Stratification of 3-D vision: projective, affine, and metric representations

Olivier Faugeras
INRIA
2004 route des Lucioles, B.P. 93
06902 Sophia-Antipolis, France
faugeras@sophia.inria.fr

Abstract

In this article we provide a conceptual framework in which to think of the relationships between the three-dimensional structure of the physical space and the geometric properties of a set of cameras which provide pictures from which measurements can be made. We usually think of the physical space as being embedded in a three-dimensional euclidean space where measurements of lengths and angles do make sense. It turns out that for artificial systems, such as robots, this is not a mandatory viewpoint and that it is sometimes sufficient to think of the physical space as being embedded in an affine or even projective space. The question then arises of how to relate these models to image measurements and to geometric properties of sets of cameras. We show that in the case of two cameras, a stereo rig, the projective structure of the world can be recovered as soon as the epipolar geometry of the stereo rig is known and that this geometry is summarized by a single 3×3 matrix, which we called the fundamental matrix [1, 2]. The affine structure can then be recovered if we add to this information a projective transformation between the two images which is induced by the plane at infinity. Finally, the euclidean structure (up to a similitude) can be recovered if we add to these two elements the knowledge of two conics (one for each camera) which are the images of the absolute conic, a circle of radius $\sqrt{-1}$ in the plane at infinity. In all three cases we show how the three-dimensional information can be recovered directly from the images without explicitely reconstructing the scene structure. This defines a natural hierarchy of geometric structures, a set of three strata, that we overlay on the physical world and which we show to be recoverable by simple procedures relying on two items, the physical space itself together with possibly, but not necessarily, some a priori information about it, and some voluntary motions of the set of cameras.

1 Introduction

This article discusses several ways of representing the geometry of three-dimensional space, which we will call the world, when viewed by a system of cameras. We usually think of the world as being euclidean, i.e. of being a place where it makes sense to measure angles and distances. When we look at this space with a system of cameras, we first make measurements in the images and then, attempt to relate them to three-dimensional quantities. This has been, and still is, one of the main research topics in computer vision in such areas as motion analysis, stereo, and camera calibration. All computer vision scientists know that going from image quantities to reliable three-dimensional metric quantities is very difficult, basically because a camera is not a metric device, unless it has been carefully calibrated, which is itself a very difficult task. We will see later in the paper that a camera is really a projective device, hence, part of the difficulty.

There are two very important ideas that have emerged in the recent years and which are related to this problematic. The first idea is that it is not always necessary, in order to perform tasks in the world, to use metric measurements and that less detailed measures such as, for example, ratios of lengths, may quite often be sufficient to achieve these tasks. This is also an active research area in robotics and vision. The second idea is that calibration in the usual sense i.e. by using special calibration grids can be entirely avoided by using active camera motions and exploiting the fact that the world can be modelled as euclidean.

In this paper I want to articulate these two ideas with the idea that we can define on the three-dimensional space that we call the world not only the structure of an euclidean metric space that we are used to, but also weaker (hence more general) structures, e.g. affine and projective. These structures can be thought of geometric strata which are overlaid one after each other upon the world. First, the projective stratum which can be specialized next into an affine stratum which can itself be specialized further into an euclidean stratum. If we are used to thinking about projective, affine, and euclidean spaces, it is rarely the case that we think of these three structures *simultaneously*. But I think that in order to really understand the relationship between the world and its images we must be able to picture in our minds those three structures overlaid upon each other.

Closely related to this stratification idea is the idea of group of geometric transformations acting on the elements of these strata and leaving invariant some properties of geometric configurations of these elements. Attached to the projective stratum is the group of projective transformations or collineations, attached to the affine stratum is the group of affine transformations, and attached to the metric stratum is the group of rigid transformations or displacements. In fact, because we rarely have an absolute yardstick for measuring distances, we are in practice interested in a subgroup of the group of displacements, the group of similitudes, otherwise known as euclidean transformations. Interestingly enough, this group appears very naturally when we build the series of strata.

It is well-known but remarkable enough to be stressed here that these four groups can be considered as subgroups of each other, e.g. the group of affine transformations can be considered as a subgroup of the group of collineations and the group of similitudes as a subgroup of the group of affine transformations. These relationships will be made clearer in the paper.

This notion of groups brings in naturally another notion which is central to this paper, i.e. the notion of invariant. In our context, an invariant is a property of a geometric configuration which does not change when a transformation of a given group is applied to that geometric configuration. For a given geometric configuration, e.g. a set of points, lines, surfaces, there may exist projective invariants which are properties of the configuration which do not change when we apply a projective transformation to the elements of the configuration. These invariants are also be affine and similitude invariants of the same configuration. From the practical standpoint, this means that if can measure invariant properties of the world within the projective stratum (the most general), then these properties will remain invariant in the next strata, i.e. affine and euclidean.

We do not present in this paper any experimental results. This is not to say that we are not interested in the actual implementation of the ideas that I will present. In fact, many of these ideas or consequences of them have been implemented, sometimes on special purpose hardware, and we refer the interested reader to the corresponding publications [2, 1, 3, 4, 5, 6]. The purpose of this paper is to provide a coherent framework in which to express these ideas in a somewhat systematic and formal, perhaps even elegant, way. We hope that the reader will accept to step back a little and take a fresh look at a number of old problems. This may provide opportunities for solving more efficiently some newer problems.

2 Related work

Koenderink and van Doorn have started the interest of the computer vision community in non metric reconstructions from sets of cameras with their pioneering work on affine structure from motion [7].

Gunnar Sparr has developed a theory based on a novel definition of *shape* [8, 9] and showed that this theory could also be used to compute affine and projective reconstructions of the world from point correspondences. Contrarily to us, he does not rely on epipolar geometry to obtain these reconstructions. The cost he has to pay is a somewhat more complicated theory than ours and a more difficult combinatorial problem of obtaining the point correspondences since he cannot rely on the epipolar constraint [10, 11, 12].

Roger Mohr and coworkers [13] have used some ideas from projective geometry to perform reconstruction of the world from a number of point correspondences. However, neither do their clearly distinguish between the three main classes of reconstructions nor do they abstract them from the relevant camera geometry.

Richard Hartley and coworkers [14] have developed simultaneously and independently of [15] a method to compute a projective reconstruction of the world from point correspondences. The second paper also included a preliminary discussion of the affine reconstruction case.

More recently, Amnon Shashua [16, 17, 18] has developed a set of similar ideas which differ slightly from those expressed in [15, 14] by the fact that, through the use of an extra plane of reference, he introduces a special projective invariant which allows him to elegantly predict the position of image points in other views (for a related approach, see [19]). No attempt has been made in this work to relate the three types of possible reconstruction.

The ideas developed in the present paper are closely related to those expressed by Luong and Viéville [20] who also looked at the problem of representing systems of cameras in the framework of projective, affine and euclidean geometries. Their main emphasis was on the characterization of invariant representations for the perspective projection operation performed by the cameras while we are interested here in obtaining invariant representations of the 3-D scene, in determining how the minimum information about the camera geometry, necessary to estimate such representations can be obtained from the images, and how the 3-D representations themselves can be obtained from the images without actually performing an explicit 3-D reconstruction.

3 Stratification of 3-D space: projective, affine and euclidean structures

3.1 Notations

We represent vectors and matrices in boldface, i.e. vector x is noted $\mathbf{x}$. Since a great deal of our discussion deals with geometric entities which can sometime be represented by vectors or matrices we sometimes differentiate between the geometric entity itself, e.g. a point x and its vector representation $\mathbf{x}$.

3.2 3-D space as a projective space

We will first consider that the world is embedded in a projective space of dimension three noted $\mathcal{P}^3$. Similarly, we will consider the retinal plane of a camera as embedded in a projective space of dimension two noted $\mathcal{P}^2$. Since we will also have to consider projective spaces of dimension one, we begin with a brief pedestrian introduction to general projective spaces of any dimension and then specialize to the cases where this dimension equals one, two, or three.

3.2.1 General projective spaces

The use of projective spaces has been made popular in three-dimensional computer graphics and robotics from the early days because they allowed a very compact representation of all changes of coordinate systems as four by four matrixes instead of a rotation matrix and a translation vector. This is because such changes are special cases of linear projective transformations called collineations. Let us look at general projective spaces in more detail.

An n-dimensional projective space, $\mathcal{P}^n$ (in this article we will use only the cases $n = 1, 2, 3$) can be thought of as arising from an $n+1$ dimensional vector space (real or complex) in which we define the following relation between non zero vectors. To help guide the reader's intuition, it is useful to think of a non zero vector as defining a line through the origin. We say that two such vectors $\mathbf{x}$ and $\mathbf{y}$ are equivalent if and only if they define the same line. Mathematically, this can be stated as the fact that there exists a non zero scalar λ such that

$$\mathbf{y} = \lambda \mathbf{x}$$

It is easily verified that this defines an equivalence relation on the vector space minus the zero vector. The equivalence class of a vector is the set of all non zero vectors which are parallel to it (it can be thought of as the line defined by this vector). The set of all equivalence classes is the projective space $\mathcal{P}^n$. A point in that space is called a projective point and is an equivalence class of vectors and can therefore be represented by any vector in the class. If $\mathbf{x}$ is such a vector, then $\lambda\mathbf{x}$, $\lambda \neq 0$ is also in the class and represents the same projective point. The coordinates of any vector in the equivalence class are the coordinates of the corresponding projective point. They are therefore not all equal to zero (we have excluded the zero vector from the beginning) and defined up to a scale factor. It is sometimes useful to differeintiate between the projective point noted x and one of its coordinate vectors noted $\mathbf{x}$.

What do we have so far? a projective point in $\mathcal{P}^n$ is represented by an $n+1$-vector of coordinates $\mathbf{x} = [x_1, \dots x_{n+1}]^T$, where at least one of the x_i is non-zero. The numbers x_i are sometimes called the homogeneous or projective coordinates of the point, and the vector $\mathbf{x}$ is called a *coordinate vector*.

Two $n+1$-vectors $[x_1, \ldots x_{n+1}]^T$ and $[y_1, \ldots y_{n+1}]^T$ represent the same point if and only if there exists a non-zero scalar λ such that $x_i = \lambda y_i$ for $1 \leq i \leq n+1$. Therefore, the correspondence between points and coordinate vectors is not one-to-one and this makes the application of linear algebra to projective geometry a little more complicated.

Collineations We now look at the linear transformations of a projective space. An $(n+1)\times(n+1)$ matrix $\mathbf{A}$ such that $det(\mathbf{A})$ is different from 0 defines a linear transformation or *collineation* from $\mathcal{P}^n$ into itself. It is easy to see that the set of collineations is a group for the usual operation of matrix multiplication. This group is also known as the *projective group* The matrix associated with a given collineation is defined up to a nonzero scale factor, which we usually denote by:

$$\rho \mathbf{y} = \mathbf{A}\mathbf{x} \quad \text{and also} \quad \mathbf{x} \overline{\wedge} \mathbf{y}$$

Quite often we will omit for simplicity the factor ρ and write simply $\mathbf{y} = \mathbf{A}\mathbf{x}$. The reader must remember that this is a *projective* equality, equivalent to the equality of n ratios.

Projective basis Another important notion is that of a projective basis. This is the extension to projective spaces of the idea of coordinate system. A projective basis is a set of $n+2$ points of $\mathcal{P}^n$ such that no $n+1$ of them are linearly dependent. A set of projective points are linearly independent if, considering any set of coordinate vectors of these points, these vectors are linearly independent. It is readily verified that this is independent of the choice of the coordinate vectors and of the choice of basis vectors. For example, the set $\mathbf{e}_i = [0, \ldots, 1, \ldots, 0]^T$, $i = 1, \cdots, n+1$, where 1 is in the ith position, and $\mathbf{e}_{n+2} = [1, 1, \cdots, 1]^T$, is a projective basis, called the standard projective basis. A projective point of $\mathcal{P}^n$ represented by any of its coordiante vectors $\mathbf{x}$ can be described as a linear combination of any $n+1$ points of the standard basis. For example:

$$\mathbf{x} = \sum_{i=1}^{n+1} x_i \mathbf{e}_i$$

We will use several time in the sequel the following result, borrowed from, for example, [21] and the proof of which can be found in [15].

Proposition 1 *Let $\mathbf{x}_1, \cdots, \mathbf{x}_{n+2}$ be $n+2$ coordinate vectors of points in $\mathcal{P}^n$, no $n+1$ of which are linearly dependent, i.e., a projective basis. If $\mathbf{e}_1, \cdots, \mathbf{e}_{n+1}, \mathbf{e}_{n+2}$ is the standard projective basis, there exist nonsingular matrices $\mathbf{A}$ such that $\mathbf{A}\mathbf{e}_i = \lambda_i \mathbf{x}_i$, $i = 1, \cdots, n+2$, where the λ_i are non-zero scalars; any two matrices with this property differ at most by a scalar factor.*

This proposition tells us that any projective basis can be transformed, via a collineation into the standard projective basis.

Change of projective basis Let us consider two sets of $n+2$ points represented by the coordinate vectors $\mathbf{x}_1, \cdots, \mathbf{x}_{n+2}$ and $\mathbf{y}_1, \cdots, \mathbf{y}_{n+2}$. It can be proved that if the points in these two sets are in general position, there exists a unique collineation that maps the first set of points onto the second.

Proposition 2 *If $\mathbf{x}_1, \cdots, \mathbf{x}_{n+2}$ and $\mathbf{y}_1, \cdots, \mathbf{y}_{n+2}$ are two sets of $n+2$ coordinate vectors such that in either set no $n+1$ vectors are linearly dependent, i.e., form two projective basis, then there exists a non-singular $(n+1)\times(n+1)$ matrix $\mathbf{P}$ such that $\mathbf{P}\mathbf{x}_i = \rho_i \mathbf{y}_i$, $i = 1, \cdots, n+2$, where the ρ_i are scalars, and the matrix $\mathbf{P}$ is uniquely determined apart from a scalar factor.*

This proposition shows that a collineation is defined by $n+2$ pairs of corresponding points. The proof can be found for example in [22].

3.2.2 Projective lines, planes and spaces

We illustrate these general notions on three examples which will be used in the rest of the paper.

The projective line The space $\mathcal{P}^1$ is known as the projective line. It is the simplest of all projective spaces, which is the first reason why we start with it. The second reason is that many structures embedded in higher dimensional projective spaces have the same structure as $\mathcal{P}^1$.

The standard projective basis of the projective line is $\mathbf{e}_1 = [1, 0]^T$, $\mathbf{e}_2 = [0, 1]^T$, and $\mathbf{e}_3 = [1, 1]^T$. A point on the line can be written as

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 \tag{1}$$

with x_1 and x_2 not both equal to 0. Let us consider a subset of $\mathcal{P}^1$ of the points such that $x_2 \neq 0$. This is the same as excluding the point represented by $\mathbf{e}_1$. Now since the homogeneous coordinates are defined up to a scalar, these points are described by a parameter α, $-\infty \leq \alpha \leq +\infty$ so that

$$\mathbf{x} = \alpha\mathbf{e}_1 + \mathbf{e}_2$$

where $\alpha = \frac{x_1}{x_2}$. The parameter α is often called the *projective parameter* of the point. Note that the point represented by $\mathbf{e}_2$ has projective parameter equal to 0.

We now define the very important concept of the cross-ratio, which is a quantity that remains invariant under the group of collineations. Let a, b, c, d be four points of $\mathcal{P}^1$ with their respective projective parameters α_a, α_b, α_c, α_d. Then the cross-ratio $\{a,\, b;\, c,\, d\}$ is defined to be

$$\{a,\, b;\, c,\, d\} = \frac{\alpha_a - \alpha_c}{\alpha_a - \alpha_d} : \frac{\alpha_b - \alpha_c}{\alpha_b - \alpha_d} \tag{2}$$

The significance of the cross-ratio is that it is invariant under collineations of $\mathcal{P}^1$. In particular, $\{a,\, b;\, c,\, d\}$ is independent of the choice of coordinates in $\mathcal{P}^1$. Note that the collineations of $\mathcal{P}^1$ are usually called homographies.

