibrated stereo rig? Proceedings of the European Conference on Computer Vision, Santa Margherita Ligure, Italy, 1992, pp. 563–578.

[7] O. Faugeras, Stratification of three-dimensional vision: projective, affine and metric representations, Journal of the Optical Society of America A 12 (1995) 465–484.

[8] R.I. Hartley, Estimation of relative camera positions for uncalibrated cameras, Proceedings of the European Conference on Computer Vision, Santa Margherita Ligure, Italy, 1992, pp. 579–587.

[9] R. Hartley, R. Gupta, T. Chang, Stereo from uncalibrated cameras, Proceedings of the European Conference on Computer Vision, Santa Margherita Ligure, Italy, 1992, pp. 761–764.

[10] F. Kahl, Critical motions and ambiguous Euclidean reconstructions in auto-calibration, Proceedings of the International Conference on Computer Vision, Kerkyra, Greece, 1999, pp. 469–475.

[11] S.J. Maybank, O.D. Faugeras, A theory of self calibration of a moving camera, International Journal on Computer Vision 8 (1992) 123–151.

[12] G.N. Newsam, D.Q. Huynh, M.J. Brooks, H.P. Pan, Recovering unknown focal lengths in self-calibration: an essentially linear algorithm and degenerate configurations. Part B3 of the proceedings of the XVIII ISPRS-Congress, Vienna, Austria, 1996, pp. 575–580.

[13] M. Pollefeys, R. Koch, L. Van Gool, Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters, Proceedings of the International Conference on Computer Vision, Bombay, India, 1998, pp. 90–95.

[14] M. Pollefeys, Self-calibration and metric 3D reconstruction from uncalibrated image sequences, PhD Thesis, Department ESAT, Katholieke Universiteit Leuven, Belgium, 1999.

[15] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical recipes in C, Cambridge University Press, Cambridge, New York, 1992.

[16] J.G. Semple, G.T. Kneebone, Algebraic Projective Geometry, Oxford Science Publication, 1952.

[17] P. Sturm, Critical motion sequences for monocular self-calibration and uncalibrated Euclidean reconstruction, Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, Puerto Rico, 1997, pp. 1100–1105.

[18] P. Sturm, Vision 3D non calibrée: contributions à la reconstruction projective et étude des mouvements critiques pour l'auto-calibrage, PhD Thesis, INPG, France, 1997.

[19] P. Sturm, Critical motion sequences for the self-calibration of cameras and stereo systems with variable focal length. Proceedings of the British Machine Vision Conference, Nottingham, 1999, pp. 63–72.