Submitted to International Journal of Computer Vision. This version July 13, 1995. 29 pages.

The Geometry of Projective Reconstruction I: Matching Constraints and the Joint Image

BILL TRIGGS

Bill.Triggs@imag.fr

LIFIA, INRIA Rhône-Alpes, 46 avenue Félix Viallet, 38031 Grenoble, France.

Abstract.

This paper studies the geometry of perspective projection into multiple images and the matching constraints that this induces between the images. The combined projections produce a 3D subspace of the space of combined image coordinates called the **joint image**. This is a complete projective replica of the 3D world defined entirely in terms of image coordinates, up to an arbitrary choice of certain scale factors. Projective reconstruction is a canonical process in the joint image requiring only the rescaling of image coordinates. The matching constraints tell whether a set of image points is the projection of a single world point. In 3D there are only three types of matching constraint: the fundamental matrix, Shashua's trilinear tensor, and a new quadrilinear 4 image tensor. All of these fit into a single geometric object, the **joint image Grassmannian** tensor. This encodes exactly the information needed for reconstruction: the location of the joint image in the space of combined image coordinates.

Keywords: Computer Vision, Visual Reconstruction, Projective Geometry, Tensor Calculus, Grassmann Geometry

1. Introduction

This is the first of two papers that examine the geometry underlying the recovery of 3D projective structure from multiple images. This paper focuses on the geometry of multi-image projection and the **matching constraints** that this induces on image measurements. The second paper will deal with projective reconstruction techniques and error models.

Matching constraints like the fundamental matrix and Shashua's trilinear tensor [19] are currently a topic of lively interest in the vision community. This paper uncovers some of the beautiful and useful structure that lies behind them and should be of interest to anyone working on the geometry of vision. We will show that in three dimensions there are only three types of constraint: the fundamental matrix, Shashua's trilinear tensor, and a new quadrilinear four image tensor. All other matching constraints reduce trivially to one of these three types. Moreover, all of the constraint tensors fit very naturally into a single underlying geometric object, the **joint image Grassmannian**. Structural constraints on the Grassmannian tensor lead to quadratic relations between the matching tensors.

The joint image Grassmannian encodes precisely the portion of the imaging geometry that can be recovered from image measurements. It specifies the location of the **joint image**, a three dimensional submanifold of the space of combined image coordinates containing the matching m-tuples of image points. The topology of the joint image is complicated, but with an arbitrary choice of certain scale factors it becomes a 3D projective space containing a projective 'replica' of the 3D world. This replica is all that can be inferred about the world from image measurements. 3D reconstruction is an intrinsic, canonical geometric process only in the joint image, however an appropriate choice of basis there allows the results to be transferred to the original 3D world up to a projectivity.

This is a paper on the geometry of vision so there will be 'too many equations, no algorithms and no real images'. However it also represents a powerful new way to *think* about projective vision and that *does* have practical consequences. To understand this paper you will need to be comfortable with the tensorial approach to projective geometry: appendix A sketches the necessary background. This approach will be unfamiliar to many vision researchers, although a mathematician should have no problems with it. The change of notation is unfortunate but essential: the traditional matrix-vector notation is simply not powerful enough to express many of the concepts discussed here and becomes a real barrier to clear expression above a certain complexity. However in my experience effort spent learning the tensorial notation is amply repaid by increased clarity of thought.

In origin this work dates from the initial projective reconstruction papers of Faugeras & Maybank [3], [6], [5]. The underlying geometry of the situation was immediately evoked by those papers, although the details took several years to gel. In that time there has been a substantial amount of work on projective reconstruction. Faugeras' book [4] is an excellent general introduction and Maybank [15] provides a more mathematically oriented synthesis. Alternative approaches to projective reconstruction appear in Hartley *et.al.* [9] and Mohr *et.al.* [17]. Luong & Viéville [14] have studied 'canonic decompositions' of projection matrices for multiple views. Shashua [19] has developed the theory of the trilinear matching constraints, with input from Hartley [8]. A brief summary of the present paper appears in [20]. In parallel with the current work, both Werman & Shashua [22] and Faugeras & Mourrain [7] independently discovered the quadrilinear constraint and some of the related structure (but not the 'big picture' — the full joint image geometry). However the deepest debt of the current paper is to time spent in the Oxford mathematical physics research group lead by Roger Penrose [18], whose notation I have 'borrowed' and whose penetrating synthesis of the geometric and algebraic points of view has been a powerful tool and a constant source of inspiration.

2. Conventions and Notation

The world and images will be treated as projective spaces and expressed in homogeneous coordinates. Many equations will apply only up to scale, denoted $a \sim b$. The imaging process will be approximated by a perspective projection. Optical effects such as radial distortion and all the difficult problems of early vision will be ignored: we will basically assume that the images have already been reduced to a smoldering heap of geometry. When token matching between images is required, divine intervention will be invoked (or more likely a graduate student with a mouse).

Our main interest is in sequences of 2D images of ordinary 3D Euclidean space, but when it is straightforward to generalize to D_i dimensional images of d dimensional space we will do so. 1D 'linear' cameras and projection within a 2D plane are also practically important, and for clarity it is often easier to see the general case first.

Our notation is fully tensorial with all indices written out explicitly (*c.f.* appendix A). It is modelled on notation developed for mathematical physics and projective geometry by Roger Penrose [18]. Explicit indices are tedious for simple expressions but make complex tensor calculations *much* easier. Superscripts denote contravariant (*i.e.* point or vector) indices, while subscripts denote covariant (*i.e.* hyperplane, linear form or covector) ones. Contravariant and covariant indices transform inversely under changes of coordinates so that the **contraction** (*i.e.* 'dot product' or sum over all values) of a covariant-contravariant pair is invariant. The 'Einstein summation convention' applies: when the same index symbol appears in covariant and contravariant positions it denotes a contraction (implicit sum) over that index pair. For example $\mathbf{T}^a_b \mathbf{x}^b$ and $\mathbf{x}^b \mathbf{T}^a_b$ both stand for standard matrix-vector multiplication $\sum_b \mathbf{T}^a_b \mathbf{x}^b$. The repeated indices give the contraction, not the order of terms. Non-tensorial labels like image number are never implicitly summed over.

Different types of index denote different space or label types. This makes the notation a little baroque but it helps to keep things clear, especially when there are tensors with indices in several distinct spaces as will be common here. $\mathcal{H}^x$ denotes the *h*omogeneous vector space of objects (*i.e.* tensors) with index type x, while $\mathcal{P}^x$ denotes the associated *p*rojective space of such objects defined only up to nonzero scale: tensors $\mathbf{T}^x$ and $\lambda\mathbf{T}^x$ in $\mathcal{H}^x$ represent the same element of $\mathcal{P}^x$ for all $\lambda \neq 0$. We will not always distinguish points of $\mathcal{P}^x$ from their homogeneous representatives in $\mathcal{H}^x$. Indices $a, b, \ldots$ denote ordinary (projectivized homogenized d-dimensional) Euclidean space $\mathcal{P}^a$ ($a = 0, \ldots, d$), while $A_i, B_i, \ldots$ denote homogeneous coordinates in the D_i-dimensional i^{th} image $\mathcal{P}^{A_i}$ ($A_i = 0, \ldots, D_i$). When there are only two images A and A' are used in place of A_1 and A_2. Indices $i, j, \ldots = 1, \ldots, m$ are *i*mage labels, while $p, q, \ldots = 1, \ldots, n$ are *p*oint labels. Greek indices $\alpha, \beta, \ldots$ denote the combined homogeneous coordinates of all the images, thought of as a single big $(D+m)$-dimensional

joint image vector ($D = \sum_{i=1}^{m} D_i$). This is discussed in section 4.

The same base symbol will be used for 'the same thing' in different spaces, for example the equations $\mathbf{x}^{A_i} \sim \mathbf{P}_a^{A_i}\mathbf{x}^a$ $(i = 1,\ldots,m)$ denote the projection of a world point $\mathbf{x}^a \in \mathcal{P}^a$ to m distinct image points $\mathbf{x}^{A_i} \in \mathcal{P}^{A_i}$ via m distinct perspective projection matrices $\mathbf{P}_a^{A_i}$. These equations apply only up to scale and there is an implicit summation over all values of $a = 0,\ldots,d$.

We will follow the mathematicians' convention and use index 0 for homogenization, *i.e.* a Euclidean vector $(x^1 \cdots x^d)^\top$ is represented projectively as $(1\, x^1 \cdots x^d)^\top$ rather than $(x^1 \cdots x^d\ 1)^\top$. This seems more natural and makes notation and coding easier.

$\mathbf{T}^{[ab\ldots c]}$ denotes the result of antisymmetrizing the tensor $\mathbf{T}^{ab\ldots c}$ over all permutations of the indices $ab\ldots c$. For example $\mathbf{T}^{[ab]} \equiv \frac{1}{2}(\mathbf{T}^{ab} - \mathbf{T}^{ba})$. In any $d+1$ dimensional linear space there is a unique-up-to-scale $d+1$ index alternating tensor $\varepsilon^{a_0 a_1 \cdots a_n}$ and its dual $\varepsilon_{a_0 a_1 \cdots a_n}$. Up to scale, these have components ± 1 and 0 as $a_0 a_1 \ldots a_n$ is respectively an even or odd permutation of $01\ldots n$, or not a permutation at all. Any antisymmetric $k+1$ index contravariant tensor $\mathbf{T}^{[a_0 \cdots a_k]}$ can be 'dualized' to an antisymmetric $d-k$ index covariant one $(*\mathbf{T})_{a_{k+1}\cdots a_d} \equiv \frac{1}{(k+1)!}\ \varepsilon_{a_{k+1}\cdots a_d b_0 \cdots b_k}\mathbf{T}^{b_0\ldots b_k}$, and vice versa $\mathbf{T}^{a_0\ldots a_k} = \frac{1}{(d-k)!}\ (*\mathbf{T})_{b_{k+1}\cdots b_d}\ \varepsilon^{b_{k+1}\cdots b_d a_0 \cdots a_k}$, without losing information.

A k dimensional projective subspace of the d dimensional projective space $\mathcal{P}^a$ can be denoted by either the span of any $k+1$ independent points $\{\mathbf{x}_i^a \mid i = 0,\ldots,k\}$ in it or the intersection of any $d-k$ independent linear forms (hyperplanes) $\{\mathbf{l}_a^i \mid i = k+1,\ldots,d\}$ orthogonal to it. The antisymmetric tensors $\mathbf{x}_0^{[a_0} \ldots \mathbf{x}_k^{a_k]}$ and $\mathbf{l}_{[a_{k+1}}^{k+1} \cdots \mathbf{l}_{a_d]}^d$ uniquely define the subspace and are (up to scale) independent of the choice of points and forms and dual to each other. They are called respectively **Grassmann coordinates** and **dual Grassmann coordinates** for the subspace. Read appendix A for more details on this.

3. Prelude in F

As a prelude to the arduous general case, we will briefly consider the important sub-case of a single pair of 2D images of 3D space. The low dimensionality of this situation allows a slightly simpler (but ultimately equivalent) method of attack. We will work rapidly in homogeneous coordinates, viewing the 2D projective image spaces $\mathcal{P}^A$ and $\mathcal{P}^{A'}$ as 3D homogeneous vector spaces $\mathcal{H}^A$ and $\mathcal{H}^{A'}$ $(A = 0,1,2;\ A' = 0',1',2')$ and the 3D projective world space $\mathcal{P}^a$ as a 4D vector space $\mathcal{H}^a$ $(a = 0,\ldots,3)$. The perspective image projections are then 3×4 matrices $\mathbf{P}_a^A$ and $\mathbf{P}_a^{A'}$ defined only up to scale. Assuming that the projection matrices have rank 3, each has a 1D kernel that corresponds to a unique world point killed by the projection: $\mathbf{P}_a^A\ \mathbf{e}^A = \mathbf{0}$ and $\mathbf{P}_a^{A'}\ \mathbf{e}'^a = \mathbf{0}$. These points are called the **centres of projection** and each projects to the **epipole** in the opposite image: $\mathbf{e}^A \equiv \mathbf{P}_a^A \mathbf{e}'^a$ and $\mathbf{e}^{A'} \equiv \mathbf{P}_a^{A'}\mathbf{e}^a$. If the centres of projection are distinct, the two projections define a 3×3 rank 2 tensor called the **fundamental matrix** $\mathbf{F}_{AA'}$ [4]. This maps any given image point $\mathbf{x}^A$ ($\mathbf{x}^{A'}$) to a corresponding **epipolar line** $\mathbf{l}_{A'} \sim \mathbf{F}_{AA'}\mathbf{x}^A$ ($\mathbf{l}_A \sim \mathbf{F}_{AA'}\mathbf{x}^{A'}$) in the other image. Two image points correspond in the sense that they could be the projections of a single world point if and only if each lies on the epipolar line of the other: $\mathbf{F}_{AA'}\ \mathbf{x}^A\mathbf{x}^{A'} = 0$. The null directions of the fundamental matrix are the epipoles: $\mathbf{F}_{AA'}\ \mathbf{e}^A = \mathbf{0}$ and $\mathbf{F}_{AA'}\ \mathbf{e}^{A'} = \mathbf{0}$, so every epipolar line must pass through the corresponding epipole. The fundamental matrix $\mathbf{F}_{AA'}$ can be estimated from image correspondences even when the image projections are unknown.

Two image vectors $\mathbf{x}^A$ and $\mathbf{x}^{A'}$ can be packed into a single 6 component vector $\mathbf{x}^\alpha = (\mathbf{x}^A\ \ \mathbf{x}^{A'})^\top$ where $\alpha = 0,1,2,0',1',2'$. The space of such vectors will be called **homogeneous joint image space** $\mathcal{H}^\alpha$. Quotienting out the overall scale factor in $\mathcal{H}^\alpha$ produces a 5 dimensional projective space called **projective joint image space** $\mathcal{P}^\alpha$. The two 3×4 image projection matrices can be stacked into a single 6×4 **joint projection matrix** $\mathbf{P}_a^\alpha \equiv (\mathbf{P}_a^A\ \ \mathbf{P}_a^{A'})^\top$. If the centres of projection are distinct, no point in $\mathcal{P}^a$ is simultaneously killed by both projections, so the joint projection matrix has a vanishing kernel and hence rank 4. This implies that the joint projection is a nonsingular linear bijection from $\mathcal{H}^a$ onto its image space in $\mathcal{H}^\alpha$. This 4 dimensional image space will be called the **homogeneous joint image** $\mathcal{I}^\alpha$. Descending to $\mathcal{P}^\alpha$, the joint projection becomes a bijective projective equivalence between $\mathcal{P}^a$ and the **projective joint image** $\mathcal{PI}^\alpha$ (the projection of $\mathcal{I}^\alpha$ into $\mathcal{P}^\alpha$). The projection of $\mathcal{PI}^\alpha$ to each image is just a trivial deletion of coordinates, so *the projective joint image is a complete projective replica of the world space in image coordinates.*

Unfortunately, $\mathcal{PI}^\alpha$ is not quite unique. Any rescaling $\{\mathbf{P}_a^A, \mathbf{P}_a^{A'}\} \to \{\lambda \mathbf{P}_a^A, \lambda' \mathbf{P}_a^{A'}\}$ of the underlying projection matrices produces a different but equivalent space $\mathcal{PI}^\alpha$. However modulo this arbitrary choice of scaling the projective joint image is canonically defined by the physical situation.

Now suppose that the projection matrices are unknown but the fundamental matrix has been estimated from image measurements. Since $\mathbf{F}$ has rank 2, it can be decomposed (non-uniquely!) as

$$\mathbf{F}_{AA'} \;=\; \mathbf{u}_A\,\mathbf{v}_{A'} - \mathbf{v}_A\,\mathbf{u}_{A'} \;=\; \mathrm{Det}\begin{pmatrix} \mathbf{u}_A & \mathbf{u}_{A'} \\ \mathbf{v}_A & \mathbf{v}_{A'} \end{pmatrix}$$

where $\mathbf{u}_A \not\sim \mathbf{v}_A$ and $\mathbf{u}_{A'} \not\sim \mathbf{v}_{A'}$ are two pairs of independent image covectors. It is easy to see that $\mathbf{u}_A \leftrightarrow \mathbf{u}_{A'}$ and $\mathbf{v}_A \leftrightarrow \mathbf{v}_{A'}$ are actually pairs of corresponding epipolar lines[1]. In terms of joint image space, the $\mathbf{u}$'s and $\mathbf{v}$'s can be viewed as a pair of 6 component covectors defining a 4 dimensional linear subspace $\mathcal{I}^\alpha$ of $\mathcal{H}^\alpha$ via the equations:

$$\mathcal{I}^\alpha \equiv \left\{ \begin{pmatrix} \mathbf{x}^A \\ \mathbf{x}^{A'} \end{pmatrix} \;\middle|\; \begin{pmatrix} \mathbf{u}_A\,\mathbf{x}^A + \mathbf{u}_{A'}\,\mathbf{x}^{A'} \\ \mathbf{v}_A\,\mathbf{x}^A + \mathbf{v}_{A'}\,\mathbf{x}^{A'} \end{pmatrix} = \begin{pmatrix} \mathbf{u}_A & \mathbf{u}_{A'} \\ \mathbf{v}_A & \mathbf{v}_{A'} \end{pmatrix} \begin{pmatrix} \mathbf{x}^A \\ \mathbf{x}^{A'} \end{pmatrix} \;=\; \mathbf{0} \right\}$$

Trivial use of the constraint equations shows that any point $(\mathbf{x}^A \;\; \mathbf{x}^{A'})^\top$ of $\mathcal{I}^\alpha$ automatically satisfies the epipolar constraint $\mathbf{F}_{AA'}\,\mathbf{x}^A\mathbf{x}^{A'} = 0$. In fact, given any $(\mathbf{x}^A \;\; \mathbf{x}^{A'})^\top \in \mathcal{H}^\alpha$, the equations

$$\begin{aligned} \mathbf{0} \;&=\; \begin{pmatrix} \mathbf{u}_A & \mathbf{u}_{A'} \\ \mathbf{v}_A & \mathbf{v}_{A'} \end{pmatrix} \begin{pmatrix} \lambda\,\mathbf{x}^A \\ \lambda'\,\mathbf{x}^{A'} \end{pmatrix} \\ &=\; \begin{pmatrix} \mathbf{u}_A\mathbf{x}^A & \mathbf{u}_{A'}\mathbf{x}^{A'} \\ \mathbf{v}_A\mathbf{x}^A & \mathbf{v}_{A'}\mathbf{x}^{A'} \end{pmatrix} \begin{pmatrix} \lambda \\ \lambda' \end{pmatrix} \end{aligned}$$

have a nontrivial solution if and only if

$$\mathbf{F}_{AA'}\,\mathbf{x}^A\mathbf{x}^{A'} \;=\; \mathrm{Det}\begin{pmatrix} \mathbf{u}_A\mathbf{x}^A & \mathbf{u}_{A'}\mathbf{x}^{A'} \\ \mathbf{v}_A\mathbf{x}^A & \mathbf{v}_{A'}\mathbf{x}^{A'} \end{pmatrix} \;=\; 0$$

In other words, the set of matching point pairs in the two images is exactly the set of pairs that can be rescaled to lie in $\mathcal{I}^\alpha$. *Up to a rescaling, the joint image is the set of matching points in the two images.*

A priori, $\mathcal{I}^\alpha$ depends on the choice of the decomposition $\mathbf{F}_{AA'} = \mathbf{u}_A\,\mathbf{v}_{A'} - \mathbf{v}_A\,\mathbf{u}_{A'}$. In fact appendix B shows that the most general redefinition of the $\mathbf{u}$'s and $\mathbf{v}$'s that leaves $\mathbf{F}$ unchanged up to scale is

$$\begin{pmatrix} \mathbf{u}_A & \mathbf{u}_{A'} \\ \mathbf{v}_A & \mathbf{v}_{A'} \end{pmatrix} \longrightarrow \Lambda \begin{pmatrix} \mathbf{u}_A & \mathbf{u}_{A'} \\ \mathbf{v}_A & \mathbf{v}_{A'} \end{pmatrix} \begin{pmatrix} 1/\lambda & 0 \\ 0 & 1/\lambda' \end{pmatrix}$$

where Λ is an arbitrary nonsingular 2×2 matrix and $\{\lambda, \lambda'\}$ are arbitrary nonzero relative scale factors. Λ is a linear mixing of the constraint vectors and has no effect on the location of $\mathcal{I}^\alpha$, but λ and λ' represent rescalings of the image coordinates that move $\mathcal{I}^\alpha$ bodily according to

$$\begin{pmatrix} \mathbf{x}^A \\ \mathbf{x}^{A'} \end{pmatrix} \longrightarrow \begin{pmatrix} \lambda\,\mathbf{x}^A \\ \lambda'\,\mathbf{x}^{A'} \end{pmatrix}$$

Hence, given $\mathbf{F}$ and an arbitrary choice of the relative image scaling the joint image $\mathcal{I}^\alpha$ is defined uniquely.