The projective plane The space $\mathcal{P}^2$ is known as the projective plane. A point in $\mathcal{P}^2$ is defined by three numbers, not all zero, $(x_1,\ x_2,\ x_3)$. They form a coordinate vector $\mathbf{x}$ defined up to a scale factor. In $\mathcal{P}^2$, there are objects other than points, such as lines. A line is also defined by a triplet of numbers $(u_1,\ u_2,\ u_3)$, not all zero. They form a coordinate vector $\mathbf{u}$ defined up to a scale factor. The equation of the line is

$$\sum_{i=1}^{3} u_i x_i = 0 \tag{3}$$

Formally, there is no difference between points and lines in $\mathcal{P}^2$. This is known as the *principle of duality*. A point represented by $\mathbf{x}$ can be thought of as the set lines through it. These lines are represented by the coordinate vectors $\mathbf{u}$ satisfying $\mathbf{u}^T\mathbf{x} = 0$. This is sometimes referred to as the *line equation* of the point. Inversely, a line represented by $\mathbf{u}$ can be thought of as the set of points represented by $\mathbf{x}$ and satisfying the same equation, called the *point equation* of the line. The principle of duality is a statement about theorems: given a theorem about points and lines, interchange the roles of the points and lines, and adjust the wording accordingly then the new statement will also be true.

Let us now generalize the notion of cross-ratio, introduced in the previous section for four points of $\mathcal{P}^1$, to four lines of $\mathcal{P}^2$ intersecting at a point. Given four lines l_1, l_2, l_3, l_4 of $\mathcal{P}^2$ that intersect at a point, their cross-ratio $\{l_1, l_2; l_3, l_4\}$ is defined as the cross-ratio $\{P_1, P_2; P_3, P_4\}$ of their four points of intersection with any line l not going through their point of intersection. This value is of course independent of the choice of l.

There is a structure of the projective plane that has numerous applications, especially in stereo and motion. The name of this structure is the *pencil of lines.* It is the set of lines in $\mathcal{P}^2$ passing through a fixed point. This is a one-dimensional projective space known as a pencil of lines. Let us consider two lines l_1 and l_2 of the pencil represented by their coordinate vectors $\mathbf{u}_1$ and $\mathbf{u}_2$. Any line l of the pencil goes through the point of intersection of l_1 and l_2 represented by $\mathbf{u}_1 \wedge \mathbf{u}_2$. Thus, its coordinate vector $\mathbf{u}$ satisfies $\mathbf{u}^T(\mathbf{u}_1 \wedge \mathbf{u}_2) = 0$, or equivalently

$$\mathbf{u} = \alpha \mathbf{u}_1 + \beta \mathbf{u}_2$$

for two scalars α and β. This equation is formally equivalent to equation (1), and therefore the structure of a pencil of lines is the same as that of the projective line $\mathcal{P}^1$.

Another, perhaps more elegant, way of proving this result is to apply the principle of duality: the set of lines going through a point is the dual of the set of points on a line, i.e., a projective line!

Collineations of $\mathcal{P}^2$ are defined by 3×3 invertible matrices, defined up to a scalar factor. According to proposition 1, such a collineation is defined by 4 pairs of corresponding points. Collineations transform points, lines, and pencils of lines into points, lines, and pencils of lines, and preserve cross-ratios.

In the projective plane, the class of conic curves is especially important for reasons which will become apparent in sections 4.4 and 7. We give some simple properties of conics that will be used in later sections. A conic ω is a curve defined by the locus of points of the projective plane that satisfy the equation

$$S(\mathbf{x}) = \sum_{i,j=1}^{3} a_{ij} x_i x_j = 0$$

where the scalars a_{ij} satisfy $a_{ij} = a_{ji}$ for all i, j and hence form a 3×3 symmetric matrix $\mathbf{A}$. We can rewrite this equation in matrix form as

$$S(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} = 0$$

$\mathbf{A}$ is defined up to a scale factor and thus the conic depends on five independent parameters. We consider only in the following non-singular conics for which matrix $\mathbf{A}$ is invertible.

Let y and z be two points of the plane represented by $\mathbf{y}$ and $\mathbf{z}$, respectively. A variable point on the line $\langle y, z \rangle$ with projective parameter θ is represented by $\mathbf{y} + \theta \mathbf{z}$, and this point lies on the conic ω if and only if

$$S(\mathbf{y} + \theta \mathbf{z}) = 0$$

By expanding this and grouping terms of similar degrees in θ we have

$$S(\mathbf{y}) + 2\theta S(\mathbf{y}, \mathbf{z}) + \theta^2 S(\mathbf{z}) = 0 \tag{4}$$

where

$$S(\mathbf{y}, \mathbf{z}) = \mathbf{y}^T \mathbf{A} \mathbf{z} = S(\mathbf{z}, \mathbf{y})$$

This means that, in general, there are two points of intersection of the line $\langle y, z \rangle$ with the conic ω. These intersection points can be real or complex and are obtained by solving the quadratic equation (4). The two points are the same if and only if the following relation holds

$$S(\mathbf{y}, \mathbf{z})^2 - S(\mathbf{y}) S(\mathbf{z}) = 0$$

If we consider that the point y is fixed, this equation is quadratic in the coordinates of z: it is the equation of the two tangents from y to ω. Specializing further, if y belongs to ω, $S(\mathbf{y}) = 0$ and the equation of the tangents becomes

$$S(\mathbf{y}, \mathbf{z}) = 0$$

which is linear in the coordinates of z: there is only one tangent to the conic at a point of the conic! note that this tangent l is represented by the vector $\mathbf{l} = \mathbf{A}\mathbf{y}$. We see that when y varies along the conic, it satisfies the equation $\mathbf{y}^T\mathbf{A}\mathbf{y} = 0$ and thus the tangent l satisfies the equation $\mathbf{l}^T\mathbf{A}^{-T}\mathbf{l} = 0$. This shows that the tangents to a conic ω defined by the matrix $\mathbf{A}$ (which we assume to be of rank 3) can be thought of belonging to a conic ω^* in the dual plane defined by a matrix proportional to $\mathbf{A}^{-T}$. This conic is called the *dual* conic of the conic ω. Let $\mathbf{B}$ be the matrix of cofactors of matrix $\mathbf{A}$. Since $\mathbf{A}^{-1} = \frac{1}{det(\mathbf{A})}\mathbf{B}^T$, we conclude that we can use $\mathbf{B}$ for representing ω^* instead of $\mathbf{A}^{-T}$.

Related to these ideas are those of poles and polars which we will use in section 4.4.2. Given a point x represented by the vector $\mathbf{x}$, the polar of x with respect to the conic ω defined by the matrix $\mathbf{A}$ is the line represented by the vector $\mathbf{A}\mathbf{x}$. Therefore, the relation $S(\mathbf{x}, \mathbf{y}) = 0$ is equivalent to saying that the point y is on the polar of the point x and vice versa. Given a line l represented by $\mathbf{l}$, the pole of l with respect to the conic ω is the point x whose polar is l. Assuming that the matrix $\mathbf{A}$ is of rank 3, this point is therefore represented by the vector $\mathbf{A}^{-1}\mathbf{l}$.

The projective space The space $\mathcal{P}^3$ is known as the projective space. A point $\mathbf{x}$ in $\mathcal{P}^3$ is defined by four numbers, (x_1, x_2, x_3, x_4), not all zero. They form a coordinate vector $\mathbf{x}$ defined up to a scale factor. In $\mathcal{P}^3$, there are objects other than just points and lines, such as planes. A plane is also defined as a four-tuple of numbers (u_1, u_2, u_3, u_4), not all zero, which form a coordinate vector $\mathbf{u}$ defined up to a scale factor. The equation of this plane is then

$$\sum_{i=1}^{4} u_i x_i = 0 \tag{5}$$

This shows that the same principle of duality that exists in $\mathcal{P}^2$ between points and lines exists in $\mathcal{P}^3$ between points and planes. A point represented by $\mathbf{x}$ can be thought of as the set of planes through it. These planes are represented by $\mathbf{u}$ satisfying $\mathbf{u}^T\mathbf{x} = 0$, which is called the *plane equation* of the point. Inversely, a plane represented by $\mathbf{u}$ can be thought of as the set of points represented by $\mathbf{x}$ and satisfying the same equation, called the *point equation* of the plane.

Let us generalize the notion of cross-ratio introduced for four points of $\mathcal{P}^1$ and four lines of $\mathcal{P}^2$ intersecting at a point, to four planes of $\mathcal{P}^3$ intersecting at a line. Given four planes $\pi_1, \pi_2, \pi_3, \pi_4$ of $\mathcal{P}^3$ that intersect at a line l, their cross-ratio $\{\pi_1, \pi_2; \pi_3, \pi_4\}$ is defined as the cross-ratio $\{l_1, l_2; l_3, l_4\}$ of their four lines of intersection with any plane π not going through l. This is of course independent of the choice of π. The cross-ratio can also be defined as the cross-ratio of the four points of intersection of any line, not lying in any of the four planes, with the four planes. This is also independent of the choice of the line.

The structure that is analogous to the pencils of lines of $\mathcal{P}^2$ is the *pencil of planes*, the set of all the planes that intersect at a given line. This structure is also a projective space of dimension one, an analog to the space $\mathcal{P}^1$ since, using the principle of duality, a pencil of planes is projectively equivalent to a set of points on a same line (this line is the dual of the line of intersection of the planes).

Let us use this concept to show that the ratios of the projective coordinates of a point M in a given projective basis can be interpreted as a cross-ratio. In order to do this, we assume without loss of generality (thanks to proposition 1) that the projective basis is the standard projective basis of the projective space. We consider the four planes $\pi_1 \equiv (e_1, e_2, e_3)$, $\pi_2 \equiv (e_1, e_2, e_4)$, $\pi_3 \equiv (e_1, e_2, e_5)$ and $\pi_4 \equiv (e_1, e_2, M)$ which all go through the line $\langle e_1, e_2 \rangle$. M is a point of projective coordinates (p, q, r, s). The equations of these four planes are readily shown to be equal to

$$\begin{array}{rcc} \pi_1 & : & x_3 = 0 \\ \pi_2 & : & x_2 = 0 \\ \pi_3 & : & x_3 - x_2 = 0 \\ \pi_4 & : & r x_3 - s x_2 = 0 \end{array}$$

We can use the two planes π_1 and π_2 and the plane of equation $x_2 + x_3 = 0$ as the projective basis of the pencil of planes of axis $\langle e_1, e_2 \rangle$. Looking at the previous equations, we see that

$$\begin{array}{rcl} \pi_3 & = & \pi_1 - \pi_2 \\ \pi_4 & = & r\pi_1 - s\pi_2 \end{array}$$

and therefore, the cross-ratio $\{\pi_1, \pi_2; \pi_3, \pi_4\}$ is equal to $\frac{0+1}{0+\frac{s}{r}} : \frac{\infty+1}{\infty+\frac{s}{r}} = \frac{r}{s}$ i.e. to the ratio of the third to the fourth projective coordinates of M. This is shown in figure 1. We will use this remarkable relation in sections 5.2 and 6.2.

Figure 1 approximately here.

Collineations of $\mathcal{P}^3$ are defined by 4×4 invertible matrices defined up to a scale factor. According to proposition 1, such a collineation is defined by 5 pairs of corresponding points. Collineations transform points, lines, planes, and pencils of planes into points, lines, planes and pencils of planes, preserving cross-ratios.

3.3 3-D space as an affine space

We now describe the second stratum that we will consider. The idea is to think of the world (and for that matter of the retina) as an affine space embedded in the corresponding projective space, i.e. $\mathcal{P}^3$ and $\mathcal{P}^2$, respectively. We consider first the case of the retina, i.e. of the projective plane and then the case of the world, i.e. of the projective space. But to make things more clear we start with the projective line $\mathcal{P}^1$ and show how we can associate an affine line to it.

3.3.1 Projective and affine lines

The point represented by $\mathbf{e}_1$ is called the point at infinity of the line $\mathcal{P}^1$. It is defined by the linear equation $x_2 = 0$. The reason for this terminology is that if we think of the projective line as containing the usual affine line under the correspondence $\alpha \rightarrow \alpha\mathbf{e}_1 + \mathbf{e}_2$, then the projective parameter α of the point gives us a one-to-one correspondence between the projective and affine lines for all values of α different from ∞ (the affine line is simply the set of real numbers). The values $\alpha = \pm\infty$ correspond to the point $\mathbf{e}_1$, which is outside the affine line but is the limit of points of the affine line with large values of α. This turns out to be an extremely useful interpretation of the relationship between the affine and projective lines and, as we show later, can be generalized to higher dimensions. Note that the choice of e_1 as the point at infinity is arbitrary and any other point will do equally well.

3.3.2 Projective and affine planes

The line at infinity Suppose we choose a line in the projective plane. Without loss of generality, we can assume its equation to be $x_3 = 0$. We call this line the *line at infinity* of $\mathcal{P}^2$, denoted l_∞. Just as in the previous case of the projective line, the choice of l_∞ as the line at infinity is arbitrary and any other line will do equally well. But it is worth noting that points and lines at infinity can be chosen consistently: i.e. if l_∞ is the line at infinity of $\mathcal{P}^2$ and l a line of $\mathcal{P}^2$ different of l_∞, then $l \cap l_\infty$ is a suitable choice for the point at infinity on l (see next paragraph). The reason for this terminology is that we can think of the projective plane as containing the usual affine plane under the correspondence $\mathbf{X} = [X_1, X_2]^T \to [X_1, X_2, 1]^T$ or $X_1\mathbf{e}_1 + X_2\mathbf{e}_2 + \mathbf{e}_3$. This is a one-to-one correspondence between the affine plane and the projective plane minus the line of equation $x_3 = 0$. For each projective point of coordinates (x_1, x_2, x_3) that is not on that line, we have

$$X_1 = \frac{x_1}{x_3} \quad X_2 = \frac{x_2}{x_3} \tag{6}$$

If $X_1 \to \infty$ while X_2 does not, we obtain $\mathbf{e}_1$, which is on l_∞. Similarly, when $X_2 \to \infty$ while X_1 does not, we obtain $\mathbf{e}_2$.

Each line in the projective plane of the form of equation (3) intersects l_∞ at the point $(-u_2, u_1, 0)$, which is that line's point at infinity. Note that the vector $[-u_2, u_1]^T$ gives the direction of the affine line of equation $u_1X_1 + u_2X_2 + u_3 = 0$. This gives us a neat interpretation of the line at infinity: each point on that line, with coordinates $(x_1, x_2, 0)$, can be thought of as a direction in the underlying affine plane, the direction parallel to the vector $[x_1, x_2]^T$. Indeed, it does not matter if x_1 and x_2 are defined only up to a scale factor since the direction does not change. We will use this observation later.

As a first, and very useful, application of the idea of thinking about the affine plane as embedded in a projective plane, let us consider the case of two parallel (but not identical) lines. Since by definition these two lines have the same direction parallel to the vector $[-u_2, u_1]^T$, this means that if we consider them as *projective* lines of the projective plane, they intersect at the point represented by $[-u_2, u_1, 0]^T$ of l_∞. Therefore, two distinct parallel lines intersect at a point of l_∞: thinking of the affine plane as embedded in the projective plane allows to avoid considering special cases.

Affine transformations of the plane We have seen that there is a one-to-one correspondence between the usual affine plane and the projective plane minus the line at infinity. In the affine plane, we know that an affine transformation defines a correspondence $\mathbf{X} \to \mathbf{X}'$, which can be expressed in matrix form as

$$\mathbf{X}' = \mathbf{B}\mathbf{X} + \mathbf{b} \tag{7}$$

where $\mathbf{B}$ is a 2×2 matrix of rank 2, and $\mathbf{b}$ is a 2×1 vector. From this equation it is clear that these transformations form a group called the *affine group*, which is a subgroup of the projective group. This subgroup has the interesting property that it preserves the line at infinity.

Indeed, let $\mathbf{A}$ be the matrix of a collineation of $\mathcal{P}^2$ that leaves l_∞ invariant. The matrix $\mathbf{A}$ can be written as

$$\mathbf{A} = \begin{bmatrix} \mathbf{C} & \mathbf{c} \\ \mathbf{0}_2^T & a_{33} \end{bmatrix}$$

where $\mathbf{C}$ is a 2×2 matrix and $\mathbf{c}$ is a 2×1 vector. The condition that the rank of $\mathbf{A}$ is 3 implies that $a_{33} \neq 0$ and the rank of $\mathbf{C}$ is equal to 2. Using the equations (6) we can write equation (7) with $\mathbf{B} = \frac{1}{a_{33}}\mathbf{C}$ and $\mathbf{b} = \frac{1}{a_{33}}\mathbf{c}$.