Appendix B also shows that given any pair of nonsingular projection matrices $\mathbf{P}_a^A$ and $\mathbf{P}_a^{A'}$ compatible with $\mathbf{F}_{AA'}$ in the sense that the projection of every point of $\mathcal{P}^a$ satisfies the epipolar constraint $\mathbf{F}_{AA'}\;\mathbf{P}_a^A\mathbf{P}_b^{A'}\;\mathbf{x}^a\mathbf{x}^b \;=\; 0$, the $\mathcal{I}^\alpha$ arising from factorization of $\mathbf{F}$ is projectively equivalent to the $\mathcal{I}^\alpha$ arising from the projection matrices. (Here, nonsingular means that each matrix has rank 3 and the joint matrix has rank 4, *i.e.* the centres of projection are unique and distinct). In fact there is a constant rescaling $\{\mathbf{P}_a^A, \mathbf{P}_a^{A'}\} \to \{\lambda\,\mathbf{P}_a^A, \lambda'\,\mathbf{P}_a^{A'}\}$ that makes the two coincide.

In summary, the fundamental matrix can be factorized to define a three dimensional projective subspace $\mathcal{PI}^\alpha$ of the space of combined image coordinates. $\mathcal{PI}^\alpha$ is projectively equivalent to the 3D world and uniquely defined by the images up to an arbitrary choice of a single relative scale factor. Projective reconstruction in $\mathcal{PI}^\alpha$ is simply a matter of rescaling the homogeneous image measurements. This paper investigates the geometry of $\mathcal{PI}^\alpha$ and its multi-image counterparts and argues that up to the choice of scale factor, they provide *the* natural canonical projective reconstruction of the information in the images: all other reconstructions are merely different ways of looking at the information contained in $\mathcal{PI}^\alpha$.

4. Too Many Joint Images

Now consider the general case of projection into $m \geq 1$ images. We will model the world and images respectively as d and D_i dimensional projective spaces

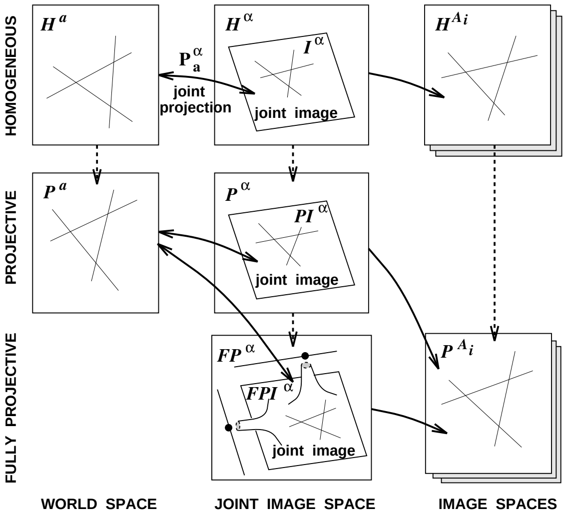

Fig. 1. The various joint images and projections.

$\mathcal{P}^a$ $(a = 0, \ldots, d)$ and $\mathcal{P}^{A_i}$ $(A_i = 0, \ldots, D_i,$ $i = 1, \ldots, m)$ and use homogeneous coordinates everywhere. It may appear more natural to use Euclidean or affine spaces, but when it comes to discussing perspective projection it is simpler to view things as (fragments of) projective space. The usual Cartesian and pixel coordinates are still inhomogeneous local coordinate systems covering almost all of the projective world and image manifolds, so projectivization does not change the essential situation too much.

In homogeneous coordinates the perspective image projections are represented by homogeneous $(D_i + 1) \times (d + 1)$ matrices $\{\mathbf{P}_a^{A_i} | i = 1, \ldots, m\}$ that take homogeneous representatives of world points $\mathbf{x}^a \in \mathcal{P}^a$ to homogeneous representatives of image points $\mathbf{x}^{A_i} \sim \mathbf{P}_a^{A_i} \mathbf{x}^a \in \mathcal{P}^{A_i}$. The homogeneous vectors and matrices representing world points $\mathbf{x}^a$, image points $\mathbf{x}^{A_i}$ and projections $\mathbf{P}_a^{A_i}$ are each defined only up to scale. Arbitrary nonzero rescalings of them do not change the physical situation because the rescaled world and image vectors still represent the same points of the underlying projective spaces $\mathcal{P}^a$ and $\mathcal{P}^{A_i}$, and the projection equations $\mathbf{x}^{A_i} \sim \mathbf{P}_a^{A_i}$ still hold up to scale.

Any collection of m image points $\{\mathbf{x}^{A_i} | i = 1, \ldots, m\}$ can be viewed as a single point in the Cartesian product $\mathcal{P}^{A_1} \times \mathcal{P}^{A_2} \times \cdots \times \mathcal{P}^{A_m}$ of the individual projective image spaces. This is a $D = \sum_{i=1}^m D_i$ dimensional differentiable manifold whose local inhomogeneous coordinates are just the combined pixel coordinates of all the image points. Since any m-tuple of matching points is an element of $\mathcal{P}^{A_1} \times \cdots \times \mathcal{P}^{A_m}$, it may seem that this space is the natural arena for multi-image projective reconstruction. This is almost true but we need to be a little more careful. Although most world points can be represented by their projections in $\mathcal{P}^{A_1} \times \cdots \times \mathcal{P}^{A_m}$, the centres of projection are missing because they fail to project to anything at all in their own images. To represent these, extra points must be glued on to $\mathcal{P}^{A_1} \times \cdots \times \mathcal{P}^{A_m}$.

When discussing perspective projections it is convenient to introduce homogeneous coordinates. A separate homogenizer is required for each image, so the result is just the Cartesian product $\mathcal{H}^{A_1} \times \mathcal{H}^{A_2} \times \cdots \times \mathcal{H}^{A_m}$ of the individual homogeneous image spaces $\mathcal{H}^{A_i}$. We will call this $D+m$ dimensional vector space **homogeneous joint image space** $\mathcal{H}^\alpha$. By quotienting out the overall scale factor in $\mathcal{H}^\alpha$ in the usual way, we can view it as a $D + m - 1$ dimensional projective space $\mathcal{P}^\alpha$ called **projective joint image space**. This is a *bona fide* projective space but it still contains the arbitrary relative scale factors of the component images. A point of $\mathcal{H}^\alpha$ can be represented as a $D + m$ component column vector $\mathbf{x}^\alpha = (\mathbf{x}^{A_1} \cdots \mathbf{x}^{A_m})^\top$ where the $\mathbf{x}^{A_i}$ are homogeneous coordinate vectors in each image. We will think of the index α as taking values $0_1, 1_1, \ldots, D_i, 0_{i+1}, \ldots, D_m$, where the subscripts indicate the image the coordinate came from. An individual image vector $\mathbf{x}^{A_i}$ can be thought of as a vector in $\mathcal{H}^\alpha$ whose non-image-i components vanish.

Since the coordinates of each image are only defined up to scale, the natural definition of the equivalence relation '$\sim$' on $\mathcal{H}^\alpha$ is 'equality up to individual rescalings of the component images': $(\mathbf{x}^{A_1} \ \cdots \ \mathbf{x}^{A_m})^\top \sim (\lambda_1 \mathbf{x}^{A_1} \ \cdots \ \lambda_m \mathbf{x}^{A_m})^\top$ for all $\{\lambda_i \neq 0\}$. So long as none of the $\mathbf{x}^{A_i}$ vectors vanish, the equivalence classes of '$\sim$' are m-dimensional subspaces of $\mathcal{H}^\alpha$ that correspond exactly to the points of $\mathcal{P}^{A_1} \times \cdots \times \mathcal{P}^{A_m}$. However when some of the $\mathbf{x}^{A_i}$ vanish the equivalence classes are lower dimensional subspaces that have no corresponding point in $\mathcal{P}^{A_1} \times \cdots \times \mathcal{P}^{A_m}$. We will call the entire stratified set of equivalence classes **fully projective joint image space** $\mathcal{FP}^\alpha$. This is basically $\mathcal{P}^{A_1} \times \cdots \times \mathcal{P}^{A_m}$ augmented with the lower dimensional product spaces $\mathcal{P}^{A_i} \times \cdots \times \mathcal{P}^{A_j}$ for each proper subset of images $i, \ldots, j$. Most world points project to 'regular' points of $\mathcal{FP}^\alpha$ in $\mathcal{P}^{A_1} \times \cdots \times \mathcal{P}^{A_m}$, but the centres of projection project into lower dimensional fragments of $\mathcal{FP}^\alpha$.

A set of perspective projections into m projective images $\mathcal{P}^{A_i}$ defines a unique **joint projection** into the fully projective joint projective image space $\mathcal{FP}^\alpha$. Given an arbitrary choice of scaling for the homogeneous representatives $\{\mathbf{P}_a^{A_i} \mid i = 1, \ldots, m\}$ of the individual image projections, the joint projection can be represented as a single $(D+m) \times (d+1)$ **joint projection matrix**

$$\mathbf{P}_a^\alpha \;\equiv\; \begin{pmatrix} \mathbf{P}_a^{A_1} \\ \vdots \\ \mathbf{P}_a^{A_m} \end{pmatrix} : \mathcal{H}^a \longrightarrow \mathcal{H}^\alpha$$

which defines a projective mapping between the underlying projective spaces $\mathcal{P}^a$ and $\mathcal{P}^\alpha$. A rescaling $\{\mathbf{P}_a^{A_i}\} \to \{\lambda_i\, \mathbf{P}_a^{A_i}\}$ of the individual image projection matrices does not change the physical situation or the fully projective joint projection on $\mathcal{FP}^\alpha$, but it *does* change the joint projection matrix $\mathbf{P}_a^\alpha$ and the resulting projections from $\mathcal{H}^a$ to $\mathcal{H}^\alpha$ and from $\mathcal{P}^a$ to $\mathcal{P}^\alpha$. An arbitrary choice of the individual projection scalings is always necessary to make things concrete.

Given a choice of scaling for the components of $\mathbf{P}_a^\alpha$, the image of $\mathcal{H}^a$ in $\mathcal{H}^\alpha$ under the joint projection $\mathbf{P}_a^\alpha$ will be called the **homogeneous joint image** $\mathcal{I}^\alpha$. This is the set of joint image space points that are the projection of some point in world space: $\{\mathbf{P}_a^\alpha\, \mathbf{x}^a \in \mathcal{H}^\alpha \mid \mathbf{x}^a \in \mathcal{H}^a\}$. In $\mathcal{I}^\alpha$, each world point is represented by its homogeneous vector of image coordinates. Similarly we can define the projective and fully projective joint images $\mathcal{PI}^\alpha$ and $\mathcal{FPI}^\alpha$ as the images of the projective world space $\mathcal{P}^a$ in the projective and fully projective joint image spaces $\mathcal{P}^\alpha$ and $\mathcal{FP}^\alpha$ under the projective and fully projective joint projections. (Equivalently, $\mathcal{PI}^\alpha$ and $\mathcal{FPI}^\alpha$ are the projections of $\mathcal{I}^\alpha$ to $\mathcal{P}^\alpha$ and $\mathcal{FP}^\alpha$).

If the $(D+m) \times (d+1)$ joint projection matrix $\mathbf{P}_a^\alpha$ has rank less than $d+1$ it will have a nontrivial kernel and many world points will project to the same set of image points, so unique reconstruction will be impossible. On the other hand if $\mathbf{P}_a^\alpha$ has rank $d+1$, the homogeneous joint image $\mathcal{I}^\alpha$ will be a $d+1$ dimensional linear subspace of $\mathcal{H}^\alpha$ and $\mathbf{P}_a^\alpha$ will be a nonsingular linear bijection from $\mathcal{H}^a$ onto $\mathcal{I}^\alpha$. Similarly, the projective joint projection will define a nonsingular projective bijection from $\mathcal{P}^a$ onto the d dimensional projective space $\mathcal{PI}^\alpha$ and the fully projective joint projection will be a bijection (and at most points a diffeomorphism) from $\mathcal{P}^a$ onto $\mathcal{FPI}^\alpha$ in $\mathcal{FP}^\alpha$. Structure in $\mathcal{P}^a$ will be mapped bijectively to projectively equivalent structure in $\mathcal{PI}^\alpha$, so $\mathcal{PI}^\alpha$ will be 'as good as' $\mathcal{P}^a$ as far as projective reconstruction is concerned. Moreover, projection from $\mathcal{PI}^\alpha$ to the individual images is a trivial throwing away of coordinates and scale factors, so structure in $\mathcal{PI}^\alpha$ has a very direct relationship with image measurements.

Unfortunately, although $\mathcal{PI}^\alpha$ is closely related to the images it is not quite canonically defined by the physical situation because it moves when the individual image projection matrices are rescaled. However, the truly canonical structure — the fully projective joint image $\mathcal{FPI}^\alpha$ — has a complex stratified structure that is not so easy to handle. When restricted to the product space $\mathcal{P}^{A_1} \times \cdots \times \mathcal{P}^{A_m}$, $\mathcal{FPI}^\alpha$ is equivalent to the projective space $\mathcal{P}^a$ with each centre of projection 'blown up' to the corresponding image space $\mathcal{P}^{A_i}$. The missing centres of projection lie in lower strata of $\mathcal{FP}^\alpha$. Given this complication, it seems easier to work with the simple projective space $\mathcal{PI}^\alpha$ or its homogeneous representative $\mathcal{I}^\alpha$ and to accept that an arbitrary choice of scale factors will be required. We will do this from now on, but it is important to verify that this arbitrary choice does not affect the final results, particularly as far as numerical methods and error models are concerned. It is also essential to realize that although *for any one point* the projection scale factors can be chosen arbitrarily, once they are chosen they apply uniformly to all other points: *no matter which scaling is chosen, there is a strong coherence between the scalings of different points*. A central theme of this paper is that the essence of projective reconstruction is the recovery of this scale coherence from image measurements.

5. The Joint Image Grassmannian Tensor

We can view the joint projection matrix $\mathbf{P}_a^\alpha$ (with some choice of the internal scalings) in two ways: *(i)* as a collection of m projection matrices from $\mathcal{P}^a$ to the m images $\mathcal{P}^{A_i}$; *(ii)* as a set of $d+1$ $(D+m)$-component column vectors $\{\mathbf{P}_a^\alpha \mid a = 0, \ldots, d\}$ that span the joint image subspace $\mathcal{I}^\alpha$ in $\mathcal{H}^\alpha$. From the second point of view the images of the standard basis $\{(10\cdots0)^\top, (01\cdots0)^\top, \ldots, (00\cdots1)^\top\}$ for $\mathcal{H}^a$ (*i.e.* the columns of $\mathbf{P}_a^\alpha$) form a basis for $\mathcal{I}^\alpha$ and a set of homogeneous coordinates $\{x^a \mid a = 0, \ldots, d\}$ can be viewed either as the coordinates of a point $\mathbf{x}^a$ in $\mathcal{P}^a$ or as the coordinates of a point $\mathbf{P}_a^\alpha \mathbf{x}^a$ in $\mathcal{I}^\alpha$ with respect to the basis $\{\mathbf{P}_a^\alpha \mid a = 0, \ldots, d\}$. Similarly, the columns of $\mathbf{P}_a^\alpha$ and the $(d+2)^{nd}$ column $\sum_{a=0}^{d} \mathbf{P}_a^\alpha$ form a projective basis for $\mathcal{PI}^\alpha$

that is the image of the standard projective basis $\{(10\cdots0)^\top,\ldots,(00\cdots1)^\top,(11\cdots1)^\top\}$ for $\mathcal{P}^a$.

This means that *any reconstruction in $\mathcal{P}^a$ can be viewed as reconstruction in $\mathcal{PI}^\alpha$ with respect to a particular choice of basis there.* This is important because we will see that (up to a choice of scale factors) $\mathcal{PI}^\alpha$ *is canonically defined by the imaging situation and can be recovered directly from image measurements.* In fact we will show that the information in the combined matching constraints is exactly the location of the subspace $\mathcal{PI}^\alpha$ in $\mathcal{P}^\alpha$, and this is exactly the information we need to make a *canonical geometric reconstruction of $\mathcal{P}^a$ in $\mathcal{PI}^\alpha$ from image measurements*.

By contrast we can not hope to recover the basis in $\mathcal{P}^a$ or the individual columns of $\mathbf{P}_a^\alpha$ by image measurements. In fact any two worlds that project to the same joint image are indistinguishable so far as image measurements are concerned. Under an arbitrary nonsingular projective transformation $\mathbf{x}^a \to \tilde{\mathbf{x}}^{a'} = (\Lambda^{-1})^{a'}{}_b\mathbf{x}^b$ between $\mathcal{P}^a$ and some other world space $\mathcal{P}^{a'}$, the projection matrices (and hence the basis vectors for $\mathcal{PI}^\alpha$) must change according to $\mathbf{P}_a^\alpha \to \tilde{\mathbf{P}}_{a'}^\alpha = \mathbf{P}_b^\alpha\,\Lambda^b{}_{a'}$ to compensate. The new basis vectors are a linear combination of the old ones so the space $\mathcal{PI}^\alpha$ they span is not changed, but the individual vectors *are* changed: all we can hope to recover from the images is the geometric location of $\mathcal{PI}^\alpha$, not its particular basis.

But how can we specify the location of $\mathcal{PI}^\alpha$ geometrically? We originally defined it as the span of the columns of the joint projection $\mathbf{P}_a^\alpha$, but that is rather inconvenient. For one thing $\mathcal{PI}^\alpha$ depends only on the span and not on the individual vectors, so it is redundant to specify every component of $\mathbf{P}_a^\alpha$. What is worse, the redundant components are exactly the things that can not be recovered from image measurements. It is not even clear how we would use a 'span' even if we did manage to obtain it.

Algebraic geometers encountered this sort of problem long ago and developed a useful partial solution called **Grassmann coordinates** (see appendix A). Recall that $[a\cdots c]$ denotes antisymmetrization over all permutations of the indices $a\cdots c$. Given $k+1$ independent vectors $\{\mathbf{x}_i^a \,|\, i=0,\ldots,k\}$ in a $d+1$ dimensional vector space $\mathcal{H}^a$, it turns out that the antisymmetric $k+1$ index **Grassmann tensor** $\mathbf{x}^{a_0\cdots a_k} \equiv \mathbf{x}_0^{[a_0}\cdots\mathbf{x}_k^{a_k]}$ uniquely characterizes the $k+1$ dimensional subspace spanned by the vectors and (up to scale) does not depend on the particular vectors of the subspace chosen to define it. In fact a point $\mathbf{y}^a$ lies in the span if and only if it satisfies $\mathbf{x}^{[a_0\cdots a_k}\mathbf{y}^{a_{k+1}]} = \mathbf{0}$, and under a $(k+1)\times(k+1)$ linear redefinition Λ_j^i of the basis elements $\{\mathbf{x}_i^a\}$, $\mathbf{x}^{a_0\cdots a_k}$ is simply rescaled by $\mathrm{Det}(\Lambda)$. Up to scale, the components of the Grassmann tensor are the $(k+1)\times(k+1)$ minors of the $(d+1)\times(k+1)$ matrix of components of the $\mathbf{x}_i^a$.

The antisymmetric tensors are global coordinates for the k dimensional subspaces in the sense that each subspace is represented by a unique (up to scale) Grassmann tensor. However the parameterization is highly redundant: for $1 \le k \le d-2$ the $k+1$ index antisymmetric tensors have many more independent components than there are degrees of freedom. In fact only the very special antisymmetric tensors that can be written in the above 'simple' form $\mathbf{x}_0^{[a_0}\cdots\mathbf{x}_k^{a_k]}$ specify subspaces. Those that can are characterized by the quadratic **Grassmann simplicity relations** $\mathbf{x}^{a_0\cdots[a_k}\,\mathbf{x}^{b_0\cdots b_k]} = \mathbf{0}$.

In the present case the $d+1$ columns of $\mathbf{P}_a^\alpha$ specify the d dimensional joint image subspace $\mathcal{PI}^\alpha$. Instead of antisymmetrizing over the image space indices α we can get the same effect by contracting the world space indices a with the $d+1$ dimensional alternating tensor. This gives the $d+1$ index antisymmetric **joint image Grassmannian** tensor

$$\begin{aligned} \mathbf{I}^{\alpha_0\alpha_1\cdots\alpha_d} &\equiv \tfrac{1}{(d+1)!}\,\mathbf{P}_{a_0}^{\alpha_0}\,\mathbf{P}_{a_1}^{\alpha_1}\cdots\mathbf{P}_{a_d}^{\alpha_d}\;\varepsilon^{a_0a_1\cdots a_d} \\ &\sim\; \mathbf{P}_0^{[\alpha_0}\,\mathbf{P}_1^{\alpha_1}\cdots\mathbf{P}_d^{\alpha_d]} \end{aligned}$$

Although we have defined the Grassmann tensor in terms of the columns of the projection matrix basis for $\mathcal{PI}^\alpha$, it is actually an intrinsic property of $\mathcal{PI}^\alpha$ that defines and is defined by it in a manner completely independent of the choice of basis (up to scale). In fact we will see that the Grassmann tensor contains exactly the same information as the complete set of matching constraint tensors. Since the matching constraints can be recovered from image measurements, the Grassmann tensor can be too.

As a simple test of plausibility, let us verify that the Grassmann tensor has the correct number of degrees of freedom to encode the imaging geometry required for projective reconstruction. The geometry of an m camera imaging system can be specified by giving each of the m projection mappings modulo an arbitrary overall choice of projective basis in $\mathcal{P}^a$. Up to an arbitrary scale factor, a $(D_i+1)\times(d+1)$ projection matrix is defined by $(D_i + 1)(d + 1) - 1$ parameters while a projective basis in $\mathcal{P}^a$ has $(d+1)(d+1)-1$ degrees of freedom. The m camera projective geometry therefore has

$$\sum_{i=1}^{m}\Big((D_i+1)\,(d+1)-1\Big) \; - \; \big((d+1)^2-1\big) \\ = \; (D+m-d-1)\,(d+1)-m+1$$

independent degrees of freedom. For example $11m - 15$ parameters are required to specify the geometry of m 2D cameras viewing 3D projective space [14].

The antisymmetric Grassmann tensor $\mathbf{I}^{\,\alpha_0\cdots\alpha_d}$ has $\binom{D+m}{d+1}$ *linearly* independent components. However the quadratic Grassmann relations reduce the number of *algebraically* independent components to the dimension $(D + m - d - 1)(d + 1)$ of the space of possible locations of the joint image $\mathcal{I}^\alpha$ in $\mathcal{P}^\alpha$. (Joint image locations are locally parameterized by the $((D+m)-(d+1)) \times (d+1)$ matrices, or equivalently by giving $d+1$ $(D+m)$-component spanning basis vectors in $\mathcal{P}^\alpha$ modulo $(d+1) \times (d+1)$ linear redefinitions). The overall scale factor of $\mathbf{I}^{\,\alpha_0\cdots\alpha_d}$ has already been subtracted from this count, but it still contains the $m-1$ arbitrary *relative* scale factors of the m images. Subtracting these leaves the Grassmann tensor (or the equivalent matching constraint tensors) with $(D+m-d-1)\,(d+1) - m + 1$ physically meaningful degrees of freedom. This agrees with the above degree-of-freedom count based on projection matrices.

6. Reconstruction Equations

Suppose we are given a set of m image points $\{\mathbf{x}^{A_i} \,|\, i = 1,\ldots,m\}$ that may correspond to an unknown world point $\mathbf{x}^a$ via some known projection matrices $\mathbf{P}_a^{A_i}$. Can the world point $\mathbf{x}^a$ be recovered, and if so, how?

As usual we will work projectively in homogeneous coordinates and suppose that arbitrary nonzero scalings have been chosen for the $\mathbf{x}^{A_i}$ and $\mathbf{P}_a^{A_i}$. The image vectors can be stacked into a $D+m$ component joint homogeneous image vector $\mathbf{x}^\alpha$ and the projection matrices can be stacked into a $(D+m)\times(d+1)$ component joint homogeneous projection matrix, where d is the world dimension and $D = \sum_{i=1}^m D_i$ is the sum of the image dimensions.

Any candidate reconstruction $\mathbf{x}^a$ must project to the correct point in each image: $\mathbf{x}^{A_i} \sim \mathbf{P}_a^{A_i}\,\mathbf{x}^a$. Inserting variables $\{\lambda_i \,|\, i = 1,\ldots,m\}$ to represent the unknown scale factors gives m homogeneous equations $\mathbf{P}_a^{A_i}\,\mathbf{x}^a - \lambda_i\,\mathbf{x}^{A_i} = \mathbf{0}$. These can be written as a single $(D+m)\times(d+1+m)$ homogeneous linear system, the **basic reconstruction equations**:

$$\left(\begin{array}{c|cccc} & \mathbf{x}^{A_1} & \mathbf{0} & \cdots & \mathbf{0} \\ & \mathbf{0} & \mathbf{x}^{A_2} & \cdots & \mathbf{0} \\ \mathbf{P}_a^\alpha & \vdots & \vdots & \ddots & \vdots \\ & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x}^{A_m} \end{array}\right) \begin{pmatrix} \mathbf{x}^a \\ \hline -\lambda_1 \\ -\lambda_2 \\ \vdots \\ -\lambda_m \end{pmatrix} = \mathbf{0}$$

Any nonzero solution of these equations gives a reconstructed world point $\mathbf{x}^a$ consistent with the image measurements $\mathbf{x}^{A_i}$, and also provides the unknown scale factors $\{\lambda_i\}$.

These equations will be studied in detail in the next section. However we can immediately remark that if there are less image measurements than world dimensions ($D < d$) there will be at least two more free variables than equations and the solution (if it exists) can not be unique. So from now on we require $D \geq d$.

On the other hand, if there are more measurements than world dimensions ($D > d$) the system will usually be overspecified and a solution will exist only when certain constraints between the projection matrices $\mathbf{P}_a^{A_i}$ and the image measurements $\mathbf{x}^{A_i}$ are satisfied. We will call these constraints **matching constraints** and the inter-image tensors they generate **matching tensors**. The simplest example is the epipolar constraint.

It is also clear that there is no hope of a unique solution if the rank of the joint projection matrix $\mathbf{P}_a^\alpha$ is less than $d+1$, because any vector in the kernel of $\mathbf{P}_a^\alpha$ can be added to a solution without changing the projection at all. So we will also require the joint projection matrix to have maximal rank (*i.e.* $d+1$). Recall that this implies that the joint projection $\mathbf{P}_a^\alpha$ is a bijection from $\mathcal{P}^a$ onto its image the joint image $\mathcal{PI}^\alpha$ in $\mathcal{P}^\alpha$. (This is necessary but not always sufficient for a unique reconstruction).

In the usual 3D→2D case the individual projections are 3×4 rank 3 matrices and each has a one dimensional kernel: the centre of projection. Provided there are at least two distinct centres of projection among the image projections, no point will project to zero in every image and the joint projection will have a vanishing kernel and hence maximal rank. (It turns out that in this case $\mathrm{Rank}(\mathbf{P}_a^\alpha) = 4$ is also *sufficient* for a unique reconstruction).

Recalling that the joint projection columns $\{\mathbf{P}_a^\alpha \,|\, a = 0, \ldots, d\}$ form a basis for the homogeneous joint image $\mathcal{I}^\alpha$ and treating the $\mathbf{x}^{A_i}$ as vectors in $\mathcal{H}^\alpha$ whose other components vanish, we can interpret the reconstruction equations as the geometrical statement that the space spanned by the image vectors $\{\mathbf{x}^{A_i} \,|\, i = 1, \ldots, m\}$ in $\mathcal{H}^\alpha$ must intersect $\mathcal{I}^\alpha$. At the intersection there is a point of $\mathcal{H}^\alpha$ that can be expressed: *(i)* as a rescaling of the image measurements $\sum_i \lambda_i \, \mathbf{x}^{A_i}$; *(ii)* as a point of $\mathcal{I}^\alpha$ with coordinates $\mathbf{x}^a$ in the basis $\{\mathbf{P}_a^\alpha \,|\, a = 0, \ldots, d\}$; *(iii)* as the projection into $\mathcal{I}^\alpha$ of a world point $\mathbf{x}^a$ under $\mathbf{P}_a^\alpha$. (Since $\mathcal{H}^a$ is isomorphic to $\mathcal{I}^\alpha$ under $\mathbf{P}_a^\alpha$, the last two points of view are equivalent).

This construction is important because although neither the coordinate system in $\mathcal{H}^a$ nor the columns of $\mathbf{P}_a^\alpha$ can be recovered from image measurements, the joint image $\mathcal{I}^\alpha$ *can* be recovered (up to an arbitrary choice of relative scaling). In fact the content of the matching constraints is *precisely* the location of $\mathcal{I}^\alpha$ in $\mathcal{H}^\alpha$. This gives a completely geometric and almost canonical projective reconstruction technique in $\mathcal{I}^\alpha$ that requires only the scaling of joint image coordinates. A choice of basis in $\mathcal{I}^\alpha$ is necessary only to map the construction back into world coordinates.

Recalling that the joint image can be located by giving its Grassmann coordinate tensor $\mathbf{I}^{\,\alpha\beta\cdots\gamma}$ and that in terms of this a point lies in the joint image if and only if $\mathbf{I}^{\,[\alpha\beta\cdots\gamma}\,\mathbf{x}^{\delta]} = \mathbf{0}$, the basic reconstruction system is equivalent to the following **joint image reconstruction equations**

$$\mathbf{I}^{\,[\alpha\beta\cdots\gamma} \cdot \left(\sum_{i=1}^{m} \lambda_i \, \mathbf{x}^{A_i]} \right) \;=\; \mathbf{0}$$

This is a redundant system of homogeneous linear equations for the λ_i given the $\mathbf{I}^{\,\alpha\beta\cdots\gamma}$ and the $\mathbf{x}^{A_i}$. It will be used in section 10 to derive implicit 'reconstruction' methods that are independent of any choice of world or joint image basis.

There is yet another form of the reconstruction equations that is more familiar and compact but slightly less symmetrical. For notational convenience suppose that $\mathbf{x}^{0_i} \neq 0$. (We use component 0 for normalization. Each image vector has at least one nonzero component so the coordinates can be relabelled if necessary so that $\mathbf{x}^{0_i} \neq 0$). The projection equations $\mathbf{P}_a^{A_i}\,\mathbf{x}^a = \lambda_i\,\mathbf{x}^{A_i}$ can be solved for the 0^{th} component to give $\lambda_i = (\mathbf{P}_a^{0_i}\,\mathbf{x}^a)/\mathbf{x}^{0_i}$. Substituting back into the projection equations for the other components yields the following constraint equations for $\mathbf{x}^a$ in terms of $\mathbf{x}^{A_i}$ and $\mathbf{P}_a^{A_i}$:

$$\left(\mathbf{x}^{0_i}\,\mathbf{P}_a^{A_i} - \mathbf{x}^{A_i}\,\mathbf{P}_a^{0_i} \right) \mathbf{x}^a = 0 \qquad A_i = 1, \ldots, D_i$$

(Equivalently, $\mathbf{x}^{A_i} \sim \mathbf{P}_a^{A_i}\,\mathbf{x}^a$ implies $\mathbf{x}^{[A_i}\,\mathbf{P}_a^{B_i]}\,\mathbf{x}^a = \mathbf{0}$, and the constraint follows by setting $B_i = 0_i$). Each of these equations constrains $\mathbf{x}^a$ to lie in a hyperplane in the d-dimensional world space. Combining the constraints from all the images gives the following $D \times (d+1)$ system of **reduced reconstruction equations**:

$$\begin{pmatrix} \mathbf{x}^{0_1}\,\mathbf{P}_a^{A_1} - \mathbf{x}^{A_1}\,\mathbf{P}_a^{0_1} \\ \vdots \\ \mathbf{x}^{0_m}\,\mathbf{P}_a^{A_m} - \mathbf{x}^{A_m}\,\mathbf{P}_a^{0_m} \end{pmatrix} \mathbf{x}^a \;=\; \mathbf{0} \qquad {\scriptstyle (A_i = 1, \ldots, D_i)}$$

Again a solution of these equations provides the reconstructed homogeneous coordinates of a world point in terms of image measurements, and again the equations are usually overspecified when $D > d$. Provided $\mathbf{x}^{0_i} \neq 0$ the reduced equations are equivalent to the basic ones. Their compactness makes them attractive for numerical work, but their lack of symmetry makes them less suitable for symbolic derivations such as the extraction of the matching constraints. In practice both representations are useful.

7. Matching Constraints

Now we are finally ready to derive the constraints that a set of image points must satisfy in order to be the projections of some world point. We will assume that there are more image than space dimensions ($D > d$) (if not there are no matching constraints) and that the joint projection matrix $\mathbf{P}_a^\alpha$ has rank $d+1$ (if not there are no unique reconstructions). We will work from the basic reconstruction equations, with odd remarks on the equivalent reduced case.

In either case there are $D-d-1$ more equations than variables and the reconstruction systems are overspecified. The image points must satisfy $D-d$ additional independent constraints for there to be a solution, since one degree of freedom is lost in the overall scale factor. For example in the usual 3D→2D case there are $2m-3$ additional scalar constraints: one for the first pair of images and two more for each additional image.

An overspecified homogeneous linear system has nontrivial solutions exactly when its coefficient matrix is rank deficient, which occurs exactly when all of its maximal-size minors vanish. For generic sets of image points the reconstruction systems typically have full rank: solutions exist only for the special sets of image points for which all of the $(d+m+1)\times(d+m+1)$ minors of the basic (or $(d+1)\times(d+1)$ minors of the reduced) reconstruction matrix vanish. These minors are exactly the matching constraints.

In either case each of the minors involves all $d+1$ (world-space) columns and some selection of $d+1$ (image-space) rows of the combined projection matrices, multiplied by image coordinates. This means that the constraints will be polynomials (*i.e.* tensors) in the image coordinates with coefficients that are $(d+1)\times(d+1)$ minors of the $(D+m)\times(d+1)$ joint projection matrix $\mathbf{P}_a^\alpha$. We have already seen in section 5 that these minors are precisely the Grassmann coordinates of the *joint image* $\mathcal{I}^\alpha$, the subspace of homogeneous joint image space spanned by the $d+1$ columns of $\mathbf{P}_a^\alpha$. The complete set of these defines $\mathcal{I}^\alpha$ in a manner entirely independent (up to a scale factor) of the choice of basis in $\mathcal{I}^\alpha$: they are the only quantities that *could* have appeared if the equations were to be invariant to this choice of basis (or equivalently, to arbitrary projective transformations of the world space).

Each of the $(d+m+1)\times(d+m+1)$ minors of the basic reconstruction system contains one column from each image, and hence is linear in the coordinates of each image separately and homogeneous of degree m in the combined image coordinates. The final constraint equations will be linear in the coordinates of each image that appears in them. Any choice of $d+m+1$ of the $D+m$ rows of the matrix specifies a minor, so naively there are $\binom{D+m}{d+m+1}$ distinct constraint polynomials, although the simple degree of freedom count given above shows that even in this naive case only $D-d$ of these can be algebraically independent. However the reconstruction matrix has many zero entries and we need to count more carefully.

Each row comes from (contains components from) exactly one image. The only nonzero entries in the image i column are those from image i itself, so any minor that does not include at least one row from each image will vanish. This leaves only $d+1$ of the $m+d+1$ rows free to apportion. On the other hand, if a minor contains only one row from some image — say the $\mathbf{x}^{A_i}$ row for some particular values of i and A_i — it will simply be the product of $\pm\mathbf{x}^{A_i}$ and an $m-1$ image minor because $\mathbf{x}^{A_i}$ is the *only* nonzero entry in its image i column. But exactly the same $(m-1)$-image minor will appear in several other m-image minors, one for each other choice of the coordinate $A_i = 0,\ldots,D_i$. At least one of these coordinates is nonzero, so the vanishing of the D_i+1 m-image minors is equivalent to the vanishing of the single $(m-1)$-image one.

This allows the full set of m-image matching polynomials to be reduced to terms involving at most $d+1$ images. ($d+1$ because there are only $d+1$ spare rows to share out). In the standard 3D$\to$2D case this leaves the following possibilities ($i\neq j\neq k\neq l = 1,\ldots,m$): (*i*) 3 rows each in images i and j; (*ii*) 3 rows in image i, and 2 rows each in images j and k; and (*iii*) 2 rows each in images i, j, k and l. We will show below that these possibilities correspond respectively to fundamental matrices (*i.e.* bilinear two image constraints), Shashua's trilinear three-image constraints [19], and a new quadrilinear four-image constraint. For 3 dimensional space this is the complete list of possibilities: there are *no* irreducible k-image matching constraints for $k>4$.