3.3.3 Projective and affine spaces

The plane at infinity Similarly to the previous case, let us assume that we choose a plane in the projective space $\mathcal{P}^3$. Without loss of generality, we can assume its equation to be $x_4 = 0$. We call this plane the plane at infinity π_∞ of $\mathcal{P}^3$. The reason for this terminology, just as in the case of $\mathcal{P}^2$, is that it is possible to think of the projective space as containing the usual affine space under the correspondence $\mathbf{X} = [X_1, X_2, X_3]^T \rightarrow [X_1, X_2, X_3, 1]^T$ or $X_1\mathbf{e}_1 + X_2\mathbf{e}_2 + X_3\mathbf{e}_3 + \mathbf{e}_4$. This is a one-to-one correspondence between the affine space and the projective space minus the plane at infinity of equation $x_4 = 0$. For each projective point of coordinates (x_1, x_2, x_3, x_4) not in that plane, we have

$$X_1 = \frac{x_1}{x_4} \quad X_2 = \frac{x_2}{x_4} \quad X_3 = \frac{x_3}{x_4}$$

Similarly to the case of $\mathcal{P}^2$, points, lines, and planes at infinity can be chosen consistently in $\mathcal{P}^3$: if π_∞ is the plane at infinity of $\mathcal{P}^3$ and π (resp. l) is a plane (resp. a line) of $\mathcal{P}^3$ not equal to (resp. not included in) π_∞, then $\pi \cap \pi_\infty$ (resp. $l \cap l_\infty$) is a suitable choice for the line at infinity (resp. the point at infinity) on π (resp. on l). Hence each plane of equation (5) intersects the plane at infinity along a line that is its line at infinity.

As in the case of the projective plane, it is often useful to think of the points in the plane at infinity as the set of directions of the underlying affine space. For example, the point of projective coordinates $[x_1, x_2, x_3, 0]^T$ represents the direction parallel to the vector $[x_1, x_2, x_3]^T$ and indeed, it does not matter whether x_1, x_2, x_3 are defined up to a scale factor, since the direction does not change. An analysis similar to the one done in the two-dimensional case shows that two distinct affine parallel planes can be considered as two projective planes intersecting at a line in the plane at infinity π_∞.

Affine transformations of the space In a similar fashion to the case of the projective plane, we can consider the subset of the projective group that preserves the plane at infinity. This set is a subgroup of the projective group called the affine group, and the transformations can be written in the same way as in equation (7)

$$\mathbf{X}' = \mathbf{B}\mathbf{X} + \mathbf{b} \tag{8}$$

where matrix $\mathbf{B}$ is 3×3 and has rank 3, and $\mathbf{b}$ is a 3×1 vector.

3.4 3-D space as a euclidean space

As a final stratum, and to complete our trilogy, we want to think of the world (and for that matter of the retina) as a euclidean space embedded in the previous affine space. We consider first the case of the retina, i.e. of the affine and projective plane and then the case of the world, i.e. of the affine and projective space.

3.4.1 Euclidean transformations of the plane: the absolute points

We can further specialize the set of affine transformations of the plane and require that they preserve not only the line at infinity but also two special points on that line called the *absolute points* I and J with coordinates $(1, \pm i, 0)$, where $i = \sqrt{-1}$.

This imposes constraints on matrix $\mathbf{B}$ in equation (7). Since we insist that I and J remain invariant, we have

$$\frac{1}{i} = \frac{b_{11}1 + b_{12}i + b_1 0}{b_{21}1 + b_{22}i + b_2 0}$$

$$\frac{1}{-i} = \frac{b_{11}1 - b_{12}i + b_1 0}{b_{21}1 - b_{22}i + b_2 0}$$

which yields

$$(b_{11} - b_{22})i - (b_{12} + b_{21}) = 0$$

and

$$-(b_{11} - b_{22})i - (b_{12} + b_{21}) = 0$$

Therefore $b_{11} - b_{22} = b_{12} + b_{21} = 0$ and we can write

$$\mathbf{X}' = c \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \mathbf{X} + \mathbf{b} \tag{9}$$

with $c > 0$ and $0 \leq \alpha < 2\pi$. This class of transformations is sometimes called the class of similitudes. It forms a subgroup of the affine group and therefore of the projective group. This group is called the similitude group or the euclidean transformations group. The affine point represented by $\mathbf{X}$ is first rotated by α around the origin, then scaled by c, and translated by $\mathbf{b}$. If we specialize the class of transformations further by assuming that $c = 1$, we obtain another subgroup called the group of (proper) rigid displacements.

As an application of the use of the absolute points, we show how they can be used to define the angle between two lines. The angle α between two lines l_1 and l_2 can be defined by considering their point of intersection m and the two lines i_m and j_m joining m to the absolute points I and J (see figure 2). The angle is given by Laguerre formula:

$$\alpha = \frac{1}{2i} \log(\{l_1, l_2; i_m, j_m\}) \tag{10}$$

Which is also equal to the cross-ratio of the four points I, J, m_1, m_2 of intersection of the four lines with the line at infinity l_∞.

Because $e^{i\pi} = \cos\pi + i\sin\pi = -1$, we see that if the cross-ratio $\{l_1, l_2; i_m, j_m\}$ is equal to -1, the two lines l_1 and l_2 are perpendicular.

Figure 2 approximately here.

3.4.2 Euclidean transformations of the space: the absolute conic

We can also further specialize the affine transformations of the space and require that they leave a special conic invariant. This conic, Ω, is obtained as the intersection of the quadric of equation $\sum_{i=1}^4 x_i^2 = 0$ with π_∞

$$\sum_{i=1}^{4} x_i^2 = x_4 = 0$$

The conic Ω is also called the *absolute conic*. Note that in π_∞, Ω can be interpreted as a circle of radius $i = \sqrt{-1}$, an imaginary circle! Therefore, all its points have complex coordinates in the standard projective basis and if m is a point of Ω, then $\overline{m}$, the complex conjugate point, is also on Ω since the absolute conic is defined by equations with real coefficients. It is not difficult to show that the affine transformations that keep Ω invariant can be written

$$\mathbf{X}' = c\mathbf{C}\mathbf{X} + \mathbf{b} \tag{11}$$

where $c > 0$ and $\mathbf{C}$ is orthogonal, i.e., satisfies the equation $\mathbf{C}\mathbf{C}^T = \mathbf{I}$ (see for example [22]). As in the two-dimensional case, this subset of the affine group is a subgroup called the *similitude group*. Similarly, the subset of the similitude group where $c = 1$ is also a subgroup called the group of (proper) rigid displacements.

3.5 Conclusion

We have shown how the world (resp. the retina) can be considered as a succession of strata. Each stratum corresponds to a specific geometric structure that we impose on the world (resp. on the retina). These geometric structures can be ordered in a hierarchy, from general (i.e. the projective structure), to more specialized (i.e the euclidean structure). To each stratum corresponds a group of transformations. These three groups are included in each other in a group theoretical sense: the group of similitudes is a subgroup of the affine group which is itself a subgroup of the projective group. Each group leaves some geometric quantities invariant: the cross-ratio is the most notable one for the projective group, the ratio of the lengths of two parallel vectors is the most notable one for the affine group, and angles and ratios of lengths are the most notable ones for the group of similitudes. We will see more of these invariants in the next sections.

4 Camera geometry

Let us now turn to the sensor that we use to measure the world: the camera. We model classically a camera as a pinhole. This has proven to be an excellent approximation for most practical purposes. Even though it is important to keep in mind that the pinhole model is only an approximation, albeit usually a very good one, of a real physical camera, we hope to convince the reader in what follows of the usefulness of forgetting for some time the actual physical device and of thinking of the camera as a projective geometric engine. We develop this line of thought in the following sections and relate the projective modelling of the camera to the three strata which were presented in the previous section.

4.1 The perspective projection model

This projective engine maps the two-dimensional projective space $\mathcal{P}^3$ onto the two-dimensional projective plane $\mathcal{P}^2$ by perspective projection from a center of projection C (the optical center of the camera) onto a plane $\mathcal{R}$ not containing C (the retinal plane of the camera).

This projection operation is projective linear in the sense that if we choose a projective basis of $\mathcal{P}^3$ and a projective basis of $\mathcal{P}^2$, the correspondence between a point M of $\mathcal{P}^3$ represented by $\mathbf{M}$ and its image m of $\mathcal{P}^2$ represented by $\mathbf{m}$ can be written in vector form

$$\mathbf{m} = \mathbf{PM} \tag{12}$$

where $\mathbf{P}$ is a 3×4 matrix of rank 3 defined up to the multiplication with a non zero scalar. This matrix depends therefore upon 11 parameters and is called the *perspective projection* matrix of the camera.

Note that if we change projective basis in the world by $\mathbf{M}' = \mathbf{KM}$ ($\mathbf{K}$ a 4×4 matrix of rank 4) and in the retinal plane by $\mathbf{m}' = \mathbf{Hm}$ ($\mathbf{H}$ a 3×3 matrix of rank 3), then the perspective projection matrix becomes $\mathbf{P}' = \mathbf{HPK}^{-1}$.

Given the perspective projection matrix $\mathbf{P}$ and without any further assumption about the world, we can recover the coordinates of the optical center in the projective basis of the world. Indeed, the optical center is the point for which the perspective projection is not defined, it has no image and must therefore satisfy

$$\mathbf{PC} = \mathbf{0}$$

which shows that C is represented by any non zero vector of the nullspace of the matrix $\mathbf{P}$ which is by definition of dimension 1 since $rank(\mathbf{P}) = 3$.

4.2 Two cameras and the fundamental matrix: the projective stratum

If we now consider a binocular stereo rig, we can bring in some more geometric information which has profound implications for computer vision problems. Let us call C' the optical center of the second camera and $\mathcal{R}'$ its retinal plane. The line $\langle C, C' \rangle$ intersects $\mathcal{R}$ (resp. $\mathcal{R}'$) in a point that we denote by e (resp. e'). These two points are called the epipoles of the stereo rig. By construction, any plane containing the line $\langle C, C' \rangle$, called an epipolar plane, intersects $\mathcal{R}$ (resp. $\mathcal{R}'$) along a line going through the epipole e (resp. along a line going through the epipole e'), see figure 3 . Two such lines are called corresponding epipolar lines and have an immense importance for stereo algorithms. We can rephrase the situation in projective terms by saying that the pencil of epipolar planes induces in each retinal plane a pencil of epipolar lines. According to section 3 these three pencils are projective lines, i.e. one-dimensional projective spaces $\mathcal{P}^1$.

Figure 3 approximately here.

The fundamental property of this geometric construction is that the "natural" correspondence between the two pencils of epipolar lines is projective linear, it is a homography between the two pencils considered as projective lines. The "natural" correspondence consists in associating with each epipolar line of the first pencil the corresponding epipolar line of the second, i.e. the intersection of the epipolar plane defined by the first one and the two optical centers with the second retinal plane. The reason why it is an homography is because it is one to one and preserves cross-ratios: The cross-ratio of four lines of the first pencil is equal to the cross-ratio of the four corresponding epipolar planes which is equal to the cross-ratio of the four corresponding epipolar lines in the second retinal plane. This homography is at the heart of many of the ideas which will be presented in the next sections (see figure 4). We call it the *epipolar* homography.

Figure 4 approximately here.

Having presented the geometric viewpoint, let us now present its algebraic face. In order to do this, we will adopt a slightly different view, namely we will characterize the relationship between a point m in the first retinal plane and its epipolar line l'_m in the second. This correspondence is also clearly projective linear (it is a projective linear mapping between the first retinal plane considered as a $\mathcal{P}^2$ and the dual of the second retinal plane, also considered as a $\mathcal{P}^2$) and therefore there exists a 3×3 matrix $\mathbf{F}$, defined up to a scale factor, such that

$$\mathbf{l}'_m = \mathbf{F}\mathbf{m}$$

The matrix $\mathbf{F}$ is *not* of rank 3 since if m coincides with the epipole e its epipolar line is undefined and therefore

$$\mathbf{F}\mathbf{e} = \mathbf{0}$$

Let us now further consider a point m' on the epipolar line l'_m of m. This point satisfies the relation

$$\mathbf{m}'^T\mathbf{F}\mathbf{m} = 0 \tag{13}$$

which shows that the epipolar line $l_{m'}$ in the first retinal plane of m' is represented by $\mathbf{F}^T\mathbf{m}'$:

$$\mathbf{l}_{m'} = \mathbf{F}^T\mathbf{m}'$$

In particular we have

$$\mathbf{F}^T\mathbf{e}' = \mathbf{0}$$

The matrix $\mathbf{F}$ expresses algebraically the epipolar correspondence between the two retinal planes. It is called the fundamental matrix [2, 1, 4]. Its rank is in general equal to 2 and it therefore depends upon seven free parameters. A set of such parameters which have a neat geometric interpretation are the four ratios of projective coordinates of the two epipoles and the three ratios of the coefficients of the homography between the two pencils of epipolar lines.
Just as in the one-camera case where we related the optical center to the perspective projection $\mathbf{P}$, in the two-cameras case, we can also relate the fundamental matrix $\mathbf{F}$ to the two perspective projection matrices $\mathbf{P}$ and $\mathbf{P}'$. The interested reader is referred to, for example [22].

4.3 Two cameras looking at planes: the affine stratum

According to our discussion of section 3.3, in order to go from a projective representation of the world to an affine representation, we have to identify the plane at infinity. Before explaining how this can be achieved, we will start with a brief description of the relationship between planes in the world and a pair of cameras.

Indeed, planes in the world have very interesting properties with respect to our stereo rig. In effect, a plane π induces in general a projective linear correspondence, a collineation, between the two retinal planes.

This can be readily seen by noting that the perspective projection from π to $\mathcal{R}$ (resp. from π to $\mathcal{R}'$) is one to one if π does not contain the optical center C (resp. the optical center C') and preserves cross-ratios and is therefore a collineation. Composing the inverse of the first collineation with the second defines a collineation from $\mathcal{R}$ to $\mathcal{R}'$ called the collineation induced by π that we note H_π and represent by the 3×3 matrix $\mathbf{H}_\pi$. Even though a collineation of $\mathcal{P}^2$ depends upon 8 parameters, there is no contradiction with the fact that a plane depends upon 3 parameters. Indeed, the collineation is related to the fundamental matrix [2] in the following manner. Let m be a point of $\mathcal{R}$. The point m' of $\mathcal{R}'$ represented by $\mathbf{H}_\pi \mathbf{m}$ is the image in the second camera of the intersection of the optical ray $\langle C, m \rangle$ with π. Therefore it belongs to the epipolar line of m and we have

$$(\mathbf{H}_\pi \mathbf{m})^T \mathbf{F} \mathbf{m} = 0$$

for all points m. This implies that the matrix $\mathbf{H}_\pi^T \mathbf{F}$ is antisymmetric:

$$\mathbf{H}_\pi^T \mathbf{F} + \mathbf{F}^T \mathbf{H}_\pi = \mathbf{0} \tag{14}$$

This imposes six homogeneous constraints on the collineation.

The interaction between the geometry of the stereo rig and H_π can also be seen as follows. Let us consider the point of intersection P of the line $\langle C, C' \rangle$ with π. The images of P in the two retinal planes are the two epipoles e and e' which therefore correspond to each other through H_Π. A consequence of this is that if the fundamental matrix is known (and thus the epipoles), three pairs of corresponding points are sufficient to determine the collineation since a fourth pair (e, e') is already available.

This observation can be turned into a very simple geometric construction: Let us assume that the plane (or its collineation) is represented by three pairs of corresponding points (m_i, m_i'), $i = 1, 2, 3$ and the pair of epipoles (e, e'). Given a point m in the first image, how do we construct its image m' in the second image under the plane collineation? This is shown in figure 5. We construct the point m_{12} intersection of the two lines $\langle m_1, m_3 \rangle$ and $\langle m_2, m \rangle$. The point m_{12}' in the second image at the intersection of the line $\langle m_1', m_3' \rangle$ and the epipolar line $l_{m_{12}}'$ of m_{12} corresponds to m_{12} under the plane collineation since the line projecting to $\langle m_1, m_3 \rangle$ in the first image and to $\langle m_1', m_3' \rangle$ in the second is certainly in the plane. Therefore, the line $\langle m_2', m_{12}' \rangle$ is the image of the line $\langle m_2, m_{12} \rangle$

and its intersection with the epipolar line l'_m of m yields the sought for point m'. We call this procedure the **Point-Plane** procedure.

We can use this procedure for solving another problem which will appear several times in the remaining of the paper. The problem is the following. Given a plane in the world represented either by its collineation or by three pairs of point correspondences, and given a line in the world represented by its pair of images (l, l'), construct the images of the point of intersection of the line with the plane. If the plane is represented by its collineation, we just apply it to the line l, obtaining d'. The point of intersection m' of l' and d' is the image in the second camera of the point M of intersection of the 3-D line with the plane. m is then obtained for example by intersecting the epipolar line of m' with l. If the plane is represented by three pairs of corresponding points (m_i, m'_i), $i = 1, 2, 3$ and the pair of epipoles (e, e') then we can solve our problem very simply by using twice **Point-Plane**. We call the resulting procedure **Line-Plane**.