We can look at all this in another way. Consider the $d+m+1$ $(D+m)$-component columns of the reconstruction system matrix. Temporarily writing $\mathbf{x}_i^\alpha$ for the image i column whose only nonzero entries are $\mathbf{x}^{A_i}$, the columns are $\{\mathbf{P}_a^\alpha \mid a=0,\ldots,d\}$ and $\{\mathbf{x}_i^\alpha \mid i=1,\ldots,m\}$ and we can form them into a $d+m+1$ index antisymmetric tensor $\mathbf{P}_0^{[\alpha_0}\cdots\mathbf{P}_d^{\alpha_d}\mathbf{x}_1^{\beta_1}\cdots\mathbf{x}_m^{\beta_m]}$. Up to scale, the components of this tensor are exactly the possible $(d+m+1)\times(d+m+1)$ minors of the system matrix. The term $\mathbf{x}_i^\alpha$ vanishes unless α is one of the components A_i, so we need at least one index from each image in the index set $\alpha_0,\ldots,\alpha_d,\beta_1,\ldots,\beta_m$. If only one component from image i is present in the set (B_i say, for some fixed value of B_i), we can extract an overall factor of $\mathbf{x}^{B_i}$ as above. Proceeding in this way the tensor can be reduced to irreducible terms of the form $\mathbf{P}_0^{[\alpha_0}\cdots\mathbf{P}_d^{\alpha_d}\mathbf{x}_i^{B_i}\mathbf{x}_j^{B_j}\cdots\mathbf{x}_k^{B_k]}$. These contain anything from 2 to $d+1$ distinct images $i,j,\ldots,k$. The indices $\alpha_0,\ldots,\alpha_d$ are an arbitrary choice of indices from images $i,j,\ldots,k$ in which each image appears at least once. Recalling that up to scale the components of the joint image Grassmannian $\mathbf{I}^{\,\alpha_0\cdots\alpha_d}$ are just $\mathbf{P}_0^{[\alpha_0}\cdots\mathbf{P}_d^{\alpha_d]}$, and dropping the redundant subscripts on the $\mathbf{x}_i^{A_i}$, we can write the final constraint equations in the compact form

$$\mathbf{I}^{\,[A_iA_j\cdots A_k\alpha\cdots\beta}\,\mathbf{x}^{B_i}\mathbf{x}^{B_j}\cdots\mathbf{x}^{B_k]} = \mathbf{0}$$

where $i, j, \ldots, k$ contains between 2 and $d+1$ distinct images. The remaining indices $\alpha \cdots \beta$ can be chosen arbitrarily from any of the images $i, j, \ldots, k$, up to the maximum of $D_i + 1$ indices from each image. (NB: the $\mathbf{x}^{B_i}$ stand for m distinct vectors whose non-i components vanish, not for the single vector $\mathbf{x}^\alpha$ containing all the image measurements. Since $\mathbf{I}^{\,\alpha_0 \cdots \alpha_d}$ is already antisymmetric and permutations that place a non-i index on $\mathbf{x}^{B_i}$ vanish, it is enough to antisymmetrize separately over the components from each image).

This is all rather intricate, but in three dimensions the possibilities are as follows ($i \neq j \neq k \neq l = 1, \ldots, m$):

$$
\begin{aligned}
\mathbf{I}^{\,[A_i B_i A_j B_j}\, \mathbf{x}^{C_i} \mathbf{x}^{C_j]} &= \mathbf{0} \\
\mathbf{I}^{\,[A_i B_i A_j A_k}\, \mathbf{x}^{C_i} \mathbf{x}^{B_j} \mathbf{x}^{B_k]} &= \mathbf{0} \\
\mathbf{I}^{\,[A_i A_j A_k A_l}\, \mathbf{x}^{B_i} \mathbf{x}^{B_j} \mathbf{x}^{B_k} \mathbf{x}^{B_l]} &= \mathbf{0}
\end{aligned}
$$

These represent respectively the epipolar constraint, Shashua's trilinear constraint and the new quadrilinear four image constraint.

We will discuss each of these possibilities in detail below, but first we take a brief look at the constraints that arise from the *reduced* reconstruction system. Each row of this system is linear in the coordinates of one image and in the corresponding rows of the joint projection matrix, so each $(d+1) \times (d+1)$ minor can be expanded into a sum of degree $d+1$ polynomial terms in the image coordinates, with $(d+1) \times (d+1)$ minors of the joint projection matrix (Grassmann coordinates of $\mathcal{PI}^\alpha$) as coefficients. Moreover, any term that contains two non-zeroth coordinates from the same image (say $A_i \neq 0$ and $B_i \neq 0$) vanishes because the row $\mathbf{P}_a^{0_i}$ appears twice in the corresponding coefficient minor. So each term is at most linear in the non-zeroth coordinates of each image. If k_i is the total number of rows from the i^{th} image in the minor, this implies that the zeroth coordinate $\mathbf{x}^{0_i}$ appears either k_i or $k_i - 1$ times in each term to make up the total homogeneity of k_i in the coordinates of the i^{th} image. Throwing away the nonzero overall factors of $(\mathbf{x}^{0_i})^{k_i - 1}$ leaves a constraint polynomial linear in the coordinates of each image and of total degree at most $d+1$, with $(d+1) \times (d+1)$ minors of the joint projection matrix as coefficients. Closer inspection shows that these are the same as the constraint polynomials found above.

7.1. Bilinear Constraints

Now we restrict attention to 2D images of a 3D world and examine each of the three constraint types in turn. First consider the bilinear joint image Grassmannian constraint $\mathbf{I}^{\,[B_1 C_1 B_2 C_2} \mathbf{x}^{A_1} \mathbf{x}^{A_2]} = \mathbf{0}$, where as usual $\mathbf{I}^{\,\alpha\beta\gamma\delta} \equiv \frac{1}{4!}\, \mathbf{P}_a^\alpha \mathbf{P}_b^\beta \mathbf{P}_c^\gamma \mathbf{P}_d^\delta\, \varepsilon^{abcd}$. Recalling that it is enough to antisymmetrize over the components from each image separately, the epipolar constraint becomes

$$
\mathbf{x}^{[A_1}\, \mathbf{I}^{\,B_1 C_1][B_2 C_2}\, \mathbf{x}^{A_2]} = \mathbf{0}
$$

Dualizing both sets of antisymmetric indices by contracting with $\varepsilon_{A_1 B_1 C_1}\, \varepsilon_{A_2 B_2 C_2}$ gives the epipolar constraint the equivalent but more familiar form

$$
\begin{aligned}
0 &= \mathbf{F}_{A_1 A_2}\, \mathbf{x}^{A_1} \mathbf{x}^{A_2} \\
&= \frac{1}{4 \cdot 4!} \left(\varepsilon_{A_1 B_1 C_1} \mathbf{x}^{A_1} \mathbf{P}_a^{B_1} \mathbf{P}_b^{C_1} \right) \cdot \\
&\qquad \cdot \left(\varepsilon_{A_2 B_2 C_2} \mathbf{x}^{A_2} \mathbf{P}_c^{B_2} \mathbf{P}_d^{C_2} \right) \varepsilon^{abcd}
\end{aligned}
$$

where the $3 \times 3 = 9$ component bilinear constraint tensor or **fundamental matrix** $\mathbf{F}_{A_1 A_2}$ is defined by

$$
\begin{aligned}
\mathbf{F}_{A_1 A_2} &\equiv \frac{1}{4}\, \varepsilon_{A_1 B_1 C_1}\, \varepsilon_{A_2 B_2 C_2}\, \mathbf{I}^{\,B_1 C_1 B_2 C_2} \\
&= \frac{1}{4 \cdot 4!} \left(\varepsilon_{A_1 B_1 C_1} \mathbf{P}_a^{B_1} \mathbf{P}_b^{C_1} \right) \cdot \\
&\qquad \cdot \left(\varepsilon_{A_2 B_2 C_2} \mathbf{P}_c^{B_2} \mathbf{P}_d^{C_2} \right) \varepsilon^{abcd} \\
\mathbf{I}^{\,B_1 C_1 B_2 C_2} &= \mathbf{F}_{A_1 A_2}\, \varepsilon^{A_1 B_1 C_1} \varepsilon^{A_2 B_2 C_2}
\end{aligned}
$$

Equivalently, the epipolar constraint can be derived by direct expansion of the 6×6 basic reconstruction system minor

$$
\mathrm{Det} \begin{pmatrix} \mathbf{P}_a^{A_1} & \mathbf{x}^{A_1} & \mathbf{0} \\ \mathbf{P}_a^{A_2} & \mathbf{0} & \mathbf{x}^{A_2} \end{pmatrix} = 0
$$

Choosing the image 1 rows and column and any two columns a and b of $\mathbf{P}$ gives a 3×3 sub-determinant $\varepsilon_{A_1 B_1 C_1} \mathbf{x}^{A_1} \mathbf{P}_a^{B_1} \mathbf{P}_b^{C_1}$. The remaining rows and columns (for image 2 and the remaining two columns c and d of $\mathbf{P}$, say) give the factor $\varepsilon_{A_2 B_2 C_2} \mathbf{x}^{A_2} \mathbf{P}_c^{B_2} \mathbf{P}_d^{C_2}$ multiplying this sub-determinant in the determinantal sum. Antisymmetrizing over the possible choices of a through d gives the above bilinear constraint equation. When there are only two images, $\mathbf{F}$ can also be written as the inter-image part of the $\mathcal{P}^\alpha$ (six dimensional) dual $\mathbf{F}_{A_1 A_2} = \frac{1}{4}\, \varepsilon_{A_1 B_1 C_1 A_2 B_2 C_2}\, \mathbf{I}^{\,B_1 C_1 B_2 C_2}$. This is why it was generated by the $6 - 4 = 2$ six dimensional constraint covectors $\mathbf{u}_\alpha$ and $\mathbf{v}_\alpha$ for $\mathcal{I}^\alpha$ in section 3.

The bilinear constraint equation

$$0 = \left(\varepsilon_{A_1B_1C_1}\mathbf{x}^{A_1}\mathbf{P}_a^{B_1}\mathbf{P}_b^{C_1}\right)\cdot$$
$$\cdot\left(\varepsilon_{A_2B_2C_2}\mathbf{x}^{A_2}\mathbf{P}_c^{B_2}\mathbf{P}_d^{C_2}\right)\,\varepsilon^{abcd}$$

can be interpreted geometrically as follows. The dualization $\varepsilon_{ABC}\ \mathbf{x}^A$ converts an image point $\mathbf{x}^A$ into covariant coordinates in the image plane. Roughly speaking, this represents the point as the pencil of lines through it: for any two lines $\mathbf{l}_A$ and $\mathbf{m}_A$ through $\mathbf{x}^A$, the tensor $\mathbf{l}_{[B}\mathbf{m}_{C]}$ is proportional to $\varepsilon_{ABC}\ \mathbf{x}^A$. Any covariant image tensor can be 'pulled back' through the linear projection $\mathbf{P}_a^A$ to a covariant tensor in 3D space. An image line $\mathbf{l}_A$ pulls back to the 3D plane $\mathbf{l}_a = \mathbf{l}_A\mathbf{P}_a^A$ through the projection centre that projects to the line. The tensor $\varepsilon_{ABC}\ \mathbf{x}^A$ pulls back to the 2 index covariant tensor $\mathbf{x}_{[bc]} \equiv \varepsilon_{ABC}\ \mathbf{x}^A\ \mathbf{P}_b^B\mathbf{P}_c^C$. This is the covariant representation of a line in 3D: the optical ray through $\mathbf{x}^A$. Given any two lines $\mathbf{x}_{[ab]}$ and $\mathbf{y}_{[ab]}$ in 3D space, the requirement that they intersect is $\mathbf{x}_{ab}\ \mathbf{y}_{cd}\ \varepsilon^{abcd} = 0$. So the above bilinear constraint equation really *is* the standard epipolar constraint, *i.e.* the requirement that the optical rays of the two image points must intersect. Similarly, the $\mathbf{F}_{A_1A_2}$ tensor really is the usual fundamental matrix. Of course this can also be illustrated by explicitly writing out terms.

7.2. Trilinear Constraints

Now consider the trilinear, three image Grassmannian constraint $\mathbf{I}^{[B_1C_1B_2B_3}\,\mathbf{x}^{A_1}\mathbf{x}^{A_2}\mathbf{x}^{A_3]} = \mathbf{0}$. This corresponds to a 7×7 basic reconstruction minor formed by selecting all three rows from the first image and two each from the remaining two. Restricting the antisymmetrization to each image and contracting with $\varepsilon_{A_1B_1C_1}$ gives the trilinear constraint

$$\mathbf{x}^{A_1}\mathbf{x}^{[A_2}\,\mathbf{G}_{A_1}{}^{B_2][B_3}\,\mathbf{x}^{A_3]} = \mathbf{0}$$

where the $3\times3\times3 = 27$ component trilinear constraint tensor $\mathbf{G}_{A_1}{}^{A_2A_3}$ is defined by

$$\mathbf{G}_{A_1}{}^{A_2A_3} \equiv \tfrac{1}{2}\,\varepsilon_{A_1B_1C_1}\,\mathbf{I}^{B_1C_1A_2A_3}$$
$$= \tfrac{1}{2\cdot 4!}\left(\varepsilon_{A_1B_1C_1}\mathbf{P}_a^{B_1}\mathbf{P}_b^{C_1}\right)\mathbf{P}_c^{A_2}\,\mathbf{P}_d^{A_3}\,\varepsilon^{abcd}$$
$$\mathbf{I}^{A_1B_1A_2A_3} = \mathbf{G}_{C_1}{}^{A_2A_3}\,\varepsilon^{C_1A_1B_1}$$

Dualizing the image 2 and 3 indices by contracting with $\varepsilon_{A_2B_2C_2}\ \varepsilon_{A_3B_3C_3}$ gives the constraint the alternative form

$$\mathbf{0} = \varepsilon_{A_2B_2C_2}\,\varepsilon_{A_3B_3C_3}\cdot\mathbf{G}_{A_1}{}^{B_2B_3}\cdot\mathbf{x}^{A_1}\mathbf{x}^{A_2}\mathbf{x}^{A_3}$$
$$= \tfrac{1}{2.4!}\left(\varepsilon_{A_1B_1C_1}\mathbf{x}^{A_1}\mathbf{P}_a^{B_1}\mathbf{P}_b^{C_1}\right)\cdot$$
$$\cdot\left(\varepsilon_{A_2B_2C_2}\mathbf{x}^{A_2}\mathbf{P}_c^{B_2}\right)\left(\varepsilon_{A_3B_3C_3}\mathbf{x}^{A_3}\mathbf{P}_d^{B_3}\right)\,\varepsilon^{abcd}$$

These equations must hold for all $3\times 3 = 9$ values of the free indices C_2 and C_3. However when C_2 is projected along the $\mathbf{x}^{C_2}$ direction or C_3 is projected along the $\mathbf{x}^{C_3}$ direction the equations are tautological because, for example, $\varepsilon_{A_2B_2C_2}\ \mathbf{x}^{A_2}\mathbf{x}^{C_2} \equiv 0$. So there are actually only $2\times 2 = 4$ linearly independent scalar constraints among the $3\times 3 = 9$ equations, corresponding to the two image 2 directions 'orthogonal' to $\mathbf{x}^{A_2}$ and the two image 3 directions 'orthogonal' to $\mathbf{x}^{A_3}$. However, each of the $3\times 3 = 9$ constraint equations and $3^3 = 27$ components of the constraint tensor are 'activated' for *some* $\mathbf{x}^{A_i}$, so none can be discarded outright.

The constraint can also be written in matrix notation as follows (*c.f.* [19]). The contraction $\mathbf{x}^{A_1}\mathbf{G}_{A_1}{}^{A_2A_3}$ has free indices A_2A_3 and can be viewed as a 3×3 matrix $[\mathbf{G}\,\mathbf{x}_1]$, and the fragments $\varepsilon_{A_2B_2C_2}\ \mathbf{x}^{A_2}$ and $\varepsilon_{A_3B_3C_3}\ \mathbf{x}^{A_3}$ can be viewed as 3×3 antisymmetric 'cross product' matrices $[\mathbf{x}_2]_\times$ and $[\mathbf{x}_3]_\times$ (where $\mathbf{x}\times\mathbf{y} = [\mathbf{x}]_\times\,\mathbf{y}$ for any 3-vector $\mathbf{y}$). The constraint is then given by the 3×3 matrix equation

$$[\mathbf{x}_2]_\times\,[\mathbf{G}\,\mathbf{x}_1]\,[\mathbf{x}_3]_\times = \mathbf{0}_{\{3\times3\}}$$

The projections along $\mathbf{x}_2^\top$ (on the left) and $\mathbf{x}_3$ (on the right) vanish identically, so again there are only 4 linearly independent equations.

The trilinear constraint formula

$$\mathbf{x}^{A_1}\mathbf{x}^{[A_2}\,\mathbf{G}_{A_1}{}^{B_2][B_3}\,\mathbf{x}^{A_3]} = \mathbf{0}$$

also implies that for all values of the free indices $[A_2B_2]$ (or dually C_2)

$$\mathbf{x}^{A_3} \sim \mathbf{x}^{A_1}\mathbf{x}^{[A_2}\,\mathbf{G}_{A_1}{}^{B_2]A_3}$$
$$\sim \varepsilon_{C_2A_2B_2}\,\mathbf{x}^{A_1}\mathbf{x}^{A_2}\,\mathbf{G}_{A_1}{}^{B_2A_3}$$

More precisely, for *matching* $\mathbf{x}^{A_1}$ and $\mathbf{x}^{A_2}$ the quantity $\mathbf{x}^{A_1}\mathbf{x}^{[A_2}\,\mathbf{G}_{A_1}{}^{B_2]A_3}$ can always be factorized as $\mathbf{T}^{[A_2B_2]}\,\mathbf{x}^{A_3}$ for some $\mathbf{x}^{A_i}$-dependent tensor $\mathbf{T}^{[A_2B_2]}$ (and similarly with $\mathbf{T}_{C_2}$ for the dual form). By fixing suitable values of $[A_2B_2]$ or C_2, these equations can be used to **transfer** points from images 1 and 2 to image 3, *i.e.* to directly predict the projection in image 3 of a 3D point whose projections in images 1 and 2 are known, without any intermediate 3D reconstruction step[2].

The trilinear constraints can be interpreted geometrically as follows. As above the quantity $\varepsilon_{ABC}\,\mathbf{x}^A\,\mathbf{P}_b^B\mathbf{P}_c^C$ represents the optical ray through $\mathbf{x}^A$ in covariant 3D coordinates. For any $\mathbf{y}^A \in \mathcal{P}^A$ the quantity $\varepsilon_{ABC}\,\mathbf{x}^A\mathbf{y}^B\mathbf{P}_c^C$ defines the 3D plane through the optical centre that projects to the image line through $\mathbf{x}^A$ and $\mathbf{y}^A$. All such planes contain the optical ray of $\mathbf{x}^A$, and as $\mathbf{y}^A$ varies the entire pencil of planes through this line is traced out. The constraint then says that for any plane through the optical ray of $\mathbf{x}^{A_2}$ and any other plane through the optical ray of $\mathbf{x}^{A_3}$, the 3D line of intersection of these planes meets the optical ray of $\mathbf{x}^{A_1}$.

The line of intersection always meets the optical rays of both $\mathbf{x}^{A_2}$ and $\mathbf{x}^{A_3}$ because it lies in planes containing those rays. If the rays are skew *every* line through the two rays is generated as the planes vary. The optical ray through $\mathbf{x}^{A_1}$ can not meet every such line, so the constraint implies that the optical rays of $\mathbf{x}^{A_2}$ and $\mathbf{x}^{A_3}$ can not be skew. In other words the image 1 trilinear constraint implies the epipolar constraint between images 2 and 3.

Given that the rays of $\mathbf{x}^{A_2}$ and $\mathbf{x}^{A_3}$ meet (say, at some point $\mathbf{x}^a$), as the two planes through these rays vary their intersection traces out every line through $\mathbf{x}^a$ not in the plane of the rays. The only way that the optical ray of $\mathbf{x}^{A_1}$ can arrange to meet each of these lines is for it to pass through $\mathbf{x}^a$ as well. In other words the trilinear constraint for each image implies that all three optical rays pass through the same point. Thus, the epipolar constraints between images 1 and 2 and images 1 and 3 also follow from the image 1 trilinear constraint.

The constraint tensor $\mathbf{G}_{A_1}{}^{A_2A_3} \equiv \varepsilon_{A_1B_1C_1}\,\mathbf{I}^{B_1C_1A_2A_3}$ treats image 1 specially. The analogous image 2 and image 3 tensors $\mathbf{G}_{A_2}{}^{A_3A_1} \equiv \varepsilon_{A_2B_2C_2}\,\mathbf{I}^{B_2C_2A_3A_1}$ and $\mathbf{G}_{A_3}{}^{A_1A_2} \equiv \varepsilon_{A_3B_3C_3}\,\mathbf{I}^{B_3C_3A_1A_2}$ are linearly independent of $\mathbf{G}_{A_1}{}^{A_2A_3}$ and give further linearly independent trilinear constraints on $\mathbf{x}^{A_1}\mathbf{x}^{A_2}\mathbf{x}^{A_3}$. Together, the 3 homogeneous constraint tensors contain $3 \times 27 = 81$ linearly independent components (including 3 arbitrary scale factors) and naïvely give $3 \times 9 = 27$ trilinear scalar constraint equations, of which $3 \times 4 = 12$ are linearly independent for any given triple $\mathbf{x}^{A_1}\mathbf{x}^{A_2}\mathbf{x}^{A_3}$.

However, although there are no *linear* relations between the $3 \times 27 = 81$ trilinear and $3 \times 9 = 27$ bilinear matching tensor components for the three images, the matching tensors are certainly not *algebraically* independent of each other: there are many *quadratic* relations between them inherited from the quadratic simplicity constraints on the joint image Grassmannian tensor. In fact, we saw in section 5 that the simplicity constraints reduce the number of algebraically independent degrees of freedom of $\mathbf{I}^{\alpha_0\cdots\alpha_3}$ (and therefore the complete set of bilinear and trilinear matching tensor components) to only $11m - 15 = 18$ for $m = 3$ images. Similarly, there are only $2m - 3 = 3$ *algebraically* independent scalar constraint equations among the *linearly* independent $3 \times 4 = 12$ trilinear and $3 \times 1 = 3$ bilinear constraints on each matching triple of points. One of the main advantages of the Grassmann formalism is the extent to which it clarifies the rich algebraic structure of this matching constraint system. The components of the constraint tensors are essentially just Grassmann coordinates of the joint image, and Grassmann coordinates are *always* linearly independent and quadratically redundant.

Since all three of the epipolar constraints follow from a single trilinear tensor it may seem that the trilinear constraint is more powerful than the epipolar ones, but this is not really so. Given a triple of image points $\{\mathbf{x}^{A_i} \,|\, i = 1, \ldots, 3\}$, the three pairwise epipolar constraints say that the three optical rays must meet pairwise. If they do not meet at a single point, this implies that each ray must lie in the plane of the other two. Since the rays pass through their respective optical centres, the plane also contains the three optical centres, and is therefore the **trifocal plane**. But this is impossible in general: most image points simply do not lie on the trifocal lines (the projections of the trifocal planes). So for general matching image points the three epipolar constraints together imply that the three optical rays meet at a unique 3D point. This is enough to imply the trilinear constraints. Since we know that only $2m - 3 = 3$ of the constraints are algebraically independent, this is as expected.

Similarly, the information contained in just one of the trilinear constraint tensors is generically $4 > 2m - 3 = 3$ linearly independent constraints, which is enough to imply the other two trilinear tensors as well as the three bilinear ones. This explains why most of the early work on trilinear constraints successfully ignores two of the three available tensors [19], [8]. However in the context of purely *linear* reconstruction all three of the tensors would be necessary.

7.3. Quadrilinear Constraints

Finally, the quadrilinear, four image Grassmannian constraint $\mathbf{I}^{[B_1B_2B_3B_4}\,\mathbf{x}^{A_1}\mathbf{x}^{A_2}\mathbf{x}^{A_3}\mathbf{x}^{A_4]} = \mathbf{0}$ corresponds to an 8×8 basic reconstruction minor that selects two rows from each of four images. As usual the antisymmetrization applies to each image separately, but in this case the simplest form of the constraint tensor is just a direct selection of $3^4 = 81$ components of the Grassmannian itself

$$\begin{aligned}\mathbf{H}^{A_1A_2A_3A_4} &\equiv \mathbf{I}^{A_1A_2A_3A_4} \\ &= \tfrac{1}{4!}\,\mathbf{P}_a^{A_1}\mathbf{P}_b^{A_2}\mathbf{P}_c^{A_3}\mathbf{P}_d^{A_4}\,\varepsilon^{abcd}\end{aligned}$$

Dualizing the antisymmetric index pairs $[A_iB_i]$ by contracting with $\varepsilon_{A_iB_iC_i}$ for $i = 1,\ldots,4$ gives the quadrilinear constraint

$$\begin{aligned}\mathbf{0} &= \varepsilon_{A_1B_1C_1}\,\varepsilon_{A_2B_2C_2}\,\varepsilon_{A_3B_3C_3}\,\varepsilon_{A_4B_4C_4}\,\cdot \\ &\qquad\qquad \cdot\,\mathbf{x}^{A_1}\mathbf{x}^{A_2}\mathbf{x}^{A_3}\mathbf{x}^{A_4}\,\mathbf{H}^{B_1B_2B_3B_4} \\ &= \tfrac{1}{4!}\left(\varepsilon_{A_1B_1C_1}\mathbf{x}^{A_1}\mathbf{P}_a^{B_1}\right)\left(\varepsilon_{A_2B_2C_2}\mathbf{x}^{A_2}\mathbf{P}_b^{B_2}\right)\cdot \\ &\quad \cdot\left(\varepsilon_{A_3B_3C_3}\mathbf{x}^{A_3}\mathbf{P}_c^{B_3}\right)\left(\varepsilon_{A_4B_4C_4}\mathbf{x}^{A_4}\mathbf{P}_d^{B_4}\right)\,\varepsilon^{abcd}\end{aligned}$$

This must hold for each of the $3^4 = 81$ values of $C_1C_2C_3C_4$. But again the constraints with C_i along the direction $\mathbf{x}^{C_i}$ for any $i = 1,\ldots,4$ vanish identically, so for any given quadruple of points there are only $2^4 = 16$ linearly independent constraints among the $3^4 = 81$ equations.

Together, these constraints say that for every possible choice of four planes, one through the optical ray defined by $\mathbf{x}^{A_i}$ for each $i = 1,\ldots,4$, the planes meet in a point. By fixing three of the planes and varying the fourth we immediately find that each of the optical rays passes through the point, and hence that they all meet. This brings us back to the two and three image sub-cases.

Again, there is nothing algebraically new here. The $3^4 = 81$ homogeneous components of the quadrilinear constraint tensor are *linearly* independent of each other and of the $4\times 3\times 27 = 324$ homogeneous trilinear and $6\times 9 = 54$ homogeneous bilinear tensor components; and the $2^4 = 16$ linearly independent quadrilinear scalar constraints are *linearly* independent of each other and of the linearly independent $4\times 3\times 4 = 48$ trilinear and $6\times 1 = 6$ bilinear constraints. However there are only $11m - 15 = 29$ *algebraically* independent tensor components in total, which give $2m - 3 = 5$ *algebraically* independent constraints on each 4-tuple of points. The quadrilinear constraint is algebraically equivalent to various different combinations of two and three image constraints. For example five scalar epipolar constraints will do: take the three pairwise constraints for the first three images, then add two of the three involving the fourth image to force the optical rays from the fourth image to pass through the intersection of the corresponding optical rays from the other three images.

7.4. Matching Constraints for Lines

It is well known that there is no matching constraint for lines in two images. Any two non-epipolar image lines $\mathbf{l}_{A_1}$ and $\mathbf{l}_{A_2}$ are the projection of some unique 3D line: simply pull back the image lines to two 3D planes $\mathbf{l}_{A_1}\mathbf{P}_a^{A_1}$ and $\mathbf{l}_{A_2}\mathbf{P}_a^{A_2}$ through the centres of projection and intersect the planes to find the 3D line $\mathbf{l}_{ab} = \mathbf{l}_{A_1}\mathbf{l}_{A_2}\,\mathbf{P}_{[a}^{A_1}\mathbf{P}_{b]}^{A_2}$.

However for three or more images of a line there are trilinear matching constraints as follows [8]. An image line is the projection of a 3D line if and only if each point on the 3D line projects to a point on the image line. Writing this out, we immediately see that the lines $\{\mathbf{l}_{A_i}\,|\,i = 1,\ldots,m\}$ correspond to a 3D line if and only if the $m\times 4$ reconstruction equations

$$\begin{pmatrix}\mathbf{l}_{A_1}\mathbf{P}_a^{A_1}\\ \vdots \\ \mathbf{l}_{A_m}\mathbf{P}_a^{A_m}\end{pmatrix}\mathbf{x}^a = \mathbf{0}$$

have a line (*i.e.* a 2D linear space) of solutions $\lambda\mathbf{x}^a + \mu\mathbf{y}^a$ for some solutions $\mathbf{x}^a \not\sim \mathbf{y}^a$.

There is a 2D solution space if and only if the coefficient matrix has rank $4 - 2 = 2$, which means that every 3×3 minor has to vanish. Obviously each minor is a trilinear function in three $\mathbf{l}_{A_i}$'s and misses out one of the columns of $\mathbf{P}_a^{\alpha}$. Labelling the missing column as a and expanding produces constraint equations like

$$\mathbf{l}_{A_1}\,\mathbf{l}_{A_2}\,\mathbf{l}_{A_3}\,\left(\mathbf{P}_b^{A_1}\mathbf{P}_c^{A_2}\mathbf{P}_d^{A_3}\,\varepsilon^{abcd}\right) = \mathbf{0}$$

These simply require that the three pulled back planes $\mathbf{l}_{A_1}\mathbf{P}_a^{A_1}$, $\mathbf{l}_{A_2}\mathbf{P}_a^{A_2}$ and $\mathbf{l}_{A_3}\mathbf{P}_a^{A_3}$ meet in some common 3D line, rather than just a single point. Note the geometry here: each line $\mathbf{l}_{A_i}$ pulls back to a hyperplane in $\mathcal{P}^{\alpha}$ under the trivial projection. This restricts to a hyperplane in $\mathcal{PI}^{\alpha}$, which can be expressed as $\mathbf{l}_{A_i}\mathbf{P}_a^{A_i}$ in the basis $\mathbf{P}_a^{\alpha}$ for $\mathcal{PI}^{\alpha}$. There are $2m - 4$ algebraically independent constraints for m images: two for each image except the first two. There are *no* irreducible higher order constraints for lines in more than 3 images, *e.g.* there is no analogue of the quadrilinear constraint for lines.

By contracting with a final $\mathbf{P}_a^\alpha$, the constraints can also be written in terms of the Grassmannian tensor as

$$\mathbf{l}_{A_1}\,\mathbf{l}_{A_2}\,\mathbf{l}_{A_3}\,\mathbf{I}^{\,\alpha A_1 A_2 A_3} \;=\; \mathbf{0}$$

for all α. Choosing α from images 1, 2 or 3 and contracting with an image 1, 2 or 3 epsilon to produce a trivalent tensor $\mathbf{G}_{A_i}{}^{A_j A_k}$, or choosing α from a fourth image and substituting the quadrivalent tensor $\mathbf{H}^{A_i A_j A_k A_l}$ reduces the line constraints to the form

$$\begin{aligned} \mathbf{l}_{A_2}\,\mathbf{l}_{A_3}\,\mathbf{l}_{[A_1}\,\mathbf{G}_{B_1]}{}^{A_2 A_3} \;&=\; \mathbf{0} \\ \mathbf{l}_{A_1}\,\mathbf{l}_{A_2}\,\mathbf{l}_{A_3}\,\mathbf{H}^{A_1 A_2 A_3 A_4} \;&=\; \mathbf{0} \end{aligned}$$

These formulae illustrate and extend Hartley's observation that the coefficient tensors of the three-image line constraints are equivalent to those of the trilinear point constraints [8]. Note that although all of these line constraints are *tri*linear, some of them do involve *quadri*valent point constraint tensors.

Since α can take any of $3m$ values A_i, for each triple of lines and $m \geq 3$ images there are very naïvely $3m$ trilinear constraints of the above two forms. However all of these constraints are derived by linearly contracting 4 underlying world constraints with $\mathbf{P}_a^\alpha$'s, so at most 4 of them can be linearly independent. For m matching images of lines this leaves $4\binom{m}{3}$ *linearly* independent constraints of which only $2m-4$ are algebraically independent.

The skew symmetrization in the trivalent tensor based constraint immediately implies the **line transfer** equation

$$\mathbf{l}_{A_1} \;\sim\; \mathbf{l}_{A_2}\,\mathbf{l}_{A_3}\,\mathbf{G}_{A_1}{}^{A_2 A_3}$$

This can be used to predict the projection of a 3D line in image 1 given its projections in images 2 and 3, without intermediate 3D reconstruction. Note that line transfer from images 1 and 2 to image 3 is most simply expressed in terms of the image 3 trilinear tensor $\mathbf{G}_{A_3}{}^{A_1 A_2}$, whereas the image 1 or image 2 tensors $\mathbf{G}_{A_1}{}^{A_2 A_3}$ or $\mathbf{G}_{A_2}{}^{A_1 A_3}$ are the preferred form for point transfer.

It is also possible to match *(i)* points against lines that contain them and *(ii)* distinct image lines that are known to intersect in 3D. Such constraints might be useful if a polyhedron vertex is obscured or poorly localized. They are most easily derived by noting that both the line reconstruction equations and the reduced point reconstruction equations are homogeneous in $\mathbf{x}^a$, the coordinates of the intersection point. So line and point rows from several images can be stacked into a single 4 column matrix. As usual there is a solution exactly when all 4×4 minors vanish. This yields two particularly simple irreducible constraints — and correspondingly simple interpretations of the matching tensors' content — for an image point against two lines containing it and four non-corresponding image lines that intersect in 3D:

$$\begin{aligned} \mathbf{x}^{A_1}\,\mathbf{G}_{A_1}{}^{A_2 A_3}\,\mathbf{l}_{A_2}\mathbf{l}'_{A_3} \;&=\; 0 \\ \mathbf{H}^{A_1 A_2 A_3 A_4}\,\mathbf{l}_{A_1}\mathbf{l}'_{A_2}\mathbf{l}''_{A_3}\mathbf{l}'''_{A_4} \;&=\; 0 \end{aligned}$$

7.5. *Matching Constraints for k-Subspaces*

More generally, the projections of a k dimensional subspace in d dimensions are (generically) k dimensional image subspaces that can be written as antisymmetric $D_i - k$ index Grassmann tensors $\mathbf{x}_{A_i \cdots B_i \cdots C_i}$. The matching constraints can be built by selecting any $d+1-k$ of these covariant indices from any set $i, j, \ldots, k$ of image tensors and contracting with the Grassmannian to leave k free indices:

$$\begin{aligned} \mathbf{0} \;=\; \mathbf{x}_{A_i \cdots B_i C_i \cdots E_i} \;\cdots\; \mathbf{x}_{A_k \cdots B_k C_k \cdots E_k} \,\cdot \\ \cdot\; \mathbf{I}^{\,\alpha_1 \cdots \alpha_k A_i \cdots B_i \cdots A_k \cdots B_k} \end{aligned}$$

Dualizing each covariant Grassmann tensor gives an equivalent contravariant form of the constraint, for image subspaces $\mathbf{x}^{A_j \cdots E_j}$ defined by the span of a set of image points

$$\mathbf{0} \;=\; \mathbf{I}^{\,\alpha_1 \cdots \alpha_k [A_i \cdots B_i \cdots A_k \cdots B_k}\,\mathbf{x}^{C_i \cdots E_i}\;\cdots\;\mathbf{x}^{C_k \cdots E_k]}$$

As usual it is enough to antisymmetrize over the indices from each image separately. Each set $A_j \cdots B_j C_j \cdots E_j$ is any choice of up to $D_j + 1$ indices from image j, $j = i, \ldots, k$.

7.6. *2D Matching Constraints & Homographies*

Our formalism also works for 2D projective images of a 2D space. This case is practically important because it applies to 2D images of a planar surface in 3D and there are many useful plane-based vision algorithms.

The joint image of a 2D source space is two dimensional, so the corresponding Grassmannian tensor has only three indices and there are only two distinct types of matching constraint: bilinear and trilinear. Let indices a and A_i represent 3D space and the i^{th} image as usual, and indices $A = 0, 1, 2$ represent homogeneous coordinates on the source plane. If the plane is given by $\mathbf{p}_a \mathbf{x}^a = 0$, the three index epsilon tensor on it is proportional to $\mathbf{p}_a \varepsilon^{abcd}$ when expressed in world coordinates, so the Grassmann tensor becomes

$$\begin{aligned} \mathbf{I}^{\alpha\beta\gamma} &\equiv \tfrac{1}{3!}\, \mathbf{P}^\alpha_A\, \mathbf{P}^\beta_B\, \mathbf{P}^\gamma_C\, \varepsilon^{ABC} \\ &\sim \tfrac{1}{4!}\, \mathbf{p}_a\, \mathbf{P}^\alpha_b\, \mathbf{P}^\beta_c\, \mathbf{P}^\gamma_d\, \varepsilon^{abcd} \end{aligned}$$

This yields the following bilinear and trilinear matching constraints with free indices respectively C_2 and $C_1 C_2 C_3$

$$\begin{aligned} \mathbf{0} \;=\; \mathbf{p}_a \,&\Big(\varepsilon_{A_1 B_1 C_1}\, \mathbf{x}^{A_1}\, \mathbf{P}^{B_1}_b \mathbf{P}^{C_1}_c\Big) \cdot \\ &\cdot \Big(\varepsilon_{A_2 B_2 C_2}\, \mathbf{x}^{A_2}\, \mathbf{P}^{B_2}_d\Big)\, \varepsilon^{abcd} \\ \mathbf{0} \;=\; \mathbf{p}_a \,&\Big(\varepsilon_{A_1 B_1 C_1}\, \mathbf{x}^{A_1}\, \mathbf{P}^{B_1}_b\Big) \Big(\varepsilon_{A_2 B_2 C_2}\, \mathbf{x}^{A_2}\, \mathbf{P}^{B_2}_c\Big) \cdot \\ &\cdot \Big(\varepsilon_{A_3 B_3 C_3}\, \mathbf{x}^{A_3}\, \mathbf{P}^{B_3}_d\Big)\, \varepsilon^{abcd} \end{aligned}$$

The bilinear equation says that $\mathbf{x}^{A_2}$ is the image of the intersection of optical ray of $\mathbf{x}^{A_1}$ with the plane $\mathbf{p}_a$: $\mathbf{x}^{A_2} \sim \Big(\mathbf{p}_a \cdot \varepsilon_{A_1 B_1 C_1}\, \mathbf{P}^{B_1}_b \mathbf{P}^{C_1}_c \cdot \mathbf{P}^{A_2}_d \cdot \varepsilon^{abcd}\Big)\, \mathbf{x}^{A_1}$. In fact it is well known that any two images of a plane are projectively equivalent under a transformation (homography) $\mathbf{x}^{A_2} \sim \mathbf{H}^{A_2}_{A_1}\, \mathbf{x}^{A_1}$. In our notation the homography is just

$$\mathbf{H}^{A_2}_{A_1} \;\equiv\; \mathbf{p}_a \cdot \varepsilon_{A_1 B_1 C_1}\, \mathbf{P}^{B_1}_b \mathbf{P}^{C_1}_c \cdot \mathbf{P}^{A_2}_d \cdot \varepsilon^{abcd}$$

The trilinear constraint says that any three image lines through the three image points $\mathbf{x}^{A_1}$, $\mathbf{x}^{A_2}$ and $\mathbf{x}^{A_3}$ always meet in a point when pulled back to the plane $\mathbf{p}_a$. This implies that the optical rays of the three points intersect at a common point on the plane, and hence gives the obvious cyclic consistency condition $\mathbf{H}^{A_1}_{A_2}\, \mathbf{H}^{A_2}_{A_3} \sim \mathbf{H}^{A_1}_{A_3}$ (or equivalently $\mathbf{H}^{A_1}_{A_2}\, \mathbf{H}^{A_2}_{A_3}\, \mathbf{H}^{A_3}_{B_1} \sim \boldsymbol{\delta}^{A_1}_{B_1}$) between the three homographies.