Figure 5 approximately here.

Returning now to the problem of going from a projective representation of the world to an affine one, we see that the problem is really to obtain at least three pairs of corresponding points which are the images of three points in the plane at infinity in order to estimate the collineation it induces between the two retinal planes. We describe several ways of doing this in section 6.

4.4 Two cameras looking at the absolute conic: the euclidean stratum

The image of the absolute conic in each camera is also a conic and this conic does not change when we move the camera around. This is because, as shown in section 3, the absolute conic is invariant with respect to similitudes of the world and hence to rigid displacements. It is hard to envision, but it is nonetheless true, that the absolute conic is a curve with only complex points (see section 3.4.2) whose image in a camera does not change when the camera is moved about the world. This phenomenon is analog to what happens to the image of a point at infinity when we translate the camera: it does not change either. Both properties are intimately tied to the structure of the world as an affine or euclidean space and to the geometric operation performed by a camera.

4.4.1 One camera and the absolute conic: Measuring the angle between two optical rays

If the image of the absolute conic is known in a camera, it then becomes a metric measurement device that can compute angles between optical rays [22]. This can be readily seen by using Laguerre's formula given in section 3.4.1 as follows. Let m and n be two image points and consider the two optical rays $\langle C, m\rangle$ and $\langle C, n\rangle$. Let us call α the angle (between 0 and π) that they form, let M and N be their intersections with the plane at infinity, and let A and B be the two intersections of the line $\langle M, N\rangle$ with the absolute conic Ω. The angle α between $\langle C, m\rangle$ and $\langle C, n\rangle$ is given by Laguerre 's formula $\frac{1}{2i}\log(\{M, N; A, B\})$. The reason for this is that the line at infinity of the plane defined by the three points C, m, n is the intersection of that plane with the plane at infinity, i.e., the line $\langle M, N\rangle$. The absolute points of that plane are the intersections A and B of that line with the absolute conic Ω.

The cross-ratio $\{M, N; A, B\}$ is preserved under the projection to the retinal plane, and thus the angle between $\langle C, m\rangle$ and $\langle C, n\rangle$ is given by $\frac{1}{2i}\log(\{m, n; a, b\})$, where the points a and b are the "images" of the points A and B.

Since a and b are the two intersections of the line $\langle m, n\rangle$ with ω, this shows that the angle can be computed only from the image ω of the absolute conic. The situation is depicted in figure 6.

Figure 6 approximately here.

In detail, the line $\langle m, n\rangle$ is represented by $\mathbf{m}+\theta\mathbf{n}$. The variable θ is a projective parameter of that line. Point m has projective parameter 0, and point n has projective parameter equal to ∞. The reader should not worry about this, since the magic of the cross-ratio will take care of it!

In order to compute the projective parameters of a and b we apply equation (4) with S being the equation of ω. The projective parameters are the roots of the quadratic equation

$$S(\mathbf{m}) + 2\theta S(\mathbf{m},\, \mathbf{n}) + S(\mathbf{n})\theta^2 = 0$$

Let θ_0 and $\overline{\theta_0}$ be the two roots, which are complex conjuguate. According to equation (2), we have

$$\{m,\, n;\, a,\, b\} = \frac{0-\theta_0}{0-\overline{\theta_0}} : \frac{\infty-\theta_0}{\infty-\overline{\theta_0}}$$

The ratio containing ∞ is equal to 1 (that is the magic!), and therefore

$$\{m,\, n;\, a,\, b\} = \frac{\theta_0}{\overline{\theta_0}} = e^{2iArg(\theta_0)}$$

where $Arg(\theta_0)$ is the argument of the complex number θ_0. In particular, we have $\alpha = Arg(\theta_0)\ (\pi)$. A straightforward computation shows that the two roots are equal to

$$\frac{-S(\mathbf{m},\, \mathbf{n}) \pm i\sqrt{S(\mathbf{m})S(\mathbf{n}) - S(\mathbf{m},\, \mathbf{n})^2}}{S(\mathbf{n})} \tag{15}$$

Simple considerations show that

$$\cos\alpha = -\frac{S(\mathbf{m},\, \mathbf{n})}{\sqrt{S(\mathbf{m})S(\mathbf{n})}} \tag{16}$$

an equation which uniquely defines α between 0 and π. The sine is therefore positive and given by $\sqrt{1-\cos^2\alpha}$.

4.4.2 The absolute conic and the intrinsic parameters

We have seen previously that the image ω of the absolute conic did not change when we moved the camera in space. This, together with the fact that ω contains only complex points, has some strong implications on the coefficients of the equation defining ω. We now examine these consequences. Let $\mathbf{A}$ be the symmetric matrix defining the equation of ω in the retinal plane:

$$S(\mathbf{m}) = \mathbf{m}^T\mathbf{A}\mathbf{m}$$

Since ω does not contain any real point, this means that $S(\mathbf{m})$ is either strictly positive or strictly negative for all points m with real coordinates. Let us assume that it is strictly positive. The quadratic form defined by matrix $\mathbf{A}$ is accordingly positive definite and we can use a theorem which says that a necessary and sufficient condition for a quadratic form to be positive definite is that its matrix can be written as $\mathbf{W}\mathbf{W}^T$, where $\mathbf{W}$ is a lower-triangular matrix with positive diagonal elements [23]. This decomposition is called the Cholesky decomposition of matrix $\mathbf{A}$ and is unique.

If we define $\mathbf{p} = \mathbf{W}^T\mathbf{m}$, the equation of ω can be written $S(\mathbf{p}) = \mathbf{p}^T\mathbf{p}$. The matrix $\mathbf{W}$ can be interpreted as defining a change of projective coordinates in the retinal plane. For reasons which will become clear later, we are more interested in the matrix $\mathbf{W}^{-T}$ which is upper triangular. Let us write this matrix

$$\mathbf{W}^{-T} = \begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix} \tag{17}$$

It is easy to see that $\mathbf{W}^{-T}$, like $\mathbf{W}$, has positive diagonal elements and since $\mathbf{W}$ is, like $\mathbf{A}$, defined up to a scale factor, we can assume that $f = 1$ (f cannot be equal to 0 because otherwise the rank of $\mathbf{A}$ would be less than 3). We also have $a > 0$ and $d > 0$. Changing notations, we write

$$\begin{array}{cc} a = \alpha_u & b = -\alpha_u \cot\theta \\ d = \frac{\alpha_v}{\sin\theta} & c = u_0 \\ e = v_0 & \end{array} \tag{18}$$

These equations uniquely define the five parameters α_u, α_v, u_0, v_0, and θ. This is clear for α_u, u_0, v_0. For θ and α_v, we see that the equation $b = -\alpha_u \cot\theta$ defines θ between 0 and π. Thus the sine is positive and since $d > 0$ this uniquely defines α_v as a positive number.

These parameters have been introduced by several authors from physical and heuristic considerations in the past [22, 2] and are called the intrinsic parameters of the camera. Here they appear without such considerations, as a consequence of the fact that the image of the absolute conic is an imaginary curve.

Let us assume that the retinal plane is an affine plane, which makes sense if we are using the image coordinates provided by the sensor. Note that we consider this affine plane as embedded in a projective plane, in agreement with our general approach. The affine plane can be considered as obtained from the projective plane by throwing away the line at infinity of equation $x_3 = 0$. A point of the affine plane of coordinates (u, v) can be considered as a projective point of coordinates $(u, v, 1)$. Inversely, a projective point of coordinates (x_1, x_2, x_3) not belonging to the line at infinity can be considered as an affine point of coordinates $(\frac{x_1}{x_3}, \frac{x_2}{x_3})$.

We now give an intuitive interpretation of the intrinsic parameters in this context. First, let us determine the center of ω. We know that the center of a conic is the pole of the line at infinity, thus it is the point represented by (see section 3.2.2)

$$\mathbf{A}^{-1}\mathbf{e}_3$$

since the vector $\mathbf{e}_3 = [0, 0, 1]^T$ represents the line at infinity.

The relation $\mathbf{A} = \mathbf{W}\mathbf{W}^T$ implies $\mathbf{A}^{-1} = \mathbf{W}^{-T}\mathbf{W}^{-1}$ and therefore, according to equations (17) and (18), the center c of ω is the point of affine coordinates (u_0, v_0).

Let us now consider the optical ray $\langle C, c\rangle$. We will show that it is perpendicular to all the directions of lines of the retinal plane. In order to do this, let us consider a point m on the line at infinity of the retinal plane. The optical ray $\langle C, m\rangle$ is therefore parallel to the retinal plane. In order to compute the angle between these two optical rays, we simply apply equation (15) to these two points. It is easy to show that $S(\mathbf{m}, \mathbf{c})$ is equal to zero and therefore that the cross-ratio $\{m, c; I, J\}$ is equal to -1. This means that $\langle C, c\rangle$ is perpendicular to $\langle C, m\rangle$ for each point m on the line at infinity of the retinal plane. Since we have seen that each such point represents a line direction in the underlying affine plane, we have proved that the line $\langle C, c\rangle$ is orthogonal to all lines in the retinal plane and therefore to the retinal plane itself. The optical ray $\langle C, c\rangle$ can be considered as the *optical axis* of the camera.

Note that this interpretation is valid only if the original (u, v) plane is a "real" affine plane, i.e. if it has not been *projectively* distorted. In that case the line at infinity represented by $(0, 0, 1)$ is not the real line at infinity and we cannot say anymore that the line $\langle C, c\rangle$ is perpendicular to the retinal plane. We call this problem the problem of the "hidden" projective transformation.

In a similar spirit, we can give an interpretation of the angle θ defined above in terms of the retinal coordinate system. Indeed, let us consider the directions of the u- and v-axes, i.e. the points at infinity of coordinates (1, 0, 0) and (0, 1, 0). The angle α beween these two directions is obtained by applying equation (16):

$$\cos\alpha = -\cos\theta$$

which shows that the angle is $\pi - \theta$ or θ since we are actually measuring angles between lines (in order to be able to talk about angles between vectors, we would have to orient the plane and this is equivalent to distinguishing between the two absolute points).

Since the matrix $\mathbf{W}^T$ is upper-triangular, it defines a collineation which preserves the line at infinity (an affine transformation). Therefore, after the change of coordinate system defined by $\mathbf{p} = \mathbf{W}^T\mathbf{m}$, the line at infinity has not changed but the directions of the new u'- and v'-axes are orthogonal since the equation of the image of the absolute conic is $S(\mathbf{p}) = \mathbf{p}^T\mathbf{p}$. Indeed, this implies that $S(\mathbf{u}'_\infty, \mathbf{v}'_\infty) = 0$ where u'_∞ and v'_∞ are the points of projective coordinates (1, 0, 0) and (0, 1, 0) in the (u', v') coordinate system and, according to equation (16) this shows that the angle between the new u'- and v'-axes is equal to $\frac{\pi}{2}$.

We can now give a familiar interpretation of the equations (18). We consider the matrix $\mathbf{W}^T$ as defining a change of coordinate system from the affine plane (u, v) to the affine plane (u', v'). We know that the directions u' and v' are orthogonal. Let us consider an orthonormal system of coordinates centered at c with axes u' and v'. Let p be a point represented by the vector $[u', v', 1]^T$ in that coordinate system. The equation $\mathbf{p} = \mathbf{W}^T\mathbf{m}$ and the equations (18) can be written as follows

$$\begin{aligned} u' &= \frac{u-u_0}{\alpha_u} + \frac{v-v_0}{\alpha_v}\cos\theta \\ v' &= \frac{v-v_0}{\alpha_v}\sin\theta \end{aligned}$$

where u and v are the affine coordinates of the point m. This shows that the (u, v) (pixel) coordinate system is obtained from the (u', v') (normalized) coordinate system by translating the origin by $(-u_0, -v_0)$, rotating the v'-axis by $\theta - \frac{\pi}{2}$ and scaling the unit vectors in the u- and v-directions by $\frac{1}{\alpha_u}$ and $\frac{1}{\alpha_v}$, respectively (see figure 7). This is precisely the definition of the intrinsic parameters given for example in [22] from heuristic considerations.

As mentioned previously, this interpretation does not hold true anymore if the plane (u, v) has been projectively distorted. The points u_∞ and v_∞ are not at infinity anymore and the interpretation of θ does not make sense ("hidden" projective transformation problem). This could happen, for example, if the retinal plane were tilted with respect to the real optical axis of the optical system of the camera (misalignment). On the other hand, and this is very important, all angle measurements based on equation (16) are still valid because they do not make any assumption about the line at infinity of the retinal plane.

To summarize, the usual interpretation of the intrinsic parameters $\alpha_u, \alpha_v, \theta, u_0, v_0$ in terms of physical parameters attached to the camera is valid only if the original retinal plane has not been distorted by a projective transformation. But, even if this is the case, the camera can still be used to perform *euclidean* measurements through equation (16) which does not require the knowledge of the line at infinity in the retinal plane.

Figure 7 approximately here.

4.4.3 Two cameras and the absolute conic

As a final property of the absolute conic, let us consider its pair of images in the two retinal planes of a stereo rig and the two epipolar planes which are tangent to the absolute conic (they are complex conjuguate, intersecting along the real line $\langle C, C' \rangle$). These two planes intersect the two retinal planes along two pairs of corresponding epipolar lines, by definition, and these epipolar lines are tangent to the images of the absolute conic in the two retinal planes. This property will be used in section 7.1.2 to derive the Kruppa equations.

5 Recovering the projective stratum

In this section we show that a stereo rig for which the fundamental matrix has been estimated allows to recover the first stratum of the world, its projective structure. Even though this can be done algebraically [15] we will develop here a purely geometric approach. But first we give some indications about the way the fundamental matrix can be estimated from a pair of images.

5.1 Learning the fundamental matrix

The estimation of the fundamental matrix of a stereo rig is a problem which has recently received a lot of attention from a variety of people [2, 1, 5]. The basic idea is to use equation (13) for a number N of known pairs of corresponding pixels (m_i, m'_i). We obtain equations which are linear in the coordinates of matrix $\mathbf{F}$. More specifically, let us note $\mathbf{f}$ the 9-dimensional vector whose coordinates are the elements of $\mathbf{F}$. Each equation (13) can be written as

$$\mathbf{a}_i^T \mathbf{f} = 0,$$

and the whole set of equations can be written in matrix form

$$\mathbf{A}\mathbf{f} = \mathbf{0}$$

where $\mathbf{A}$ is an $N \times 9$ matrix. Let $\mathbf{a}_8$ and $\mathbf{a}_9$ be the last two column vectors of $\mathbf{A}$, $\mathbf{f} = [\mathbf{g}^T, f_8, f_9]^T$, and let us rewrite the previous equation as

$$\mathbf{B}\mathbf{g} = -f_8\mathbf{a}_8 - f_9\mathbf{a}_9$$

Assuming that the rank of the $N \times 7$ matrix $\mathbf{B}$ is seven, we can solve for the first seven components $\mathbf{g}$ of $\mathbf{f}$ in the usual way

$$\mathbf{g} = -f_8(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{a}_8 - f_9(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{a}_9$$

The solution depends upon two free parameters f_8 and f_9 which can be determined by using the constraint $det(\mathbf{F}) = 0$. We obtain a third-degree homogeneous equation in f_8 and f_9 and we can solve for their ratio. Since a third degree equation has at least one real root we are guaranteed to obtain at least one solution for $\mathbf{F}$. This solution is defined up to a scale factor and some normalization must be performed in order to make comparisons. One possibility is to normalize $\mathbf{f}$ such that its vector norm is equal to 1. If there are three real roots, we choose the one which minimizes the vector norm of $\mathbf{A}\mathbf{f}$, subject to the previous constraint. In fact we can do the same computation for any of the 36 choices of pairs of coordinates of $\mathbf{f}$ and choose, among the possibly 108 solutions, the one which minimizes the previous vector norm.

This approach has the problem that it does not minimize a meaningful criterion in terms of image measurements. Even though equation (13) can be normalized by imposing that the vector norms of $\mathbf{m}$ and $\mathbf{m}'$ are equal to 1, what we really would like to happen is that the *image distance* of m' (resp. of m) to the epipolar line of m (resp. of m') is small and this is *not* guaranteed by the previous approach. This has led people to minimize the sum over the pairs of corresponding pixels of the sum of the distance of m' to the epipolar line of m and the distance of m to the epipolar line of m'. The reader can easily verify that this criterion is not polynomial in the elements of $\mathbf{F}$ and that its minimization poses the usual problems of minimizing a criterion which is not a positive quadratic form in the unknowns. The best results have been obtained by initializing he nonlinear criterion with the result of the first method [5]. A stereo rig for which the fundamental matrix $\mathbf{F}$ is known is said to be weakly calibrated.