7.7. Matching Constraints for 1D Cameras

If some of the images are taken with one dimensional 'linear' cameras, a similar analysis applies but the corresponding entries in the reconstruction equations have only two rows instead of three. Constraints that would require three rows from a 1D image no longer exist, and the remaining constraints lose their free indices. In particular, when all of the cameras are 1D there are no bilinear or trilinear tensors and the only irreducible matching constraint is the quadrilinear scalar:

$$\begin{aligned} 0 &= \mathbf{H}_{A_1 A_2 A_3 A_4}\, \mathbf{x}^{A_1} \mathbf{x}^{A_2} \mathbf{x}^{A_3} \mathbf{x}^{A_4} \\ &= \Big(\varepsilon_{A_1 B_1}\, \mathbf{x}^{A_1}\, \mathbf{P}^{B_1}_a\Big) \Big(\varepsilon_{A_2 B_2}\, \mathbf{x}^{A_2}\, \mathbf{P}^{B_2}_b\Big) \cdot \\ &\quad \cdot \Big(\varepsilon_{A_3 B_3}\, \mathbf{x}^{A_3}\, \mathbf{P}^{B_3}_c\Big) \Big(\varepsilon_{A_4 B_4}\, \mathbf{x}^{A_4}\, \mathbf{P}^{B_4}_d\Big)\, \varepsilon^{abcd} \end{aligned}$$

This says that the four planes pulled back from the four image points must meet in a 3D point. If one of the cameras is 2D and the other two are 1D a scalar trilinear constraint also exists.

7.8. 3D to 2D Matching

It is also useful to be able to match known 3D structure to 2D image structure, for example when building a reconstruction incrementally from a sequence of images. This case is rather trivial as the 'constraint tensor' is just the projection matrix, but for comparison it is perhaps worth writing down the equations. For an image point $\mathbf{x}^A$ projected from a world point $\mathbf{x}^a$ we have $\mathbf{x}^A \sim \mathbf{P}^A_a\, \mathbf{x}^a$ and hence the equivalent constraints

$$\mathbf{x}^{[A}\, \mathbf{P}^{B]}_a\, \mathbf{x}^a = \mathbf{0} \quad \Longleftrightarrow \quad \varepsilon_{ABC}\, \mathbf{x}^A\, \mathbf{P}^B_a\, \mathbf{x}^a = \mathbf{0}$$

There are three bilinear equations, only two of which are independent for any given image point. Similarly, a world line $\mathbf{l}_{[ab]}$ (or dually, $\mathbf{l}^{[ab]}$) and a corresponding image line $\mathbf{l}_A$ satisfy the equivalent bilinear constraints

$$\mathbf{l}_A\, \mathbf{P}^A_{[a}\, \mathbf{l}_{bc]} = \mathbf{0} \quad \Longleftrightarrow \quad \mathbf{l}_A\, \mathbf{P}^A_a\, \mathbf{l}_{bc}\, \varepsilon^{abcd} = \mathbf{0}$$

or dually

$$\mathbf{l}_A\, \mathbf{P}^A_a\, \mathbf{l}^{ab} \;=\; \mathbf{0}$$

Each form contains four bilinear equations, only two of which are linearly independent for any given image line. For example, if the line is specified by giving two points on it $\mathbf{l}^{ab} \sim \mathbf{x}^{[a} \mathbf{y}^{b]}$, we have the two scalar equations $\mathbf{l}_A\, \mathbf{P}^A_a\, \mathbf{x}^a \;=\; 0$ and $\mathbf{l}_A\, \mathbf{P}^A_a\, \mathbf{y}^a \;=\; 0$.

7.9. Epipoles

There is still one aspect of $\mathbf{I}^{\alpha_0 \cdots \alpha_d}$ that we have not yet seen: the Grassmannian tensor also directly contains the *epipoles*. In fact, the epipoles are most naturally viewed as the first order term in the sequence of matching tensors, although they do not themselves induce any matching constraints.

Assuming that it has rank d, the $d \times (d+1)$ projection matrix of a $d-1$ dimensional image of d dimensional space defines a unique **centre of projection** $\mathbf{e}_i{}^a$ by $\mathbf{P}_a^{A_i}\, \mathbf{e}_i{}^a = \mathbf{0}$. The solution of this equation is given (*c.f.* section 8) by the vector of $d \times d$ minors of $\mathbf{P}_a^{A_i}$, *i.e.*

$$\mathbf{e}_i{}^a \;\sim\; \varepsilon_{A_i \cdots C_i}\, \mathbf{P}_{a_1}^{A_i} \cdots \mathbf{P}_{a_d}^{C_i}\, \varepsilon^{a a_1 \cdots a_d}$$

The projection of a centre of projection in another image is an **epipole**

$$\mathbf{e}_i{}^{A_j} \;\sim\; \varepsilon_{A_i \cdots C_i}\, \mathbf{P}_{a_0}^{A_j} \mathbf{P}_{a_1}^{A_i} \cdots \mathbf{P}_{a_d}^{C_i}\, \varepsilon^{a_0 a_1 \cdots a_d}$$

Recognizing the factor of $\mathbf{I}^{A_j A_i B_i \cdots C_i}$, we can fix the scale factors for the epipoles so that

$$\begin{aligned} \mathbf{e}_i{}^{A_j} &\equiv \tfrac{1}{d!}\, \varepsilon_{A_i B_i \cdots C_i}\, \mathbf{I}^{A_j A_i B_i \cdots C_i} \\ \mathbf{I}^{A_j A_i B_i \cdots C_i} &= \mathbf{e}_i{}^{A_j}\, \varepsilon^{A_i B_i \cdots C_i} \end{aligned}$$

The d-dimensional joint image subspace $\mathcal{PI}^\alpha$ of $\mathcal{P}^\alpha$ passes through the d-*co*dimensional projective subspace $\mathbf{x}^{A_i} = \mathbf{0}$ at the **joint image epipole**

$$\mathbf{e}_i{}^\alpha \;\equiv\; \left(\mathbf{e}_i{}^{A_1}, \ldots, \mathbf{e}_i{}^{A_{i-1}}, \mathbf{0}, \mathbf{e}_i{}^{A_{i+1}}, \ldots, \mathbf{e}_i{}^{A_m} \right)^\top$$

As usual, an arbitrary choice of the relative scale factors is required.

Counting up the components of the $\binom{m}{4}$ quadrilinear, $3\binom{m}{3}$ trilinear, $\binom{m}{2}$ bilinear and $m(m-1)$ monolinear (epipole) tensors for m images of a 3D world, we find a total of

$$\binom{3m}{4} \;=\; 81 \cdot \binom{m}{4} \;+\; 27 \cdot 3\binom{m}{3} \;+\; 9 \cdot \binom{m}{2} \;+\; 3 \cdot m(m-1)$$

linearly independent components. These are linearly equivalent to the complete set of $\binom{3m}{4}$ linearly independent components of $\mathbf{I}^{\alpha_0 \cdots \alpha_d}$, so the joint image Grassmannian tensor can be reconstructed *linearly* given the entire set of (appropriately scaled) matching tensors.

8. Minimal Reconstructions and Uniqueness

The matching constraints found above are closely associated with a set of **minimal reconstruction techniques** that produce candidate solutions $\mathbf{x}^a$ from minimal sets of d image measurements (three in the 3D case). Geometrically, measuring an image coordinate restricts the corresponding world point to a hyperplane in $\mathcal{P}^a$. The intersection of any d independent hyperplanes gives a unique solution candidate $\mathbf{x}^a$, so there is a minimal reconstruction technique based on any set of d independent image measurements. Matching is equivalent to the requirement that this candidate lies in the hyperplane of each of the remaining measurements. If d measurements are not independent the corresponding minimal reconstruction technique will fail to give a unique candidate, but so long as the images contain *some* set of d independent measurements at least one of the minimal reconstructions will succeed and the overall reconstruction solution will be unique (or fail to exist altogether if the matching constraints are violated).

Algebraically, we can restate this as follows. Consider a general $k \times (k+1)$ system of homogeneous linear equations with rank k. Up to scale the system has a unique solution given by the $(k+1)$-component vector of $k \times k$ minors of the system matrix[3]. Adding an extra row to the system destroys the solution unless the new row is orthogonal to the existing minor vector: this is exactly the requirement that the determinant of the $(k+1) \times (k+1)$ matrix vanish so that the system still has rank k. With an overspecified rank k system: any choice of k rows gives a minor vector; at least one minor vector is nonzero by rank-k-ness; every minor vector is orthogonal to every row of the system matrix by non-rank-$(k+1)$-ness; and all of the minor vectors are equal up to scale because there is only one direction orthogonal to any given k independent rows. In other words the existence of a solution can be expressed as a set of simple orthogonality relations on a candidate solution (minor vector) produced from any set of k independent rows.

We can apply this to the $(d+m) \times (d+m)$ minors of the $(D+m) \times (d+m+1)$ basic reconstruction system, or equivalently to the $d \times d$ minors of the $D \times (d+1)$ reduced reconstruction system. The situation is very similar to that for matching constraints and a similar analysis applies. The result is that if $i, j, \ldots, k$ is a set of $2 \leq m' \leq d$ distinct images and $\gamma, \ldots, \delta$ is any selection of $d - m'$ indices from images $i, j, \ldots, k$ (at most $D_i - 1$ from any one image), there is a pair of equivalent minimal reconstruction techniques for $\mathbf{x}^a \in \mathcal{P}^a$ and $\mathbf{x}^\alpha \in \mathcal{P}^\alpha$:

$$\begin{aligned} \mathbf{x}^a &\sim \mathbf{P}^{a[B_i B_j \cdots B_k \gamma \cdots \delta}\, \mathbf{x}^{A_i} \mathbf{x}^{A_j} \ldots \mathbf{x}^{A_k]} \\ \mathbf{x}^\alpha &\sim \mathbf{I}^{\,\alpha[B_i B_j \cdots B_k \gamma \cdots \delta}\, \mathbf{x}^{A_i} \mathbf{x}^{A_j} \ldots \mathbf{x}^{A_k]} \end{aligned}$$

where

$$\mathbf{P}^{a[\alpha_1 \cdots \alpha_d]} \equiv \tfrac{1}{d!}\, \mathbf{P}_{a_1}^{\alpha_1} \cdots \mathbf{P}_{a_d}^{\alpha_d}\, \varepsilon^{a a_1 \cdots a_d}$$

In these equations, the right hand side has tensorial indices $[B_i \cdots B_k \gamma \cdots \delta A_i \cdots A_k]$ in addition to a or α, but so long as the matching constraints hold any value of these indices gives a vector parallel to $\mathbf{x}^a$ or $\mathbf{x}^\alpha$ (*i.e.* for *matching* image points the tensor $\mathbf{P}^{a[B_i \cdots B_k \gamma \cdots \delta}\, \mathbf{x}^{A_i} \cdots \mathbf{x}^{A_k]}$ can be factorized as $\mathbf{x}^a\, \mathbf{T}^{[B_i \cdots B_k \gamma \cdots \delta A_i \cdots A_k]}$ for some tensors $\mathbf{x}^a$ and $\mathbf{T}$). Again it is enough to antisymmetrize over the indices of each image separately. For 2D images of 3D space the possible minimal reconstruction techniques are $\mathbf{P}^{a[B_1 C_1 B_2}\, \mathbf{x}^{A_1} \mathbf{x}^{A_2]}$ and $\mathbf{P}^{a[B_1 B_2 B_3}\, \mathbf{x}^{A_1} \mathbf{x}^{A_2} \mathbf{x}^{A_3]}$:

$$\begin{aligned} \mathbf{x}^a &\sim \left(\varepsilon_{A_1 B_1 C_1}\, \mathbf{x}^{A_1}\, \mathbf{P}_b^{B_1} \mathbf{P}_c^{C_1}\right) \cdot \\ &\qquad \cdot \left(\varepsilon_{A_2 B_2 C_2}\, \mathbf{x}^{A_2}\, \mathbf{P}_d^{C_2}\right) \varepsilon^{abcd} \\ \mathbf{x}^a &\sim \left(\varepsilon_{A_1 B_1 C_1}\, \mathbf{x}^{A_1}\, \mathbf{P}_b^{C_1}\right) \left(\varepsilon_{A_2 B_2 C_2}\, \mathbf{x}^{A_2}\, \mathbf{P}_c^{C_2}\right) \cdot \\ &\qquad \cdot \left(\varepsilon_{A_3 B_3 C_3}\, \mathbf{x}^{A_3}\, \mathbf{P}_d^{C_2}\right) \varepsilon^{abcd} \end{aligned}$$

These correspond respectively to finding the intersection of the optical ray from one image and the constraint plane from one coordinate of the second one, and to finding the intersection of three constraint planes from one coordinate in each of three images.

To recover the additional matching constraints that apply to the minimal reconstruction solution with indices $[B_i \cdots B_k \gamma \cdots \delta A_i \cdots A_k]$, project the solution to some image l to get

$$\mathbf{P}_a^{C_l}\, \mathbf{x}^a = \mathbf{I}^{\,C_l[B_i \cdots B_k \gamma \cdots \delta}\, \mathbf{x}^{A_i} \ldots \mathbf{x}^{A_k]}$$

If the constraint is to hold, this must be proportional to $\mathbf{x}^{C_l}$. If l is one of the existing images (i, say) $\mathbf{x}^{A_l}$ is already in the antisymmetrization, so if we extend the antisymmetrization to C_l the result must vanish: $\mathbf{I}^{\,[C_l B_i \cdots B_k \gamma \cdots \delta}\, \mathbf{x}^{A_i} \ldots \mathbf{x}^{A_k]} = \mathbf{0}$. If l is distinct from the existing images we can explicitly add $\mathbf{x}^{A_l}$ to the antisymmetrization list, to get $\mathbf{I}^{\,[C_l B_i \cdots B_k \gamma \cdots \delta}\, \mathbf{x}^{A_i} \ldots \mathbf{x}^{A_k} \mathbf{x}^{A_l]} = \mathbf{0}$.

Similarly, the minimal reconstruction solution for 3D lines from two images is just the pull-back

$$\mathbf{l}_{ab} \sim \mathbf{l}_{A_1} \mathbf{l}_{A_2}\, \mathbf{P}_{[a}^{A_1} \mathbf{P}_{b]}^{A_2}$$

or in contravariant form

$$\mathbf{l}^{ab} \sim \mathbf{l}_{A_1} \mathbf{l}_{A_2}\, \mathbf{P}_c^{A_1} \mathbf{P}_d^{A_2}\, \varepsilon^{abcd}$$

This can be projected into a third image and dualized to give the previously stated line transfer equation

$$\begin{aligned} \mathbf{l}_{A_3} &\sim \mathbf{l}_{A_1} \mathbf{l}_{A_2} \cdot \varepsilon_{A_3 B_3 C_3}\, \mathbf{P}_a^{B_3} \mathbf{P}_b^{C_3} \mathbf{P}_c^{A_1} \mathbf{P}_d^{A_2}\, \varepsilon^{abcd} \\ &\sim \mathbf{l}_{A_1} \mathbf{l}_{A_2}\, \mathbf{G}_{A_3}{}^{A_1 A_2} \end{aligned}$$

More generally, the covariant form of the k-subspace constraint equations given in section 7.5 generates basic reconstruction equations for k dimensional subspaces of the j^{th} image or the world space by dropping one index A_j from the contraction and using it as the α_0 of a set of $k+1$ free indices $\alpha_0 \cdots \alpha_k$ designating the reconstructed k-subspace in $\mathcal{P}^{A_j}$. To reconstruct the k-subspace in world coordinates, the projection tensors $\mathbf{P}_{a_i}^{\alpha_i}$ corresponding to the free indices must also be dropped, leaving free world indices $a_0 \cdots a_k$.

9. Grassmann Relations between Matching Tensors

The components of any Grassmann tensor must satisfy a set of quadratic 'simplicity' constraints called the **Grassmann relations**. In our case the joint image Grassmannian satisfies

$$\begin{aligned} \mathbf{0} &= \mathbf{I}^{\,\alpha_0 \cdots \alpha_{d-1}[\beta_0}\, \mathbf{I}^{\,\beta_0 \cdots \beta_{d+1}]} \\ &= \tfrac{1}{d+2} \sum_{a=0}^{d+1} (-1)^a\, \mathbf{I}^{\,\alpha_0 \cdots \alpha_{d-1} \beta_a}\, \mathbf{I}^{\,\beta_0 \cdots \beta_{a-1} \beta_{a+1} \cdots \beta_{d+1}} \end{aligned}$$

Mechanically substituting expressions for the various components of $\mathbf{I}^{\,\alpha_0 \cdots \alpha_d}$ in terms of the matching tensors produces a long list of quadratic relations between the matching tensors. For reference, table 1 gives a (hopefully complete) list of the identities that can be generated between the matching tensors of two and three images in $d = 3$ dimensions, modulo image permutation, traces of identities with covariant and contravariant indices from the same image, and (anti-)symmetrization operations on identities with several covariant or contravariant indices from the same image. (For example, $\mathbf{F}_{A_2A_3}\,\mathbf{G}_{A_1}{}^{A_2A_3} = 2\,\mathbf{F}_{A_1A_2}\,\mathbf{e}_3{}^{A_2}$ and $\mathbf{F}_{A_3(A_1}\,\mathbf{G}_{B_1)}{}^{A_2A_3} = \mathbf{0}$ follow respectively from tracing $[112, 22233]$ and symmetrizing $[112, 11333]$). The constraint tensors are assumed to be normalized as in their above definitions, in terms of an arbitrary choice of scale for the underlying image projections. In practice, these scale factors must often be recovered from the Grassmann relations themselves. Note that with these conventions, $\mathbf{F}_{A_1A_2} = \mathbf{F}_{A_2A_1}$ and $\mathbf{G}_{A_1}{}^{A_2A_3} = -\mathbf{G}_{A_1}{}^{A_3A_2}$. For clarity the free indices have been displayed on the (zero) left-hand side tensors. The labels indicate one choice of image numbers for the indices of the Grassmann simplicity relation that will generate the identity (there may be others).

As an example of the use of these identities, $\mathbf{G}_{A_1}{}^{A_2A_3}$ follows from linearly from $\mathbf{F}_{A_1A_2}$, $\mathbf{F}_{A_1A_3}$ and the corresponding epipoles $\mathbf{e}_1{}^{A_2}$, $\mathbf{e}_3{}^{A_1}$ and $\mathbf{e}_3{}^{A_2}$ by applying $[112, 11333]$ and $[112, 22333]$.

10. Reconstruction in Joint Image Space

We have argued that multi-image projective reconstruction is essentially a matter of recovering a coherent set of projective scale factors for the measured image points, that it canonically takes place in the joint image space $\mathcal{P}^\alpha$, and that reconstruction in world coordinates is best seen as a choice of basis in the resulting joint image subspace $\mathcal{PI}^\alpha$. To emphasize these points it is interesting to develop 'reconstruction' techniques that work directly in joint image space using measured image coordinates, without reference to *any* 3D world or basis.

First suppose that the complete set of matching tensors between the images has been recovered. It is still necessary to fix an arbitrary overall scale factor for each image. This can be done by choosing any coherent set of relative scalings for the matching tensors, so that they verify the Grassmann simplicity relations as given above. Then, since the components of the joint

Table 1. The Grassmann identities between the matching tensors of two and three images.

$\mathbf{0}_{A_1}$	$= \mathbf{F}_{A_1A_2}\,\mathbf{e}_1{}^{A_2}$	$[111, 11122]$
$\mathbf{0}^{A_1A_2}$	$= \mathbf{F}_{B_1B_2}\,\mathbf{F}_{C_1C_2}\,\varepsilon^{B_1C_1A_1}\,\varepsilon^{B_2C_2A_2} + 2\,\mathbf{e}_2{}^{A_1}\,\mathbf{e}_1{}^{A_2}$	$[112, 11222]$
$\mathbf{0}_{A_3}$	$= \mathbf{F}_{A_2A_3}\,\mathbf{e}_1{}^{A_2} - \varepsilon_{A_3B_3C_3}\,\mathbf{e}_1{}^{B_3}\,\mathbf{e}_2{}^{C_3}$	$[111, 22233]$
$\mathbf{0}^{A_3}_{A_1A_2}$	$= \varepsilon_{A_2B_2C_2}\,\mathbf{e}_1{}^{B_2}\,\mathbf{G}_{A_1}{}^{C_2A_3} + \mathbf{e}_1{}^{A_3}\,\mathbf{F}_{A_1A_2}$	$[111, 11223]$
$\mathbf{0}^{A_1}_{A_2A_3}$	$= \varepsilon_{A_2B_2C_2}\,\mathbf{e}_1{}^{B_2}\,\mathbf{G}_{A_3}{}^{A_1C_2} + \varepsilon_{A_3B_3C_3}\,\mathbf{e}_1{}^{B_3}\,\mathbf{G}_{A_2}{}^{A_1C_3}$	$[111, 12233]$
$\mathbf{0}^{A_1A_2A_3}_{B_2}$	$= \mathbf{F}_{B_1B_2}\,\mathbf{G}_{C_1}{}^{A_2A_3}\,\varepsilon^{B_1C_1A_1} - \mathbf{e}_1{}^{A_2}\,\mathbf{G}_{B_2}{}^{A_1A_3} + \boldsymbol{\delta}_{B_2}{}^{A_2}\,\mathbf{e}_1{}^{C_2}\,\mathbf{G}_{C_2}{}^{A_1A_3}$	$[112, 11223]$
$\mathbf{0}^{A_2B_2}_{A_1B_1A_3}$	$= \varepsilon_{A_3B_3C_3}\,\mathbf{G}_{A_1}{}^{A_2B_3}\,\mathbf{G}_{B_1}{}^{B_2C_3} - \mathbf{e}_1{}^{A_2}\,\varepsilon_{A_1B_1C_1}\,\mathbf{G}_{A_3}{}^{C_1B_2}$	$[112, 11233]$
	$-\mathbf{F}_{A_1C_2}\,\varepsilon^{C_2A_2B_2}\,\mathbf{F}_{B_1A_3}$	
$\mathbf{0}^{A_2}_{A_1B_1}$	$= \mathbf{F}_{A_1A_3}\,\mathbf{G}_{B_1}{}^{A_2A_3} + \varepsilon_{A_1B_1C_1}\,\mathbf{e}_3{}^{C_1}\,\mathbf{e}_1{}^{A_2}$	$[112, 11333]$
$\mathbf{0}^{B_1B_2}_{A_1A_2A_3}$	$= \varepsilon_{A_3B_3C_3}\,\mathbf{G}_{A_1}{}^{B_2B_3}\mathbf{G}_{A_2}{}^{B_1C_3} - \mathbf{F}_{A_1A_2}\,\mathbf{G}_{A_3}{}^{B_1B_2} + \boldsymbol{\delta}_{A_2}{}^{B_2}\,\mathbf{F}_{A_1C_2}\,\mathbf{G}_{A_3}{}^{B_1C_2}$	$[112, 12233]$
	$+\boldsymbol{\delta}_{A_1}{}^{B_1}\,\mathbf{e}_1{}^{B_2}\,\mathbf{F}_{A_2A_3}$	
$\mathbf{0}^{B_1A_2B_2}_{A_1}$	$= \mathbf{G}_{C_3}{}^{B_1B_2}\,\mathbf{G}_{A_1}{}^{A_2C_3} + \mathbf{e}_3{}^{B_1}\,\mathbf{F}_{A_1C_2}\,\varepsilon^{C_2A_2B_2} + \boldsymbol{\delta}_{A_1}{}^{B_1}\,\mathbf{e}_1{}^{A_2}\,\mathbf{e}_3{}^{B_2}$	$[112, 12333]$
$\mathbf{0}^{A_2}_{A_1A_3}$	$= \varepsilon_{A_3B_3C_3}\,\mathbf{e}_2{}^{B_3}\,\mathbf{G}_{A_1}{}^{A_2C_3} - \mathbf{F}_{A_1B_2}\mathbf{F}_{C_2A_3}\,\varepsilon^{B_2C_2A_2}$	$[112, 22233]$
$\mathbf{0}^{B_2}_{A_1A_2}$	$= \mathbf{F}_{A_2A_3}\,\mathbf{G}_{A_1}{}^{B_2A_3} + \mathbf{F}_{A_1A_2}\,\mathbf{e}_3{}^{B_2} - \boldsymbol{\delta}_{A_2}{}^{B_2}\,\mathbf{F}_{A_1C_2}\,\mathbf{e}_3{}^{C_2}$	$[112, 22333]$
$\mathbf{0}^{A_1A_2B_2A_3B_3}$	$= \mathbf{G}_{B_1}{}^{A_2A_3}\,\mathbf{G}_{C_1}{}^{B_2B_3}\,\varepsilon^{B_1C_1A_1} - \mathbf{G}_{C_2}{}^{A_1A_3}\,\varepsilon^{C_2A_2B_2}\,\mathbf{e}_1{}^{B_3}$	$[123, 11123]$
	$-\mathbf{G}_{C_3}{}^{A_1A_2}\,\varepsilon^{C_3A_3B_3}\,\mathbf{e}_1{}^{B_2}$	
$\mathbf{0}^{B_1B_2A_3B_3}_{A_1A_2}$	$= \mathbf{G}_{A_2}{}^{B_1A_3}\,\mathbf{G}_{A_1}{}^{B_2A_3} - \mathbf{G}_{A_2}{}^{B_1B_3}\,\mathbf{G}_{A_1}{}^{B_2A_3} - \mathbf{F}_{A_1A_2}\,\mathbf{G}_{C_3}{}^{B_1B_2}\,\varepsilon^{C_3A_3B_3}$	$[123, 11223]$
	$-\boldsymbol{\delta}_{A_2}{}^{B_2}\,\mathbf{G}_{C_2}{}^{B_1A_3}\,\mathbf{G}_{A_1}{}^{C_2B_3} + \boldsymbol{\delta}_{A_1}{}^{B_1}\,\mathbf{G}_{A_2}{}^{C_1B_3}\,\mathbf{G}_{C_1}{}^{B_2A_3}$	

image Grassmann tensor $\mathbf{I}^{\,\alpha\beta\cdots\gamma}$ can be recovered directly from the matching tensors, the location of the joint image $\mathcal{PI}^\alpha$ has been fixed.

Now consider a matching set of image points $\{\mathbf{x}^{A_1},\ldots,\mathbf{x}^{A_m}\}$ with arbitrary relative scalings. As discussed in section 6, the matching constraints are equivalent to the requirement that there be a rescaling of the image points that places the joint image space vector $\sum_{i=1}^m \lambda_i\,\mathbf{x}^{A_i}$ in the joint image $\mathcal{PI}^\alpha$. Expressed in terms of the Grassmannian, this becomes the **joint image reconstruction system**

$$\mathbf{I}^{\,[\alpha\beta\cdots\gamma}\cdot\left(\sum_{i=1}^m \lambda_i\,\mathbf{x}^{A_i]}\right) \;=\; \mathbf{0}$$

This is a redundant set of homogeneous multilinear equations in the Grassmannian $\mathbf{I}^{\,\alpha\beta\cdots\gamma}$, the image points $\mathbf{x}^{A_i}$, and the scale factors λ_i, that can be used to 'reconstruct' the scale factors given the Grassmannian and the image measurements.

These equations can be reexpressed in terms of the matching tensors, in much the same way as the Grassmann simplicity relations can. The types of constraint that can arise for 2D images of 3D space are shown in table 1. The left hand sides are zero tensors and the labels give index image numbers that will generate the equation. The numerical coefficients are valid only for correctly scaled matching tensors. Permuting the images generates further equations. Note that since the equations are algebraically redundant it is only necessary to apply a subset of at least $m-1$ of them to solve for the m scale factors. The optimal choice of equations probably depends on the ease and accuracy with which the various matching tensor components can be estimated.

Recovery of the scale factors locates the reconstructed joint image point $\mathbf{x}^\alpha$ unambiguously in the subspace $\mathcal{PI}^\alpha$. Its coordinates in any chosen basis (*i.e.* with respect to any given choice of the basis-vector columns of the joint projection matrix $\mathbf{P}_a^\alpha$) can easily be obtained, if required. Although this process is arguably too abstract to be called 'reconstruction', all of the relevant structure is certainly present in the joint image representation and can easily be extracted from it.

Given an efficient numerical technique for the resolution of sets of bilinear equations and a sufficient number of matching points, it would also be possible to solve the above equations simultaneously for the vector of matching tensor components and the vector of scale factors, given the measured image coordinates as coefficients. Algebraic elimination of the scale factors from these equations should ultimately lead back to the various matching constraints (modulo probably heavy use of the Grassmann relations). Elimination of the matching tensors (modulo the matching constraints viewed as constraints on the matching tensor components) for sufficiently many matching points would lead to (high degree!) basic reconstruction methods for the recovery of the scale factors directly from measured image coordinates.

Geometrically, the reconstruction process can be pictured as follows. Each image point is a D_i-codimensional subset of its D_i-dimensional image, so under the trivial projection it can be pulled back to a D_i-codimensional subspace of the joint image space $\mathcal{P}^\alpha$. Intersecting the subspaces pulled back from the different images results in an $(m-1)$-dimensional projective subspace of $\mathcal{P}^\alpha$. This is precisely the set of all possible rescalings of the $\mathbf{x}^{A_i}$. The joint image $\mathcal{PI}^\alpha$ intersects

Table 1. The five basic types of reconstruction equation for a point in the joint image.

$\mathbf{0}_{A_2}$	$=$	$(\mathbf{F}_{A_1A_2}\,\mathbf{x}^{A_1})\lambda_1 + (\varepsilon_{A_2B_2C_2}\,\mathbf{e}_1{}^{B_2}\,\mathbf{x}^{C_2})\lambda_2$	[11122]
$\mathbf{0}^{A_2A_3}$	$=$	$(\mathbf{G}_{A_1}{}^{A_2A_3}\,\mathbf{x}^{A_1})\lambda_1 - (\mathbf{e}_1{}^{A_3}\,\mathbf{x}^{A_2})\lambda_2 + (\mathbf{e}_1{}^{A_2}\,\mathbf{x}^{A_3})\lambda_3$	[11123]
$\mathbf{0}_{A_1A_2}^{A_3}$	$=$	$(\varepsilon_{A_1B_1C_1}\,\mathbf{G}_{A_2}{}^{B_1A_3}\,\mathbf{x}^{C_1})\lambda_1 + (\varepsilon_{A_2B_2C_2}\,\mathbf{G}_{A_1}{}^{B_2A_3}\,\mathbf{x}^{C_2})\lambda_2 - (\mathbf{F}_{A_1A_2}\,\mathbf{x}^{A_3})\lambda_3$	[11223]
$\mathbf{0}_{A_1}^{A_2A_3A_4}$	$=$	$(\varepsilon_{A_1B_1C_1}\,\mathbf{H}^{B_1A_2A_3A_4}\,\mathbf{x}^{C_1})\lambda_1 + (\mathbf{G}_{A_1}{}^{A_4A_3}\,\mathbf{x}^{A_2})\lambda_2 - (\mathbf{G}_{A_1}{}^{A_2A_4}\,\mathbf{x}^{A_3})\lambda_3$ $+(\mathbf{G}_{A_1}{}^{A_2A_3}\,\mathbf{x}^{A_4})\lambda_4$	[11234]
$\mathbf{0}^{A_1A_2A_3A_4A_5}$	$=$	$(\mathbf{H}^{A_2A_3A_4A_5}\,\mathbf{x}^{A_1})\lambda_1 - (\mathbf{H}^{A_1A_3A_4A_5}\,\mathbf{x}^{A_2})\lambda_2 + (\mathbf{H}^{A_1A_2A_4A_5}\,\mathbf{x}^{A_3})\lambda_3$ $-(\mathbf{H}^{A_1A_2A_3A_5}\,\mathbf{x}^{A_4})\lambda_4 + (\mathbf{H}^{A_1A_2A_3A_4}\,\mathbf{x}^{A_5})\lambda_5$	[12345]

this subspace if and only if the matching constraints are satisfied, and the intersection is of course the desired reconstruction. So the problem of multi-image projective reconstruction from points can be viewed as the search for the $(d+m-1)$-dimensional subspace of $\mathcal{P}^\alpha$ that contains (or comes closest to containing) a given set of $(m-1)$-dimensional joint-image-point subspaces, followed by an arbitrary choice (the scale factors) of a d-dimensional subspace (the joint image) of the $(d+m-1)$-dimensional space that meets each joint-image-point subspace transversally. The reconstruction of lines and higher dimensional subspaces can be viewed in similarly geometric terms.

11. Perspectives

The theoretical part of the paper is now finished, but before closing it may be worthwhile to reflect a little on our two principal themes: projective reconstruction and the tensor calculus. We will take it for granted that the projective and algebraic-geometric approaches to vision are here to stay: the 'unreasonable efficacy of mathematics in the physical sciences' can only lead to an increasing mathematization of the field.

11.1. Matching & Reconstruction

Clearly visual scene reconstruction is a large and complex problem that is not going to be 'solved' by any one contribution, so we will restrict ourselves to a few technical remarks. To the extent that the problem can be decomposed at all, the most difficult parts of it will probably always be the low level feature extraction and token matching. 3D reconstruction seems relatively straightforward once image tokens have been put into correspondence, although much remains to be done on the practical aspects, particularly on error models [17], [4], [21] and the recovery of Euclidean structure [17].

Given the complexity and algebraic redundancy of the trilinear and quadrilinear constraints it is certainly legitimate to ask whether they are actually likely to be *useful* in practice. I think that the answer is a clear 'yes' for the trilinear constraints and the overall joint image/Grassmannian picture, but the case for the quadrilinear constraints is still open.

The principal application of the matching tensors must be for token matching and verification. The trilinear constraints can be used directly to verify the correspondence of a triple of points or lines, or indirectly to transfer a hypothesized feature location to a third image given its location in two others, in a hypothesize-and-test framework. Image synthesis (*e.g.* image sequence compression and interpolation) is likely to be another important application of transfer [11].

Fundamental matrices can also be used for these applications, but because the higher order constraints 'holistically' combine data from several images and there is built-in redundancy in the constraint equations, it is likely that they will prove less prone to mismatches and numerically more stable than a sequence of applications of the epipolar constraint. For example Shashua [19] has reported that a single trilinear constraint gives more reliable transfer results than two epipolar ones, and Faugeras and Mourrain [7] have pointed out that bilinear constraint based transfer breaks down when the 3D point lies in the trifocal plane or the three optical centres are aligned, whereas trilinear transfer continues to be reasonably well conditioned.

When there are four images the quadrilinear constraint can also be used for point matching and transfer, but the equations are highly redundant and it seems likely that bilinear and trilinear methods will prove adequate for the majority of applications. The trilinear constraint is nonsingular for almost all situations involving points, provided the optical centres do not coincide and the points avoid the lines passing between them.

The most important failure for lines is probably that for lines lying in an epipolar plane of two of the images. In this case the constraints mediated by trivalent tensors are vacuous (although there is still enough information to reconstruct the corresponding 3D line unless it lies in the trifocal plane or the optical centres are aligned) and those mediated by quadrivalent tensors are rank deficient. But given the linear dependence of the various line constraints it is not clear that the quadrivalent ones have any advantage over an equivalent choice of trivalent ones.

A closely related issue is that of linear versus higher order methods. Where possible, linear formulations are usually preferred. They tend to be simpler, faster, better understood and numerically more stable than their nonlinear counterparts, and they are usually much easier to adapt to redundant data, which is common in vision and provides increased accuracy and robustness.

On the other hand, nonlinear constraints can not be represented accurately within a linear framework.

This is especially relevant to the estimation of the matching tensors. We have emphasized that the matching tensor components and constraint equations are *linearly* independent but *quadratically* highly dependent. It is straightforward to provide linear minimum-eigenvector methods to estimate: the 9-component fundamental matrix from at least 8 pairs of corresponding points in two images [12], [13]; each of the three linearly independent 27-component trilinear tensors from at least 7 triples of points in three images; and the 81-component quadrilinear tensor from at least 6 quadruples of corresponding points in four images [21]. For complex applications several of these tensors might be needed, for example a fundamental constraint might provide initial feature pairings that can be used to check for corresponding features in a third image using further fundamental or trilinear constraints. Also, different trilinear tensors are required for point transfer and line transfer.

Unfortunately, it turns out that the above linear estimation techniques (particularly that for the fundamental matrix) are numerically rather poorly conditioned, so that the final estimates are very sensitive to measurement errors and outliers. Moreover, even in the case of a single fundamental matrix there is a nonlinear constraint that can not be expressed within the linear framework. The quadratic epipolar relation $\mathbf{F}_{A_1A_2}\,\mathbf{e}_1{}^{A_2} = \mathbf{0}$ implies the cubic constraint $\mathrm{Det}(\mathbf{F}) = 0$. If this constraint is ignored, one finds that the resulting estimates of $\mathbf{F}$ and the epipoles tend to be rather inaccurate [13]. In fact, the linear method is often used only to initialize nonlinear optimization routines that take account of the nonlinearity and the estimated measurement errors in the input data.