5.2 Recovering the projective structure of the world: the projective stratum

Let us choose five pair of point correspondences (a_i, a'_i), $i = 1, \cdots, 5$ in the two images. These correspondence may have been obtained, for example, in the process of estimating the fundamental matrix. We choose the five points A_i in the world as a projective basis. Note that these points are not known in the usual sense: the only thing we know is their projections in the two images are the pairs (a_i, a'_i). They must be such that no four of them are coplanar (section 3.2.1) but this property can be checked directly from the pair of images [15]. In order to show, for example, that the point A_4 is not in the plane defined by A_1, A_2, A_3, it is sufficient to show that the projective coordinates of a_4 in the projective basis (e, a_1, a_2, a_3) are different of those of a'_4 in the projective basis (e', a'_1, a'_2, a'_3). Given any futher point correspondence (m, m') it defines a 3-D point M. We will show that the ratios of its projective coordinates in the previous projective basis can be computed from the pair of images. In order to do this we will use the fact that each such ratio is the cross-ratio of four planes and use the **Line-Plane** construction described in section 4.2.

Suppose for example that we want to compute the ratio of the third to the fourth projective coordinates, in the previous projective basis, of M of images (m, m'). We have seen in section to be equal to the cross-ratio of the four planes (A_1, A_2, A_5), (A_1, A_2, M), (A_1, A_2, A_3) and (A_1, A_2, A_4). Let P and Q be the points of intersection of the line $\langle A_3, A_4 \rangle$ with the planes (A_1, A_2, A_5) and (A_1, A_2, M), respectively. Our cross-ratio is therefore equal to the cross-ratio of the four points (P, Q, A_3, A_4) which can be computed from either one of the two images after we construct the images (p, p') and (q, q') of P and Q which we can do using the **Line-Plane** construction of section 4.2.

These coordinates are invariant, by definition, under any collineation of the world. We have therefore computed a projective invariant representation of the world from a pair of weakly calibrated cameras.

6 Recovering the affine stratum

In this section, we show that a stereo rig for which the fundamental matrix and the collineation induced by the plane at infinity have been estimated allows to recover the second stratum of the world, its affine structure.

6.1 Estimating the plane at infinity

In some cases, some affine invariant information about the scene may be available. For example, we may know that two lines are parallel. Two parallel lines intersect in the plane at infinity and therefore the points of intersection of their images in the two retinal planes are the images of that point in Π_∞. Another example is if we know the midpoint of a segment. Let a_1 (resp.a'_1) and a_2 (resp. a'_2) be the images of the two endpoints and let a (resp. a') be the known images of the midpoint. What does it teach us about the plane at infinity? well, let us consider the point at infinity B of the line of support of our line segment. Since A is the midpoint of A_1A_2, the cross-ratio $\{A, B; A_1, A_2\}$ equals -1. Since the cross-ratio is preserved by perspective projection, The image b (resp. b') of B satisfies $\{a, b; a_1, a_2\} = -1$ (resp. $\{a', b'; a'_1, a'_2\} = -1$). In order to construct b (resp. b') we only have to construct the harmonic conjugate of a with respect to a_1 and a_2 (resp. the harmonic conjugate of a' with respect to a'_1 and a'_2) and this is a standard geometric construction that can be performed with a straight-edge only [21]. The correspondence (b, b') yields one point in the plane at infinity.

More generally, if we have three pairs of correspondences (a_i, a'_i), $i = 1, 2, 3$ such that a_1, a_2, a_3 (resp. a'_1, a'_2, a'_3) are aligned, then the corresponding 3-D points A_1, A_2, A_3 are aligned if and only if the two cross-ratios $\{e, a_1; a_2, a_3\}$ and $\{e', a'_1; a'_2, a'_3\}$ are equal. If we happen to know the ratio of lengths $\frac{A_1A_2}{A_1A_3}$, then it determines the vanishing point b (resp. b') of the two image lines and thus one point in the plane at infinity, since we must have $\{a_1, b; a_2, a_3\} = \{a'_1, b'; a'_2, a'_3\} = \frac{A_1A_2}{A_1A_3}$.

If no such information is available but if we can control the displacement of our stereo rig then we can exploit the fact that if we translate it without rotating it, straight lines remain parallel to themselves. More precisely, suppose we have a line L with images l and l' before the translation of the stereo rig. After the translation, the images of L are l_1 and l'_1 and because the rig has translated, the points a and a', intersections of l and l_1 in the first image and of l' and l'_1 in the second, are the images of a point A at infinity, i.e. the point at infinity of L. Note that in order to obtain this information we must have obtained the correspondence (l, l') by some other process and kept track of l (resp. l') while the rig was translating in order to obtain the correspondence (l_1, l'_1). A variant of this idea which has been implemented in the author's laboratory is to obtain point correspondences between the two images (this is needed to estimate the fundamental matrix) and then track them while the rig is translating [6]. Since two points define a line we are back to the line case.

6.2 Affine reconstruction

Suppose that we have identified the plane at infinity. As strange as this may sound, the plane at infinity has nothing special to it and, just as a regular plane, it induces a collineation between the two images. We know that this collineation is in general defined by four point correspondences but, if the fundamental matrix has been estimated, three point correspondences are sufficient. We have described in the previous section very simple ways of obtaining these correspondences by actively moving the cameras or by using some information about the scene.

We now choose four pairs of point correspondences (a_i, a'_i) in the two images. We choose the corresponding four points A_i in the world as an affine basis. More precisely, we choose A_1 as the origin of the affine frame and the three vectors $\mathbf{A}_1\mathbf{A}_i \equiv \mathbf{e}_{i-1}$, $i = 2, 3, 4$ as the basis vectors. This assumes that none of the four points A_i lies in the plane at infinity. This can be checked since the collineation H_∞ induced by Π_∞ between the retinal planes is known. We simply have to check that $H_\infty a_i$ is sufficiently different from a'_i for each i.

Given any further point correspondence (m, m'), not in the plane at infinity, we will show that the affine coordinates of the corresponding 3-D point M in the affine basis $(A_1, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ can be computed from the pair of images. We will do it in two different ways.

First we will simply adapt the method presented in section 5.2 to this case and second, we will present a somewhat more intuitive construction. Both methods are of course equivalent.

The affine basis $(A_1, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ can be considered as a projective basis $(A_1, A_{12\infty}, A_{13\infty}, A_{14\infty}, A_5)$ where the points $A_{1i\infty}$, $i = 2, 3, 4$ are the points at infinity of the lines $\langle A_1, A_i \rangle$ and A_5 is the point of coordinates (1, 1, 1) in the affine basis. Since the images of the three points $A_{1i\infty}$, $i = 2, 3, 4$ can be constructed using the procedure **Line-Plane** and if the images of A_5 can be constructed from the images, we can apply exactly the projective scheme described in section 5.2. Indeed, we know from section 3.3 that when the projective coordinates in $\mathcal{P}^3$ are chosen in such a way that the equation of the plane at infinity is $x_4 = 0$ the affine coordinates of a point in $\mathcal{P}^3 \div \Pi_\infty$ are the ratios of its first three projective coordinates to the fourth. This construction is shown in figure 8.

Figure 8 approximately here.

It remains to show how to construct (a_5, a_5'). According to figure 9, this can be done in three main steps, each of them being implementable in the images:

1. Construct P, intersection of the line going through A_2 and parallel to $\langle A_1, A_3 \rangle$ with the line going through A_3 and parallel to $\langle A_1, A_2 \rangle$.
2. Construct the line going through P and parallel to $\langle A_1, A_4 \rangle$.
3. Construct the line through A_4 parallel to $\langle A_1, P \rangle$.

These last two lines intersect in A_5.

Figure 9 approximately here.

The corresponding construction in the first image plane follows the same pattern and is shown in figure 10. In what follows, we denote by v_{pq} the vanishing point of the image line $\langle p, q \rangle$ where p and q can take any of the four values a, b, c and d. For example, v_{ab} is the vanishing point of the line $\langle a, b \rangle$.

1. Construct, using the procedure **Line-Plane**, the vanishing points $v_{a_1 a_2}$ and $v_{a_1 a_3}$. p is at the intersection of $\langle a_3, v_{a_1 a_2} \rangle$ and of $\langle a_2, v_{a_1 a_3} \rangle$.
2. Construct, using the procedure **Line-Plane**, the vanishing points $v_{a_1 a_4}$ of the line $\langle a_1, a_4 \rangle$.
3. Construct, using the procedure **Line-Plane**, the vanishing point r of the line $\langle a_1, p \rangle$.

The point a_5 is at the intersection of $\langle a_4, r \rangle$ and $\langle p, v_{a_1 a_4} \rangle$.

Figure 10 approximately here.

The second method may be more intuitive and can be found in [24]. According to figure 11, in order to compute the affine coordinates of M, we need to construct the images of two points: the point Q_4 on the line $\langle A_1, A_4 \rangle$ such that the line $\langle M, Q_4 \rangle$ is parallel to the plane (A_1, A_2, A_3), and the point Q, intersection of the line going through M and parallel to $\langle A_1, A_4 \rangle$ with the plane (A_1, A_2, A_3). From Q we then compute Q_2 (resp. Q_3), intersection of the line going through Q and parallel to $\langle A_1, A_3 \rangle$ (resp. parallel to $\langle A_1, A_2 \rangle$) with $\langle A_1, A_2 \rangle$ (resp. with $\langle A_1, A_3 \rangle$). The three affine coordinates of M are the ratios $\frac{\overline{A_1 Q_i}}{\overline{A_1 A_i}}$, $i = 2, 3, 4$. Introducing the points at infinity $A_{1i\infty}$, $i = 2, 3, 4$ of the four lines $\langle A_1, A_i \rangle$, these ratios are in fact equal to the cross-ratios $\{A_1, A_{1i\infty}; Q_i, A_i\}$ which are preserved by perspective projection and can be computed from the images.

Figure 11 approximately here.

The images of Q are readily obtained: $v_{a_1 a_4}$ (resp. $v_{a'_1 a'_4}$) can be constructed through the procedure **Line-Plane**. This pair, together with the pair (m, m'), define a 3-D line and, applying our procedure **Line-Plane** a second time we construct the images (q, q') of the intersection of that line with the plane (A_1, A_2, A_3). Once this construction has been completed, one notices that the line $\langle M, Q_4 \rangle$ is parallel to the line $\langle A_1, Q \rangle$. We thus construct, using again the procedure **Line-Plane**, the vanishing points $v_{a_4 q}$ and $v_{a'_4 q'}$ of the lines $\langle a_4, q \rangle$ and $\langle a'_4, q' \rangle$. q_4 (resp. q'_4) is then obtained as the intersection of the line $\langle v_{a_4 q}, m \rangle$ (resp. $\langle v_{a'_4 q'}, m' \rangle$) with the line $\langle a_1, a_4 \rangle$ (resp. $\langle a'_1, a'_4 \rangle$.

From $v_{a_1 a_2}$ (resp. $v_{a'_1 a'_2}$) $v_{a_1 a_3}$ (resp. $v_{a'_1 a'_3}$) we construct the points of intersection q_2 and q_3 (resp. q'_2 and q'_3) of the lines $\langle q, v_{a_1 a_3} \rangle$ and $\langle q, v_{a_1 a_2} \rangle$ (resp. of the lines $\langle q', v_{a'_1 a'_3} \rangle$ and $\langle q', v_{a'_1 a'_2} \rangle$ with $\langle a_1, a_2 \rangle$ and $\langle a_1, a_3 \rangle$). The affine coordinates (X, Y, Z) of M are then obtained in either one of the two images as the following cross-ratios

$$
\begin{aligned}
X &= \{a_1, v_{a_1 a_2}; q_2, a_2\} = \{a'_1, v_{a'_1 a'_2}; q'_2, a'_2\} \\
Y &= \{a_1, v_{a_1 a_3}; q_3, a_3\} = \{a'_1, v_{a'_1 a'_3}; q'_3, a'_3\} \\
Z &= \{a_1, v_{a_1 a_4}; q_4, a_4\} = \{a'_1, v_{a'_1 a'_4}; q'_4, a'_4\}
\end{aligned}
$$

These coordinates are invariant, by definition, under any collineation of the world. We have therefore computed an affine invariant representation of the the world from a pair of weakly calibrated cameras for which the collineation induced by the plane at infinity is known.

As a final remark to conclude this section, in many cases one may not be interested in computing these affine coordinates, only in computing affine invariant three-dimensional properties of the scene such as ratios of lengths, midpoints, in checking affine invariant three-dimensional properties of the scene such as parallelism of lines, planes, or even in performing affine invariant constructions such as drawing a line going through a given point and parallel to a given line, constructing the midpoint of a line segment, etc... All these operations can be performed without choosing coordinates just by using the fundamental matrix and the knowledge of the plane at infinity. If coordinates must be computed that can also be done directly from the images themselves and without explicitely *reconstructing* the points.

7 Recovering the euclidean stratum

In this section we want to push our ideas to their final stage and show that a stereo rig for which the fundamental matrix, the collineation induced by the plane at infinity and the two images of the absolute conic have been estimated allows to recover the third stratum of the world, its euclidean structure or, more precisely, its structure up to a similitude. It is in fact redundant to know the collineation at infinity and the two images of the absolute conic as shown later.

7.1 Estimating the image of the absolute conic

We now describe several ways of estimating the image of the absolute conic. First, in some cases, some similitude invariants of the scene may be available. For example, we may know the angle between two lines, or the ratio of the lengths of two non parallel segments. Each such bit of information yields a constraint on the image of the absolute conic, an idea that is used, at least in the case of angles, in [25].

7.1.1 A priori information about the scene

If we know the angle α between two lines in the world, according to the analysis of section 3.4.2, and to equation (16), this yields the following constraint on the coefficients of the equation of ω

$$S(\mathbf{m}, \mathbf{n})^2 = S(\mathbf{m})S(\mathbf{n})\cos^2\alpha \tag{19}$$

This equation is seen to be a quadratic constraint on the coefficients of the equation of ω.

If we have two images of the scene for which we know the plane at infinity (see section 6), then we can obtain the vanishing point v_{pq} of any line $\langle p, q\rangle$. Now if we know the ratio of the lengths of two non coplanar segments AB and CD, we can use equation (27) which will be derived in the section 7.3.2 to derive another constraint on the coefficients of the equation of ω. Let us call r the (known) ratio $\frac{AB}{CD}$ and define

$$D(\mathbf{v}_{pq}, \mathbf{v}_{st}) = S(\mathbf{v}_{pq}, \mathbf{v}_{st})^2 - S(\mathbf{v}_{pq})S(\mathbf{v}_{st})$$

Using equations (19) and (27) we obtain the following constraint on the coefficients of the equation of ω

$$\begin{array}{l} D(\mathbf{v}_{ac}, \mathbf{v}_{bc})D(\mathbf{v}_{cd}, \mathbf{v}_{bd})S(\mathbf{v}_{ab}) = \\ r^2 D(\mathbf{v}_{ac}, \mathbf{v}_{ab})D(\mathbf{v}_{bc}, \mathbf{v}_{bd})S(\mathbf{v}_{cd}) \end{array} \tag{20}$$

which is seen to be a polynomial of degree 5 in the coefficients of the equation of ω.

A similar constraint on the image of the absolute conic in the second image can be written. If that image has been obtained with the same camera without changing the internal parameters, then we obtain a second constraint on ω.

7.1.2 Moving the camera and using Kruppa's equations

If no a priori information about the scene is available, we can still estimate the image of the absolute conic by using motions of the cameras. Note that the camera motions do not have to be known and can be anything as long as they are not pure translations or pure rotations, as shown in section 7.2.4. This observation was made in [26] and turned into an algorithm and a working method in [2, 1].

We now show that each such motion yields two quadratic polynomial equations in the coefficients of the equation of the dual of the image of the absolute conic. In order to do this, we note that if we move a camera from position 1 to position 2 without changing its internal parameters, the image of the absolute conic remains the same, as was pointed out in sections 3.4.2 and 4.4. Also, as we noticed in section 4.4.3, the two tangents from the epipoles to this image correspond to each other in the epipolar homography. Expressing these two facts algebraically yields the two equations.

Let $\mathbf{B}$ be the matrix of the dual of the image of the absolute conic. Let m be a point in the retinal plane, e the epipole. The line $\langle e, m\rangle$ is represented by the cross-product $\mathbf{e}\wedge\mathbf{m}$ which we write in matrix form $[\mathbf{e}]_\times\mathbf{m}$ with $[\mathbf{e}]_\times$ being the antisymmetric matrix representing the cross-product with the vector $\mathbf{e}$. To say that this line is tangent to ω is equivalent to saying that the point represented by $[\mathbf{e}]_\times\mathbf{m}$ is on the dual conic ω^*. Hence we write the algebraic equation

$$([\mathbf{e}]_\times\mathbf{m})^T\mathbf{B}[\mathbf{e}]_\times\mathbf{m} = 0$$

or

$$\mathbf{m}^T[\mathbf{e}]_\times\mathbf{B}[\mathbf{e}]_\times\mathbf{m} = 0 \tag{21}$$

This quadratic equation in the coordinates of $\mathbf{m}$ is the equation of the two tangents from e to ω.

From the previous considerations, for each point m on either one of these two tangents, its epipolar line must also be tangent to the image of the absolute conic in the second image. But we have seen that the image of the absolute conic in the second image is identical to its image in the first. The same is of course true of the dual conics. Therefore, introducing the fundamental matrix $\mathbf{F}$, we can write that

$$\mathbf{m}^T \mathbf{F}^T \mathbf{B} \mathbf{F} \mathbf{m} = 0 \tag{22}$$

if and only if the point m is on one of the previous two tangents. This second quadratic equation in the coordinates of $\mathbf{m}$ therefore also defines the two tangents from e to ω and, thus, The two equations (21) and (22) are equivalent.

This yields a priori five quadratic equations in the coefficients of $\mathbf{B}$. But in fact, because equations (21) and (22) represent a pair of lines and not a general conic, only two of these five equations are independent since it is sufficient to look at the intersection of the tangents with another line not going through e. In [27, 28, 26, 2], the line was chosen to be the line at infinity but in principle any line not going through e will do.

The two equations are called the Kruppa equations in recognition of the work of this Austrian mathematician [29] who worked on a variant of a problem posed by Chasles [30]. For details about the implementation of these ideas and experimental results, see [1, 4].

7.2 From a pair of images of the absolute conic to the plane at infinity

We have now estimated the images ω and ω' of the absolute conic in the two retinal planes of our stereo rig. Since the absolute conic lies in the plane at infinity, from the two conics and the epipolar geometry defined by the fundamental matrix, we should be able to recover the plane at infinity.

7.2.1 From two images of the absolute conic to H_∞

As shown previously, this is equivalent to estimating the collineation that it induces between the two retinal planes and, for that matter, three point correspondences are sufficient. The question is of course how to obtain such correspondences. We may think of choosing a point m on ω and compute the intersection of its epipolar line with ω' to obtain a corresponding point m' on ω'. The problems with this approach is that m has complex coordinates and that its epipolar line (also a complex line) intersects ω' in general in two complex points. Thus we have an ambiguity. One way to go around this difficulty is to do a bit of geometry.

Let m be a point in the first retina. The optical ray $\langle C, m \rangle$ intersects the plane at infinity in a point which we denote by M_∞. How can we build the image m'_∞ of M_∞ in the second retinal plane? this point is on the epipolar line l'_m of m and is such that the angle between the line $\langle C, C' \rangle$ and the the optical ray $\langle C', m'_\infty \rangle$ is the same as the angle between $\langle C, C' \rangle$ and the optical ray $\langle C, m \rangle$. But this angle is known since we know ω and the epipoles: consider the two points a and b of intersection of the epipolar line $\langle e, m \rangle$ with ω, the angle is given by the cross-ratio $\{e, m\,;\, a, b\}$. Therefore, considering the two points of intersection a' and b' of l'_m with ω', m'_∞ can be built as the point of l'_m such that the cross-ratio $\{e', m'_\infty\,;\, a', b'\}$ is equal to $\{e, m\,;\, a, b\}$. These two cross-ratios being of course equal to the cross-ratio $\{E, M_\infty\,;\, A, B\}$ as shown in figure 12. The details of the computation can be found in appendix A.

Figure 12 approximately here.

The situation is of course symmetric between the two retinal planes and, given a point m' in the second retinal plane, we can similarly build the image m_∞ of the intersection M'_∞ of the optical ray $\langle C', m' \rangle$ with the plane at infinity. We can thus build an arbitrary large number of pairs of point correspondences (m, m'_∞) or (m_∞, m') corresponding to points in the plane at infinity. From these pairs, the collineation H_∞ can be estimated.

7.2.2 A three-dimensional euclidean interpretation of H_∞

Let us show that after an affine change of coordinates in the two retinal planes such that the equation of the absolute conic in the first image (resp. the second) is $\mathbf{p}^T\mathbf{p} = 0$ (resp. $\mathbf{p}'^T\mathbf{p}' = 0$, its matrix $\mathbf{H}_\infty$ is proportional to a rotation matrix. Indeed, let us suppose that we change coordinate systems in the two retinal planes and define $\mathbf{p} = \mathbf{W}^T\mathbf{m}$ and $\mathbf{p}' = \mathbf{W}'^T\mathbf{m}'$ where $\mathbf{W}$ and $\mathbf{W}'$ are defined from the equations of ω and ω' as explained in section 4.4.2. The equations of the images ω and ω' of the absolute conic are $S(\mathbf{p}) = \mathbf{p}^T\mathbf{p} = 0$ and $S'(\mathbf{p}') = \mathbf{p}'^T\mathbf{p}' = 0$. Let p be a point of ω and p' its image under H_∞. p' belongs to ω' and therefore $S'(\mathbf{p}') = 0$. But this is also equal to $\mathbf{p}^T\mathbf{H}_\infty^T\mathbf{H}_\infty\mathbf{p}$ and must equal 0 for all points p of ω, i.e. $\mathbf{H}_\infty^T\mathbf{H}_\infty$ is proportional to the identity matrix $\mathbf{I}$. This shows that the matrix $\mathbf{H}_\infty$ is proportional to a rotation matrix $\mathbf{R}^T$. This matrix has a very intuitive interpretation. If we consider an orthonormal system of coordinates centered at the optical center C_1 (resp. C_2) with vectors parallel to the optical axis $\langle C_1, c_1 \rangle$ (resp. $\langle C_2, c_2 \rangle$ and the u'_1- and v'_1- orthogonal directions (resp. the u'_2- and v'_2- orthogonal directions) the matrix $\mathbf{R}$ is the one transforming the directions of the axes of the first coordinate system into those of the second, see figure 13.

If we express H_∞ in the pixel coordinate systems:

$$\mathbf{H}_\infty = \mathbf{W}'^{-T}\mathbf{R}\mathbf{W}^T$$

which shows that in the case where the two cameras are identical, i.e. $\mathbf{W} = \mathbf{W}'$, H_∞ is such that

$$\mathbf{R} = \mathbf{W}^T\mathbf{H}_\infty\mathbf{W}^{-T}.$$

Hence, using the fact that $\mathbf{R}$ is orthogonal, and introducing the matrix $\mathbf{A}$, we write:

$$\mathbf{A} = \mathbf{H}_\infty^T\mathbf{A}\mathbf{H}_\infty$$

This shows that if the collineation H_∞ has been estimated by some means, perhaps by matching points between the two views, then this matrix equation can be used to solve for the coefficients of $\mathbf{A}$ and then, by the Cholesky decomposition, for those of $\mathbf{W}$. This idea has been proposed and put into a working algorithm by Hartley [31].

Figure 13 approximately here.

The idea that was used to construct H_∞ from ω and ω' can be used to compute directly the vanishing points of a pair (l, l') of image lines without using the procedure **Plane-Line**. Here is how it works. Suppose that l is defined by the two points (m, p) (resp. l' is defined by the two points (m', p')). Note that we do not require that either m and m' or p and p' be corresponding points. We then build m'_∞ and p'_∞ in the second retinal plane, images of the points at infinity of the optical rays $\langle C, m \rangle$ and $\langle C, p \rangle$, and m_∞ and p_∞ in the first retinal plane, images of the points at infinity of the optical rays $\langle C', m' \rangle$ and $\langle C', p' \rangle$. The vanishing point q of l (resp. q' of l') is then obtained as the intersection of l and $\langle m_\infty, p_\infty \rangle$ (resp. of l' and $\langle m'_\infty, p'_\infty \rangle$). We call this procedure the **Vanishing-Point** procedure.

7.2.3 How H_∞ constrains ω and ω'

We show in the next section how to use both the homography of the plane at infinity and the images of the absolute conic to compute ratios of distances, an invariant for the group of similitudes. But before going into this, we study how the knowledge of the collineation H_∞ of the plane at infinity constrains ω and ω'. Since ω and ω' are the images of Ω in the two retinas, their duals ω^* and ω'^* are the images of Ω^*. As H_∞ is the collineation from the first retinal plane to the second induced by π_∞, $\mathbf{H}_\infty^{-T}$ represents the collineation from the dual of the first retinal plane to the dual of the second. In other words, lines of $\mathcal{R}$ are transformed into lines of $\mathcal{R}'$ by the collineation $\mathbf{H}_\infty^{-T}$. Indeed, let l' be a line of $\mathcal{R}'$, represented by $\mathbf{l}'$. Its equation can be written

$$\mathbf{l}'^T \mathbf{m}' = 0 \tag{23}$$

But for all points of $\mathcal{R}'$ there exists a unique point m of $\mathcal{R}$ such that

$$\mathbf{m}' = \mathbf{H}_\infty \mathbf{m}$$

Replacing $\mathbf{m}'$ by its value in (23), we obtain

$$(\mathbf{H}_\infty^T \mathbf{l}')^T \mathbf{m} = 0$$

and therefore, l' is the image of the line l of $\mathcal{R}$ represented by $\mathbf{H}_\infty^{-T}\mathbf{l}$.

Let now l and l' be corresponding lines for the collineation H_∞. If l belongs to ω^*, we have $\mathbf{l}^T\mathbf{B}\mathbf{l} = 0$. But l' must then belong to ω'^* and satisfy $\mathbf{l}'^T\mathbf{B}'\mathbf{l}' = 0$. This implies the following relation between $\mathbf{B}$ and $\mathbf{B}'$

$$\mathbf{H}_\infty \mathbf{B} \mathbf{H}_\infty^T = \mathbf{B}' \tag{24}$$

and imposes six homogeneous linear constraints on the coefficients of $\mathbf{B}$ and $\mathbf{B}'$.

If the epipolar geometry is also known, a reasoning similar to that of the previous section shows that

$$\mathbf{F}^T \mathbf{B} \mathbf{F} = \mathbf{B}' \tag{25}$$

which yields another set of six homogeneous linear equations which according to (14) are not independent of (24).

7.2.4 The Longuet-Higgins equation, pure camera translation

Let us finish this section by giving an interpretation of the fundamental matrix $\mathbf{F}$ when each retinal plane is referred to its normalized coordinates. We know that in this case we can choose $\mathbf{H}_\infty = \mathbf{R}^T$. We then write equation (14)

$$\mathbf{R}\mathbf{F} + \mathbf{F}^T\mathbf{R}^T = 0$$

which says that the matrix $\mathbf{F}^T\mathbf{R}^T$ is antisymmetric. Let us write it $[\mathbf{t}]_\times$, where $\mathbf{t}$ is a vector. We thus have:

$$\mathbf{F}^T = [\mathbf{t}]_\times \mathbf{R}$$

which shows that the transpose of the fundamental matrix is nothing but the essential matrix $\mathbf{E}$ defined by Longuet-Higgins in his 1981 paper [32]. The properties of this matrix have been subsequentally studied by several authors [33, 34, 22]. The vector $\mathbf{t}$ which appears in the definition is parallel to the line $\langle C_1, C_2 \rangle$. The equation (13) becomes the well known Longuet-Higgins equation

$$\mathbf{m}^T \mathbf{E} \mathbf{m}' = 0$$

. This allows us to interpret the vector $\mathbf{t}$ just introduced as the direction of the translation between the two optical centers, i.e. the direction of the line $\langle C_1, C_2 \rangle$ in figure 13.

Returning to equations (21) and (22), we see that if the displacement between the two camera positions is a pure translation, i.e. $\mathbf{R} = \mathbf{I}$, we have $\mathbf{F}^T = [\mathbf{t}]_\times = [\mathbf{e}]_\times$. The equations (21) and (22) are identical and there are no Kruppa equations in that case.

7.3 From a pair of images of the absolute conic, the plane at infinity to similitude invariants

7.3.1 Angles

We know from a previous section that the knowledge of the image of the absolute conic in one camera allows to compute angles between optical rays. Knowing the images of the absolute conic in two cameras allows to compute angles between any lines. For example, given three points A, B, C in the world with images a, b, c (resp. a', b', c') in the first (resp. second) retinal plane, how do we compute the angle between the lines $\langle A, B \rangle$ and $\langle A, C \rangle$? Let P (resp. Q) be the points at infinity of $\langle A, B \rangle$ (resp. $\langle A, C \rangle$). The angle θ is obtained by a straightforward application of Laguerre's formula. In order to be able to compute the cross-ratio that appears into it, we must be able to compute the images of P and Q, i.e. the vanishing points of the image lines $\langle a, b \rangle$ (resp. $\langle a', b' \rangle$) and $\langle a, c \rangle$ (resp. $\langle a', c' \rangle$). This is possible since, according to the previous section, we know the plane at infinity. We can thus call upon our **Line-Plane** or **Vanishing-Point** procedures. More precisely, and according to equation (16), we have

$$\cos(\langle A, B \rangle, \langle A, C \rangle) = -\frac{S(\mathbf{p}, \mathbf{q})}{\sqrt{S(\mathbf{p})S(\mathbf{q})}} = -\frac{S'(\mathbf{p}', \mathbf{q}')}{\sqrt{S'(\mathbf{p}')S'(\mathbf{q}')}} \tag{26}$$

and of course the sine is obtained as $\sqrt{1 - \cos^2(\langle A, B \rangle, \langle A, C \rangle)}$.

More generally, the angle between two general lines $\langle A, B \rangle$ and $\langle C, D \rangle$, not necessarily coplanar is obtained by considering the point at infinity P of $\langle A, B \rangle$ and the point at infinity Q of $\langle C, D \rangle$ (i.e. the directions of the two lines) and computing the cross-ratio of P and Q and the points of intersection of the line $\langle P, Q \rangle$ with the absolute conic, this computation being of course performed in the images, i.e. using equation (26).

7.3.2 Ratios of lengths

Let us now describe how we can compute the other type of similitude invariants, the ratios of lengths. Using the fact that angles can be computed, we show how to use them to compute ratios of lengths. Let us consider four points A, B, C and D and suppose we want to compute the ratio $\frac{AB}{CD}$. Considering the two triangles ABC and BCD, as shown in figure 14, we can write, for the first triangle

$$\frac{AB}{\sin\gamma} = \frac{BC}{\sin\alpha}$$

and, for the second

$$\frac{BC}{\sin\delta} = \frac{CD}{\sin\beta}$$

from which we obtain the ratio $\frac{AB}{CD}$ as a function of the four angles α, β, γ and δ

$$\frac{AB}{CD} = \frac{\sin\gamma \cdot \sin\delta}{\sin\alpha \cdot \sin\beta} \tag{27}$$

Figure 14 approximately here.

Figure 15 shows the computation in the first image plane. Using the procedure **Line-Plane** or **Vanishing-Point**, we construct the four vanishing points v_{ab}, v_{ac}, v_{bc}, v_{bd}, and v_{cd} of the image lines $\langle a, b\rangle$, $\langle a, c\rangle$, $\langle b, c\rangle$, $\langle b, d\rangle$ and $\langle c, d\rangle$. The sines which appear in equation (27) are then obtained through equation (16). More specifically, we obtain the neat formula for the ratio $\frac{AB}{CD}$ computed in the first image in which appears only the equation of ω:

$$\frac{AB}{CD} = \sqrt{\frac{S(\mathbf{v}_{ab})}{S(\mathbf{v}_{cd})} \cdot \frac{S(\mathbf{v}_{ac})S(\mathbf{v}_{bc}) - S(\mathbf{v}_{ac},\, \mathbf{v}_{bc})^2}{S(\mathbf{v}_{ab})S(\mathbf{v}_{ac}) - S(\mathbf{v}_{ab},\, \mathbf{v}_{ac})^2} \cdot \frac{S(\mathbf{v}_{bd})S(\mathbf{v}_{cd}) - S(\mathbf{v}_{bd},\, \mathbf{v}_{cd})^2}{S(\mathbf{v}_{bc})S(\mathbf{v}_{bd}) - S(\mathbf{v}_{bc},\, \mathbf{v}_{bd})^2}} \tag{28}$$

The geometry is shown is figure 15.

Figure 15 approximately here.

8 Conclusion

Table 1 approximately here.

We have reached the end of a fairly long journey in which we have seen that when we look at the physical world with a set of two cameras, there appears a natural hierarchical set of geometric descriptions of this world which involve a trilogy of groups of transformations which leave these descriptions invariant, i.e. among which we cannot discriminate. These three groups are the projective, the affine, and the similitude groups. For each of these descriptions, we have indicated a corresponding geometric property of the set of two cameras which, once known, allows to recover the related description from pairs of corresponding image features. We have also indicated how these properties of the pair of cameras could be estimated from images of the physical world. This is interesting in itself as well as in connection with the psychophysical work of Droulez and Cornilleau [35] in which they showed that humans with normal uncorrected vision and wearing distorting lenses could recover, after some time, the ability to perform correct metric judgments. Another important aspect of our work is that it clearly shows that, for each subgroup of interest, all three-dimensional invariants of the scene can be estimated directly from the images *without* performing an explicit 3-D reconstruction of the scene. This may buy stability in applications in particular because it avoids the problem, mentioned in section 4.4.2, of the "hidden" projective transformation. But this has to be checked experimentally. Table 1 summarizes the relations between the three strata of the physical world, the geometric properties of the stereo rig, and some of the three-dimensional quantities that can be recovered directly from the images.

A Computing m'_∞

From section 4 we know that the angle θ between the optical ray $\langle C, m\rangle$ and the baseline $\langle C, C'\rangle$ is given by:

$$\cos\theta = -\frac{S(e,m)}{\sqrt{S(e)S(m)}}$$

Therefore we want to find m'_∞ on the epipolar line l'_m such that

$$\frac{S'(e', m'_\infty)}{\sqrt{S'(e')S'(m'_\infty)}}$$

is equal to $-\cos\theta$. Let us choose any point m' on l'_m and write

$$\mathbf{m}'_\infty = \mathbf{m}' + \alpha\mathbf{e}'$$

The problem is to determine α. We write

$$\begin{aligned} S'(e', m'_\infty) &= S'(e', m') + \alpha S'(e') \\ S'(m'_\infty) &= S'(m') + 2\alpha S'(e', m') + \alpha^2 S'(e') \end{aligned}$$

We then express the fact that $\cos^2\theta = \frac{S'^2(e', m'_\infty)}{S'(e')S'(m'_\infty)}$. We obtain a quadratic equation in the unknown α:

$$\alpha^2 S'^2(e')\sin^2\theta + 2\alpha S'(e')S'(e', m')\sin^2\theta + S'^2(e', m') - S'(e')S'(m')\cos^2\theta = 0$$

In order to compute its roots, we compute the discriminant

$$\Delta' = [S'(e')S'(m') - S'^2(e', m')]S'^2(e')\sin^2\theta\cos^2\theta$$

Since, according to equation (15), the quantity $S'(e')S'(m') - S'^2(e', m')$ is positive, our equation has two real roots

$$\alpha = -\frac{S'(e', m')}{S'(e')} \pm \frac{|\cos\theta|}{S'(e')\sin\theta}\sqrt{S'(e')S'(m') - S'^2(e', m')}$$

The sign of the cosine of the angle between $\langle C', m'_\infty\rangle$ and $\langle C, C'\rangle$ is given by $S'(e', m'_\infty)$ which is equal to

$$S'(e', m'_\infty) = S'(e', m') + \alpha S'(e') = \pm\frac{|\cos\theta|}{\sin\theta}\sqrt{S'(e')S'(m') - S'^2(e', m')}$$

therefore only one of the two roots provides the correct sign and the solution is unique, as expected.

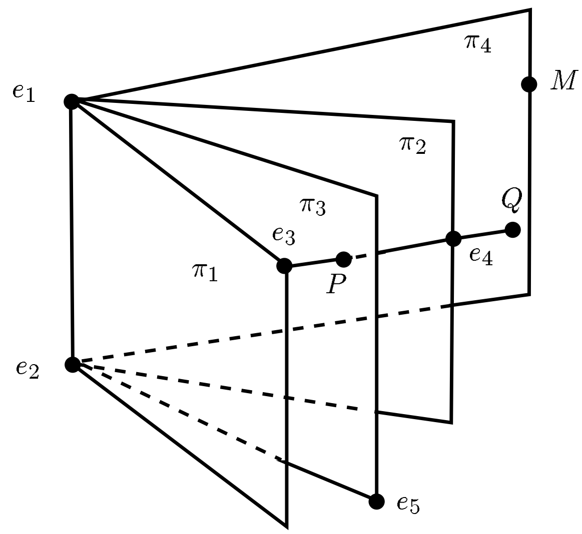

Figure 1: The ratio of the third to the last projective coordinates of the point M in the projective basis $(e_1, e_2, e_3, e_4, e_5)$ is equal to the cross-ratio of the four planes (e_1, e_2, e_3), (e_1, e_2, e_4), (e_1, e_2, e_5), and (e_1, e_2, M).

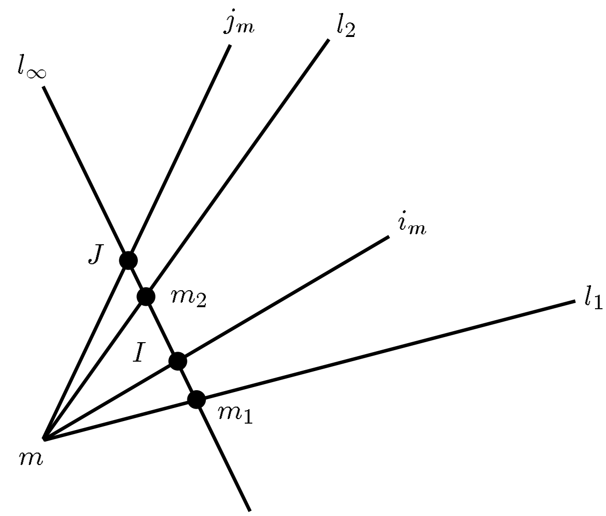

Figure 2: The angle α between l_1 and l_2 is given by Laguerre formula: $\alpha = \frac{1}{2i} \log(\{l_1, l_2; i_m, j_m\})$.

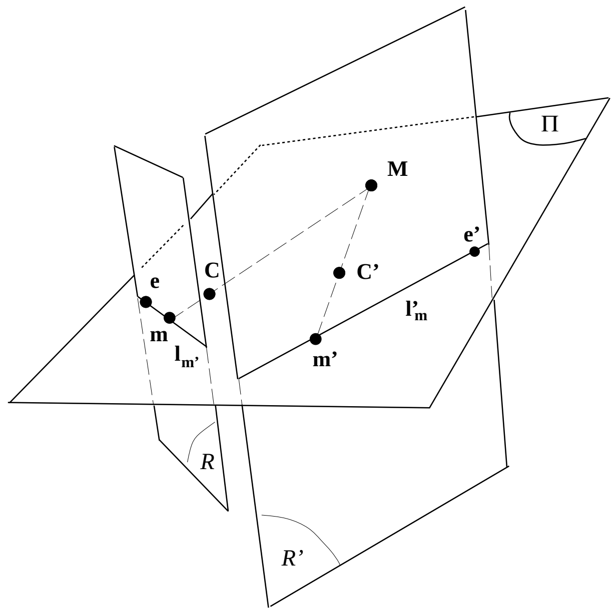

Figure 3: The epipolar geometry.

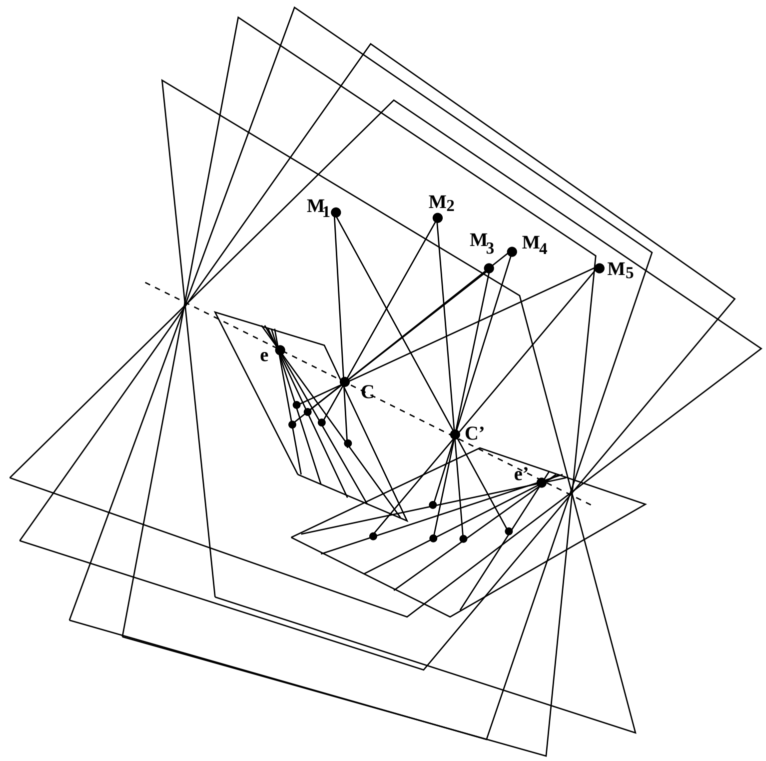

Figure 4: The epipolar pencils.

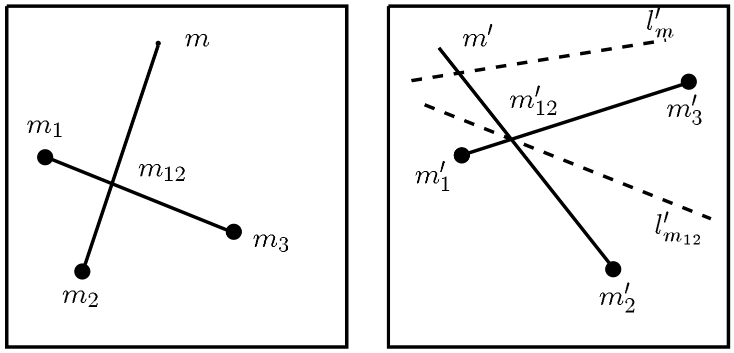

Figure 5: Construction of the image m' of the point m under the collineation induced by the plane defined by the three point correspondences $(m_i, m'_i),\ i = 1, 2, 3$ and the epipolar geometry.

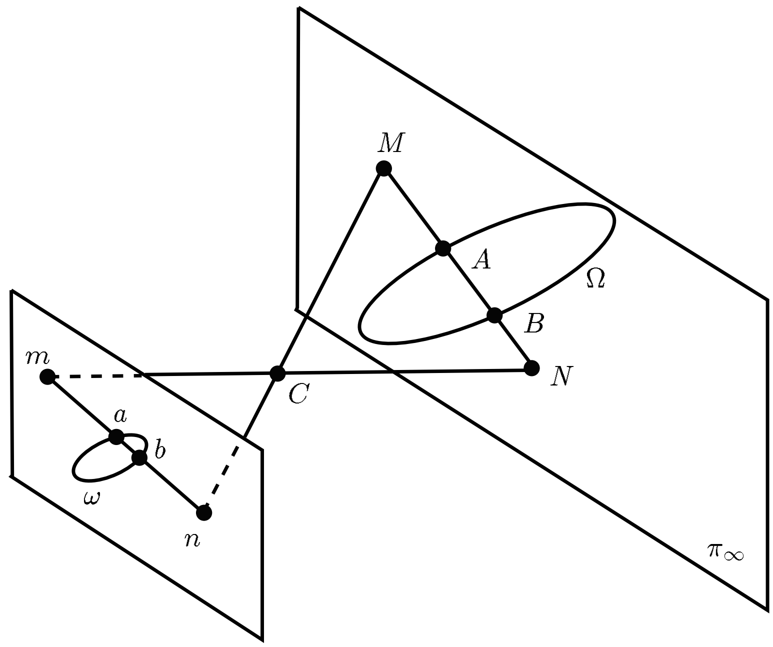

Figure 6: How to compute the angle between the optical rays $\langle C, m\rangle$ and $\langle C, n\rangle$ using the image of the absolute conic.

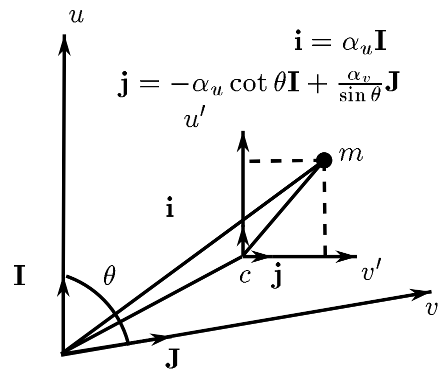

Figure 7: From pixel-coordinates (u, v) to normalized coordinates (u', v').

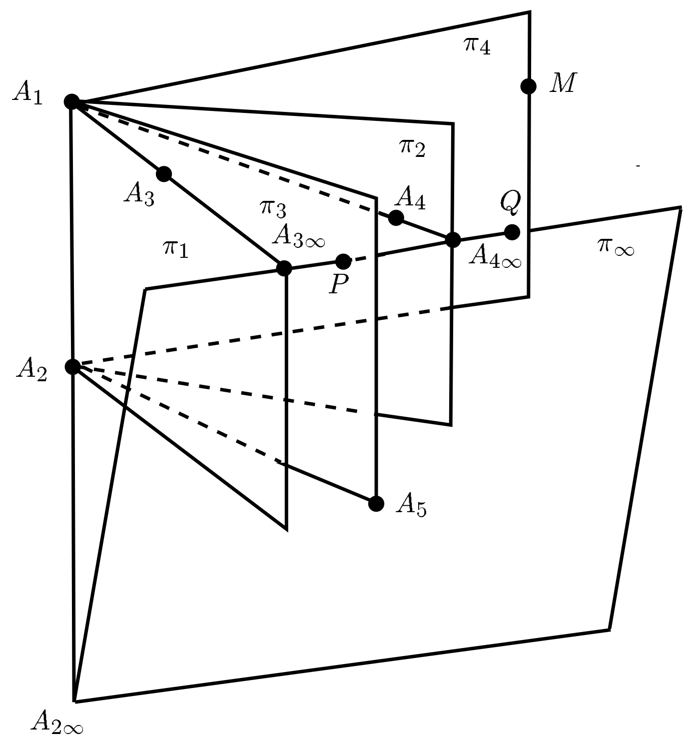

Figure 8: The third affine coordinates of M in the affine basis of origin A_1 and basis vectors $\mathbf{A}_1\mathbf{A}_i$, $i = 2, 3, 4$ is equal to the cross-ratio of the four planes (A_1, A_2, A_3), (A_1, A_2, A_4), (A_1, A_2, A_5) and (A_1, A_2, M).

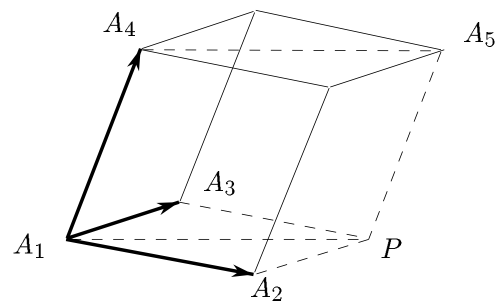

Figure 9: Three dimensional construction of A_5.

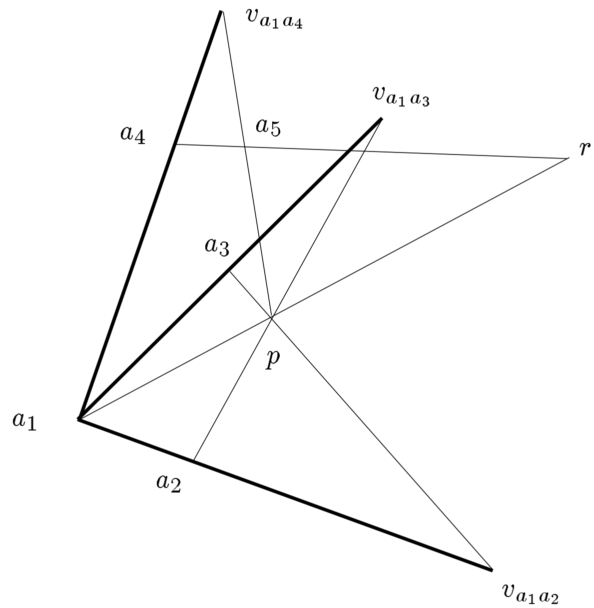

Figure 10: Construction of a_5, image of A_5 in the first retina.

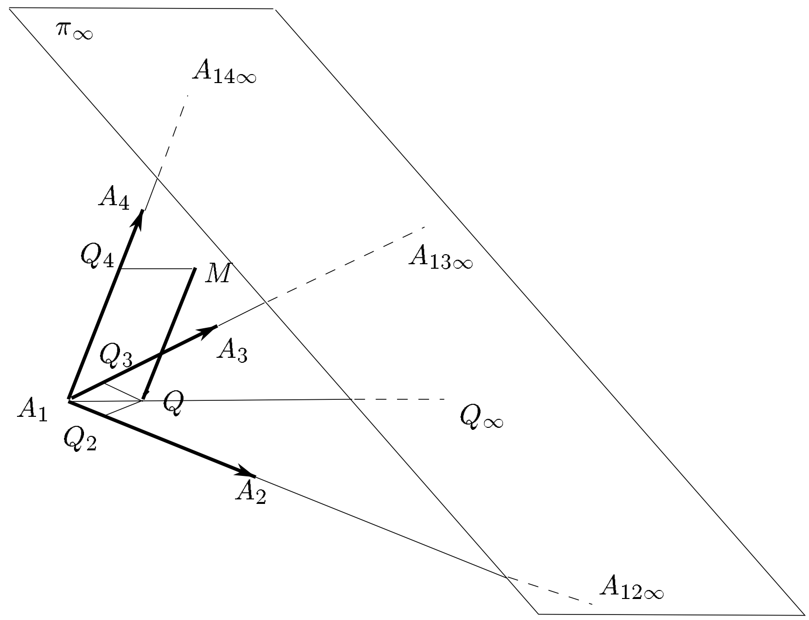

Figure 11: Computation of the affine coordinates of M.

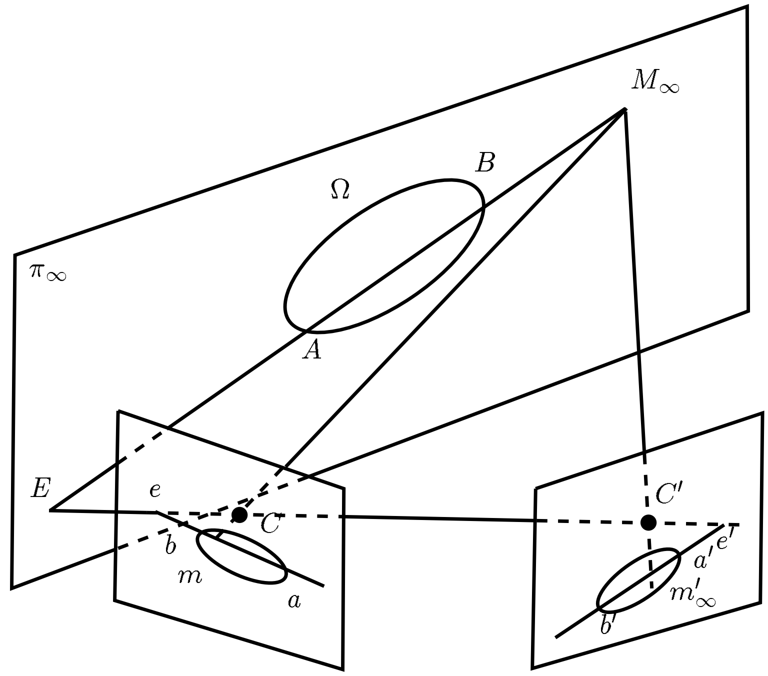

Figure 12: Construction of the pair (m, m'_∞) corresponding to the point M_∞ of the plane at infinity π_∞.

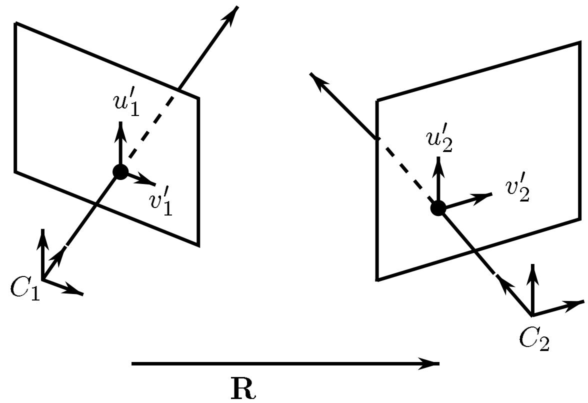

Figure 13: The collineation induced by the plane at infinity π_∞ is proportional to the matrix $\mathbf{R}$ of the 3-D rotation.

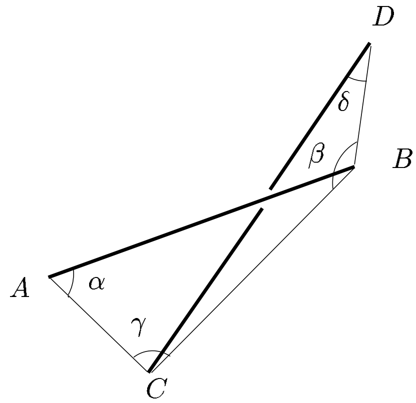

Figure 14: Computing the ratio of the lengths of the two segments AB and CD from the angles α, β, γ and δ.

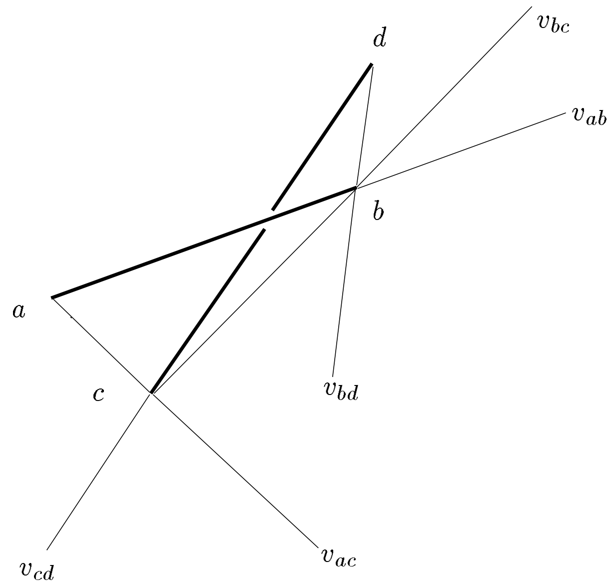

Figure 15: Computing the ratio of the lengths of the two segments AB and CD from the vanishing points v_{ab}, v_{ac}, v_{bc}, v_{bd}, and v_{cd} of the image lines $\langle a, b\rangle$, $\langle a, c\rangle$, $\langle b, c\rangle$, $\langle b, d\rangle$ and $\langle c, d\rangle$.

Geometric Structure	Stereo Rig	Invariant Measures
Projective	Fundamental Matrix	Cross-ratios
Affine	Collineation of the plane at infinity	Ratios of lengths of parallel segments
Similitude	Images of the absolute conic	Angles, ratios of lengths of non-parallel segments

Table 1: Relations between the three strata and the geometric properties of the stereo rig.

References

[1] Olivier D. Faugeras, Tuan Luong, and Steven Maybank. Camera self-calibration: theory and experiments. In Giulio Sandini, editor, *Proceedings of the 2nd European Conference on Computer Vision*, pages 321–334. Springer-Verlag, Lecture Notes in Computer Science 588, 1992.

[2] Quang-Tuan Luong. *Matrice Fondamentale et Calibration Visuelle sur l'Environnement-Vers une plus grande autonomie des systèmes robotiques.* PhD thesis, Université de Paris-Sud, Centre d'Orsay, December 1992.

[3] Olivier Faugeras, Bernard Hotz, Hervé Mathieu, Thierry Viéville, Zhengyou Zhang, Pascal Fua, Eric Théron, Laurent Moll, Gérard Berry, Jean Vuillemin, Patrice Bertin, and Catherine Proy. Real time correlation based stereo: algorithm implementations and applications. Research Report 2013, INRIA Sophia-Antipolis, France, 1993. Submitted to *The International Journal of Computer Vision*.

[4] Quang-Tuan Luong and Olivier Faugeras. The Fundamental matrix: theory, algorithms, and stability analysis. *The International Journal of Computer Vision*, 1994. To appear.

[5] R. Deriche, Z. Zhang, Q.-T Luong, and O. Faugeras. Robust Recovery of the Epipolar Geometry for an Uncalibrated Stereo Rig. In Jan-Olof Eklundh, editor, *Proceedings of the 3rd European Conference on Computer Vision*, pages 567–576, Vol.I, Stockholm,Sweden, 1994. Springer-Verlag, Lecture Notes in Computer Science 800-801.

[6] C. Zeller and Olivier Faugeras. Applications of non-metric vision to some visual guided tasks. In *Proc. International Conference on Pattern Recognition*, Israel, 1994. To appear.

[7] Jan J. Koenderink and Andrea J. van Doorn. Affine Structure from Motion. *Journal of the Optical Society of America*, A8:377–385, 1991.

[8] Gunnar Sparr. An algebraic-analytic method for reconstruction from image correspondences. In *Proceedings 7th Scandinavian Conference on Image Analysis*, pages 274–281, 1991.

[9] Gunnar Sparr. Projective invariants for affine shapes of points configurations. In J.L. Mundy and A. Zisserman, editors, *Proceedings of DARPA-ESPRIT Workshop on Applications of Invariance in Computer Vision*, pages 151–169, 1991.

[10] A. Sparr and G. Sparr. On a theorem of M. Chasles. Technical Report LUFTD2/TFMA-7001-SE, Lund University, Dept. of Mathematics, 1993.

[11] Gunnar Sparr. A common framework for kinetic depth, reconstruction and motion for deformable objects. In Jan-Olof Eklundh, editor, *Proceedings of the 3rd European Conference on Computer Vision*, volume 801 of *Lecture Notes on Computer Science*, pages 471–482. Springer-Verlag, 1994.

[12] Gunnar Sparr. Applications of a theorem of Chasles to computer vision. Technical Report LUFTD2/TFMA-7002-SE, Lund University, Dept. of Mathematics, 1994.

[13] Roger Mohr, Luce Morin, and Enrico Grosso. Relative positioning with uncalibrated cameras. In J.L. Mundy and A. Zisserman, editors, *Proceedings of DARPA-ESPRIT Workshop on Applications of Invariance in Computer Vision*, Artificial Intelligence Series, chapter 22, pages 440–460. MIT Press, 1992.

[14] Richard Hartley, Rajiv Gupta, and Tom Chang. Stereo from Uncalibrated Cameras. In *Proceedings of CVPR92, Champaign, Illinois*, pages 761–764, June 1992.

[15] Olivier D. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig. In Giulio Sandini, editor, *Proceedings of the 2nd European Conference on Computer Vision*, pages 563–578. Springer-Verlag, Lecture Notes in Computer Science 588, May 1992.

[16] Amnon Shashua. Projective Structure from two Uncalibrated Images: Structure from Motions and Recognition. Technical Report A.I. Memo No. 1363, MIT, September 1992.

[17] Amnon Shashua. Projective Depth: A Geometric Invariant for 3D Reconstruction From Two Perspective\Orthographic Views and For Visual Recognition. In *Proc. Fourth International Conference on Computer Vision*, pages 583–590, 1993.

[18] Amnon Shashua. Algebraic Functions for Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1994. To appear.

[19] Olivier Faugeras and Luc Robert. What can two images tell us about a third one? *The International Journal of Computer Vision*, 1994. To appear. Also INRIA Technical report 2018.

[20] Q.-T. Luong and T. Viéville. Canonic representations for the geometries of multiple projective views. In Jan-Olof Eklundh, editor, *Proceedings of the 3rd European Conference on Computer Vision*, pages 589–599, Vol. I. Springer-Verlag, Lecture Notes in Computer Science 800-801, 1994.

[21] J.G. Semple and G.T. Kneebone. *Algebraic Projective Geometry*. Oxford: Clarendon Press, 1952. Reprinted 1979.

[22] Olivier D. Faugeras. *Three-Dimensional Computer Vision: a Geometric Viewpoint*. MIT Press, 1993.

[23] Gene H. Golub and Charles F. Van Loan. *Matrix computations*. The John Hopkins University Press, Baltimore, Maryland, 1983.

[24] T. Moons, L. Van Gool, M. Van Dienst, and E. Pauwels. Affine structure from perspective image pairs under relative translations between object and camera. Technical Report KUL/ESAT/MI2/9306, Katholieke Universiteit Leuven, 1993.

[25] Boubakeur Boufama, Roger Mohr, and Françoise Veillon. Euclidean Constraints for Uncalibrated Reconstruction. Technical Report RT96 IMAG 17 LIFIA, LIFIA, INSTITUT IMAG, March 1993.

[26] S.J. Maybank and O.D. Faugeras. A Theory of Self-Calibration of a Moving Camera. *The International Journal of Computer Vision*, 8(2):123–152, August 1992.

[27] Olivier D. Faugeras and Steven Maybank. Motion from point matches: multiplicity of solutions. *The International Journal of Computer Vision*, 4(3):225–246, June 1990. also INRIA Tech. Report 1157.

[28] Olivier D. Faugeras and Steven Maybank. Mouvement à partir de points : nombre de solutions. *Comptes rendus de l'Académie des Sciences de Paris*, pages 177–183, 1991. t. 312, Série II.

[29] E. Kruppa. Zur Ermittlung eines Objektes aus zwei Perspektiven mit innerer Orientierung. *Sitz.-Ber. Akad. Wiss., Wien, math. naturw. Kl., Abt. IIa.*, 122:1939–1948, 1913.

[30] M. Chasles. Question No. 296. *Nouv. Ann. Math.*, 14:50, 1855.

[31] Richard Hartley. Self-Calibration from Multiple Views with a Rotating Camera. In Jan-Olof Eklundh, editor, *3rd European Conference on Computer Vision*, volume 800 of *Lecture Notes on Computer Science*, pages 471–478, Stockholm, May 1994. Springer-Verlag.

[32] H.C. Longuet-Higgins. A Computer Algorithm for Reconstructing a Scene from Two Projections. *Nature*, 293:133–135, 1981.

[33] Thomas S. Huang and Olivier D. Faugeras. Some Properties of the E Matrix in Two-View Motion Estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(12):1310–1312, December 1989.

[34] S.J. Maybank. Properties of Essential Matrices. *International Journal of Imaging Systems and technology*, 2:380–384, 1990.

[35] Jacques Droulez and Valérie Cornilleau. Adaptive changes in perceptual responses and visuomanual coordination during exposure to visual metrical distortion. *Vision Research*, 26(11):1783–1792, 1986.

List of Figures

1 The ratio of the third to the last projective coordinates of the point M in the projective basis $(e_1, e_2, e_3, e_4, e_5)$ is equal to the cross-ratio of the four planes (e_1, e_2, e_3), (e_1, e_2, e_4), (e_1, e_2, e_5), and (e_1, e_2, M). 32
2 The angle α between l_1 and l_2 is given by Laguerre formula: $\alpha = \frac{1}{2i}\log(\{l_1, l_2; i_m, j_m\})$. 33
3 The epipolar geometry. 34
4 The epipolar pencils. 35
5 Construction of the image m' of the point m under the collineation induced by the plane defined by the three point correspondences (m_i, m'_i), $i = 1, 2, 3$ and the epipolar geometry. 36
6 How to compute the angle between the optical rays $\langle C, m\rangle$ and $\langle C, n\rangle$ using the image of the absolute conic. 37
7 From pixel-coordinates (u, v) to normalized coordinates (u', v'). 38
8 The third affine coordinates of M in the affine basis of origin A_1 and basis vectors $\mathbf{A}_1\mathbf{A}_i$, $i = 2, 3, 4$ is equal to the cross-ratio of the four planes (A_1, A_2, A_3), (A_1, A_2, A_4), (A_1, A_2, A_5) and (A_1, A_2, M). 39
9 Three dimensional construction of A_5. 40
10 Construction of a_5, image of A_5 in the first retina. 41
11 Computation of the affine coordinates of M. 42
12 Construction of the pair (m, m'_∞) corresponding to the point M_∞ of the plane at infinity π_∞. 43
13 The collineation induced by the plane at infinity π_∞ is proportional to the matrix $\mathbf{R}$ of the 3-D rotation. 44
14 Computing the ratio of the lengths of the two segments AB and CD from the angles α, β, γ and δ. 45
15 Computing the ratio of the lengths of the two segments AB and CD from the vanishing points v_{ab}, v_{ac}, v_{bc}, v_{bd}, and v_{cd} of the image lines $\langle a, b\rangle$, $\langle a, c\rangle$, $\langle b, c\rangle$, $\langle b, d\rangle$ and $\langle c, d\rangle$. 46

List of Tables

1 Relations between the three strata and the geometric properties of the stereo rig. . . 47