This leads to the following open question: *When several matching tensors are being estimated, to what extent is it possible or necessary to take account of the quadratic constraints between them?* The full set of quadratic relations is very complex and it is probably not practical to account for all of them individually: it would be much simpler just to work directly in terms of the 3D joint image geometry. Moreover, many of the relations depend on the relative scaling of the constraint tensors and the recovery of these further complicates the issue (it is a question of exactly which combinations of components need to be fixed to ensure consistency and numerical stability). On the other hand, experience with the fundamental matrix suggests that it is dangerous to ignore the constraints entirely. Some at least of them are likely to be important in any given situation. Our current understanding of these matters is very sketchy: essentially all we have is a few *ad hoc* comparisons of particular techniques.

As a final point, a few people seem to have been hoping for some 'magic' reconstruction technique that completely avoids the difficulties of image-to-image matching, perhaps by holistically combining data from a large number of images (or a single dense image sequence). The fact that the matching constraints stop at four images (or equivalently three time derivatives) does not preclude this, but perhaps makes it seem a little less likely. On the other hand, the simplicity of the joint image picture makes incremental recursive reconstruction techniques that correctly handle the measurement errors and constraint geometry seem more likely (*c.f.* [16]).

11.2. Tensors vs. the Rest

This paper is as much about the use of tensors as a vehicle for mathematical vision as it is about image projection geometry. Tensors have seldom been used in vision and many people appear to be rather tensor-phobic, so it seems appropriate to say a few words in their favour: "*Don't panic!*" [1].

First of all, what *is* a tensor? — It is a collection (a multidimensional array) of components that represent a single geometric object with respect to some system of coordinates, and that are intermixed when the coordinate system is changed. This immediately evokes the two principal concerns of tensor calculus: *(i)* to perform manipulations *abstractly* at the object level rather than explicitly at the component level; and *(ii)* to ensure that all expressions are properly *covariant* (*i.e.* have the correct transformation laws) under changes of basis. The advantages are rather obvious: the higher level of abstraction brings greater compactness, clarity and insight, and the guaranteed covariance of well-formed tensorial expressions ensures that no hidden assumptions are made and that the correct algebraic symmetries and relationships between the components are automatically preserved.

Vectors are the simplest type of tensor and the familiar 3D vector calculus is a good example of the above points: it is much simpler and less error prone to write a single vector $\mathbf{x}$ instead of three components $(x^1,\ x^2,\ x^3)$ and a symbolic cross product $\mathbf{z} = \mathbf{x} \times \mathbf{y}$ instead of three equations $z^1 = x^2 y^3 - x^3 y^2$, $z^2 = x^3 y^1 - x^1 y^3$ and $z^3 = x^1 y^2 - x^2 y^1$. Unfortunately, the simple index-free matrix-vector notation seems to be difficult to extend to higher-order tensors with the required degree of flexibility. (Mathematicians sometimes define tensors as multilinear functions $\mathbf{T}(\mathbf{x}, \ldots, \mathbf{z})$ where $\mathbf{x}, \ldots, \mathbf{z}$ are vectors of some type and the result is a scalar, but this notation becomes hopelessly clumsy when it comes to inter-tensor contractions, antisymmetrization and so forth). In fact, the index-free notation becomes as much a dangerous weapon as a useful tool as soon as one steps outside the realm of simple vector calculations in a single Euclidean space. It is only too easy to write $\mathbf{x}^\top \mathbf{x} = 1$ in a projective space where no transpose (metric tensor) exists, or a meaningless 'epipolar equation' $\mathbf{l}^\top \mathbf{F}\, \mathbf{x} = 0$ where $\mathbf{l}$ is actually the 3-component vector of an image line (rather than an image *point*) and $\mathbf{x}$ belongs to the wrong image for the fundamental matrix $\mathbf{F}$ (which should have been transposed in any case).

To avoid this sort of confusion, it is essential to use a notation that clearly distinguishes the space and covariant/contravariant type of each index. Although it can not be denied that this sometimes leads to rather baroque-looking formulae — especially when there are many indices from many different spaces as in this paper — it is much preferable to the alternatives of using either no indices at all or i, j, and k for everything, so that one can never quite see what is supposed to be happening. It is important not to be fooled into thinking that tensor equations are intrinsically difficult just because they have indices. For simple calculations the indexed notation is not significantly more difficult to use than the traditional index-free one, and it becomes *much* clearer and more powerful in complex situations. For a visually appealing (but typographically inconvenient) pictorial notation, see the appendix of [18].

Simultaneously with the work presented in this paper, at least two other groups independently converged on parts of the constraint geometry from component-based points of view: Faugeras & Mourrain [7] using the Grassmann-Cayley algebra of skew linear forms, and Werman & Shashua [22] using Gröbner bases and algebraic elimination theory. Both approaches make very heavy use of computer algebra whereas all of the calculations in the present paper were done by hand, and neither (notwithstanding the considerable value of their results) succeeded in obtaining anything like a complete picture of the constraint geometry. My feeling is that it is perhaps no accident that in each of the three categories: level of geometric abstraction, efficiency of calculation, and insight gained, the relative ordering is the same: tensor calculus $>$ Grassmann-Cayley algebra $>$ elimination theory.

Elimination-theoretic approaches using resultants and Gröbner bases seem to be intrinsically component-based. They take no account of the tensorial structure of the equations and therefore make no use of the many symmetries between them, so even when the coordinate systems are carefully adapted to the problem they tend to carry a significant amount of computational redundancy. Werman & Shashua [22] suggest that an advantage of such approaches is the fact that very little geometric insight is required. Unfortunately, one might also suggest that very little geometric insight is *gained*: the output is a complex set of equations with no very clearly articulated structure.

The Grassmann-Cayley algebra [7], [2] is spiritually much closer to the tensorial point of view. Indeed, it can be viewed as a specialized index-free notation for manipulating completely antisymmetric covariant and contravariant tensors. It supports operations such as antisymmetrization over indices from several tensors (wedge product), contractions over corresponding sets of covariant and contravariant antisymmetric indices (hook product), and contravariant-covariant dualization (sometimes used to identify the covariant and contravariant algebras and then viewed as the identity, in which case the hook product is replaced by the join product). Given the connection with Grassmann coordinates, the Grassmann-Cayley algebra can be viewed as a calculus of intersection and union (span) for projective subspaces: clearly a powerful and highly relevant concept. It is likely that this approach would have lead fairly rapidly to the full Grassmannian matching constraint geometry, notwithstanding the relative opacity of the initial component-oriented formulations.

Despite its elegance, there are two problems with the Grassmann-Cayley algebra as a general formalism. The first is that it is not actually very general: it is good for calculations with linear or projective subspaces, but it does not extend gracefully to more complex situations or higher-degree objects. For example quadric surfaces are represented by *symmetric* tensors which do not fit at all well into the antisymmetric algebra. Tensors are much more flexible in this regard. The second problem with the Grassmann-Cayley algebra is that it is often infuriatingly vague about geometric (covariance) issues. Forms of different degree with indices from different spaces can be added formally within the algebra, but this makes no sense at all tensorially: such objects do not transform reasonably under changes of coordinates, and consequently do not have any clear *geometric* meaning, whatever the status of the algebra. The fact that the algebra has a stratified tensorial structure is usually hidden in the definitions of the basic product operations, but it becomes a central issue as soon as geometric invariance is called into question.

In summary, my feeling is that the tensorial approach is ultimately the most promising. The indexed notation is an extraordinarily powerful, general and flexible tool for the algebraic manipulation of geometric objects. It displays the underlying the structure and covariance of the equations very clearly, and it naturally seems to work at about the right level of abstraction for practical calculations: neither so abstract nor so detailed as to hide the essential structure of the problem. Component-based approaches are undoubtedly useful, but they are probably best reserved until *after* a general tensorial derivation has been made, to specialize and simplify a set of abstract tensorial equations to the particular application in hand.

As an example of this, a $k+1$ index antisymmetric tensor representing a k dimensional subspace of a d dimensional projective space has (very naïvely) $(d+1)^{k+1}$ components, but only $\binom{d+1}{k+1}$ of these are linearly independent owing to antisymmetry. The independent components can easily be enumerated (the indices $i_0 i_1 \cdots i_k$ for $0 \le i_0 < i_1 < \ldots < i_k \le d$ form a spanning set) and gathered into an explicit $\binom{d+1}{k+1}$ component vector for further numerical or symbolic manipulation. In fact, these components span exactly one tensorial stratum of the Grassmann-Cayley algebra.

It is perhaps unfortunate that current computer algebra systems seem to have very few tools for manipulating general tensorial expressions, as these would greatly streamline the derivation and specialization processes. However, there does not appear to be any serious obstacle to the development of such tools and it is likely that they will become available in the near future.

12. Summary

Given a set of perspective projections into m projective image spaces, there is a 3D subspace of the space of combined image coordinates called the **joint image**. This is a complete projective replica of the 3D world expressed directly in terms of scaled image coordinates. It is defined intrinsically by the physical situation up to an arbitrary choice of some internal scalings. Projective reconstruction in the joint image is a canonical process requiring only a rescaling of the image coordinates. A choice of basis in the joint image allows the reconstruction to be transferred to world space.

There are multilinear **matching constraints** between the images that determine whether a set of image points could be the projection of a single world point. For 3D worlds only three types of constraint exist: the epipolar constraint generated by the fundamental matrix between pairs of images, Shashua's trilinear constraints between triples of images and a new quadrilinear constraint on sets of corresponding points from four images.

Moreover, the entire set of constraint tensors for all the images can be combined into a single compact geometric object, the antisymmetric 4 index **joint image Grassmannian** tensor. This can be recovered from image measurements whenever the individual constraint tensors can. It encodes precisely the information needed for reconstruction: the location of the joint image in the space of combined image coordinates. It also generates the matching constraints for images of lines and a set of **minimal reconstruction techniques** closely associated with the matching constraints. Structural constraints on the Grassmannian tensor produce quadratic identities between the various constraint tensors.

Acknowledgements

This work was supported by the European Community through Esprit programs HCM and SECOND and benefited from discussions with many colleagues. The notation and the tensorial treatment of projective space are based on Roger Penrose's approach to relativistic spinors and twistors [18], which partly derives (I believe) from Hodge & Pedoe's excellent treatise on algebraic geometry [10]. Without these powerful tools this work would not have been possible.

Appendix A

Mathematical Background

This appendix provides a very brief overview of the linear algebra and projective geometry need to understand this paper, and a little background information on our notation. For more details on using tensor calculus for projective space see [10], [18].

A.1. Vectors and Tensors

A **vector space** $\mathcal{H}^a$ is a space on which addition and scaling of elements are defined: $\lambda \mathbf{x}^a + \mu \mathbf{y}^a$ is in $\mathcal{H}^a$ for all scalars λ and μ and elements $\mathbf{x}^a$ and $\mathbf{y}^a$ of $\mathcal{H}^a$. The **span** of a set $\{\mathbf{e}^a_1, \ldots, \mathbf{e}^a_k\}$ of elements of $\mathcal{H}^a$ is the vector space of linear combinations $x^1 \mathbf{e}^a_1 + \cdots + x^k \mathbf{e}^a_k$ of elements of the set. A minimal set that spans the entire space is called a **basis** and the number of elements in the set is the **dimension** of the space. Given a basis $\{\mathbf{e}^a_1, \ldots, \mathbf{e}^a_d\}$ for a d dimensional vector space $\mathcal{H}^a$, any element $\mathbf{x}^a$ of the space can be expressed as $x^1 \mathbf{e}^a_1 + \cdots + x^d \mathbf{e}^a_d$ and associated with the coordinate vector $(x^1, \ldots, x^d)$.

It is helpful to view the superscript a as an **abstract index** [18], *i.e.* an abstract label or placeholder denoting the space the object belongs to. However given a choice of basis it can also be thought of as a variable indexing the coordinate vector that represents the object in that basis.

For every vector space $\mathcal{H}^a$ there is a dual vector space of linear mappings on $\mathcal{H}^a$, denoted $\mathcal{H}_a$. An element $\mathbf{l}_a$ of $\mathcal{H}_a$ acts linearly on an element $\mathbf{x}^a$ of $\mathcal{H}^a$ to produce a scalar. This action is denoted symbolically by $\mathbf{l}_a \mathbf{x}^a$ and called **contraction**. Any basis $\{\mathbf{e}^a_1, \ldots, \mathbf{e}^a_d\}$ for $\mathcal{H}^a$ defines a unique **dual basis** $\{\mathbf{e}^1_a, \ldots, \mathbf{e}^d_a\}$ for $\mathcal{H}_a$ with $\mathbf{e}^i_a \mathbf{e}^a_j = \delta^i_j$, where δ^i_j is 1 when $i = j$ and 0 otherwise. The i^{th} coordinate of $\mathbf{x}^a$ in the basis $\{\mathbf{e}^a_j\}$ is just $x^i \equiv \mathbf{e}^i_a \, \mathbf{x}^a$. If elements of $\mathcal{H}^a$ are represented in the basis $\{\mathbf{e}^a_i\}$ as d index column vectors, elements of $\mathcal{H}_a$ in the dual basis $\{\mathbf{e}^i_a\}$ behave like d index row vectors. Contraction is then just the dot product of the coordinate vectors: $(u_1 \, \mathbf{e}^1_a + \cdots + u_d \, \mathbf{e}^d_a)(x^1 \, \mathbf{e}^a_1 + \cdots + x^d \, \mathbf{e}^a_d) = u_1 \, x^1 + \cdots + u_d \, x^d$. Contraction involves a sum over coordinates but we do not explicitly write the summation signs: whenever a superscript label also appears as a subscript a summation is implied. This is called the **Einstein summation convention**. The order of terms is unimportant: $\mathbf{u}_a \, \mathbf{x}^a$ and $\mathbf{x}^a \, \mathbf{u}_a$ both denote the contraction of the dual vector $\mathbf{u}_a$ with the vector $\mathbf{x}^a$.

Suppose we change the basis in $\mathcal{H}^a$ according to $\mathbf{e}^a_i \to \tilde{\mathbf{e}}^a_i = \sum_j \mathbf{e}^a_j \, \Lambda^j{}_i$ for some matrix $\Lambda^j{}_i$. To keep the resulting abstract element of $\mathcal{H}^a$ the same, coordinate vectors must transform inversely according to $x^i \to \tilde{x}^i = \sum_j (\Lambda^{-1})^i{}_j \, x^j$. To preserve the relations $\tilde{\mathbf{e}}^i_a \, \tilde{\mathbf{e}}^a_j = \delta^i_j$, the dual basis must also transform as $\mathbf{e}^i_a \to \tilde{\mathbf{e}}^i_a = \sum_j (\Lambda^{-1})^i{}_j \, \mathbf{e}^j_a$. Finally, to leave the abstract element of the dual space the same, dual coordinate vectors must transform as $u_i \to \tilde{u}_i = \sum_j u_j \, \Lambda^j{}_i$. Because of the transformations of their coordinates under changes of basis, vectors $\mathbf{x}^a$ are called **contravariant** and dual vectors $\mathbf{u}_a$ are called **covariant**.

An element $\mathbf{x}^a$ of $\mathcal{H}^a$ can also be viewed as a linear mapping on elements of $\mathcal{H}_a$ defined by $\mathbf{u}_a \mathbf{x}^a$, in other words as an element of the dual of the dual of $\mathcal{H}^a$. For finite dimensional spaces every linear mapping on $\mathcal{H}_a$ can be written this way, so there is a complete symmetry between $\mathcal{H}^a$ and $\mathcal{H}_a$: neither is 'more primitive'.

Any nonzero element of $\mathcal{H}_a$ defines a $d-1$ dimensional subspace of $\mathcal{H}^a$ by the equations $\mathbf{u}_a \mathbf{x}^a = 0$, and conversely any $d-1$ dimensional subspace defines a unique element of $\mathcal{H}_a$ up to scale.

It is possible to take formal ('tensor' or 'outer') products of n-tuples of elements of vector spaces, for example a formal element $\mathbf{T}^{aA}{}_\alpha \equiv \mathbf{x}^a \, \mathbf{y}^A \, \mathbf{z}_\alpha$ can be made from elements $\mathbf{x}^a$, $\mathbf{y}^A$, $\mathbf{z}_\alpha$ of vector spaces $\mathcal{H}^a$, $\mathcal{H}^A$ and $\mathcal{H}_\alpha$. The vector space of linear combinations of such objects (for different choices of $\mathbf{x}^a$, $\mathbf{y}^A$ and $\mathbf{z}_\alpha$) is called the tensor product space $\mathcal{H}^{aA}{}_\alpha = \mathcal{H}^a \otimes \mathcal{H}^A \otimes \mathcal{H}_\alpha$. When there are several distinct copies of $\mathcal{H}^a$ we use distinct letters to denote them, *e.g.* $\mathcal{H}^{ab}{}_c = \mathcal{H}^a \otimes \mathcal{H}^b \otimes \mathcal{H}_c$ contains two copies of $\mathcal{H}^a$. Elements of a tensor product space are called **tensors** and can be thought of as multidimensional arrays of components in some chosen set of bases. Under changes of basis each of the indices must be transformed individually.

There are a number of important generic operations on tensors. A set of tensors can be contracted together over any appropriate subset of their indices, for example $\mathbf{u}_{ab} \, \mathbf{x}^a \in \mathcal{H}_b$, $\mathbf{u}_a \, \mathbf{T}^{aB}{}_c \, \mathbf{x}^c \in \mathcal{H}^B$. Self contractions $\mathbf{T}^{ab\cdots}{}_{ac\cdots} \in \mathcal{H}^{b\cdots}{}_{c\cdots}$ are called **traces**. A group of indices can be **(anti-)symmetrized** by averaging over all possible permutations of their positions, with an additional minus sign for odd permutations during antisymmetrization. On indices, $(\cdots)$ denotes symmetrization and $[\cdots]$ antisymmetrization. For example $\mathbf{T}^{(ab)} = \frac{1}{2}(\mathbf{T}^{ab} + \mathbf{T}^{ba})$ and $\mathbf{T}^{[ab]} = \frac{1}{2}(\mathbf{T}^{ab} - \mathbf{T}^{ba})$ can be viewed as symmetric and antisymmetric matrices, and $\mathbf{T}^{[abc]} = \frac{1}{3!}(\mathbf{T}^{abc} - \mathbf{T}^{bac} + \mathbf{T}^{bca} - \mathbf{T}^{cba} + \mathbf{T}^{cab} - \mathbf{T}^{acb})$ is an antisymmetric 3 index tensor. A group of indices is **(anti-)symmetric** if (anti-)symmetrization over them does not change the tensor: $(\cdots)$ and $[\cdots]$ are also used to denote this, for example $\mathbf{T}^{[ab]}_{(cd)} \in \mathcal{H}^{[ab]}{}_{(cd)}$ is antisymmetric in ab and symmetric in cd. Permutation of (anti-)symmetric indices changes at most the sign of the tensor.

In d dimensions antisymmetrizations over more than d indices vanish: in any basis each index must take a distinct value between 1 and d. Up to scale there is a unique antisymmetric d index tensor $\varepsilon^{a_1 a_2 \cdots a_d} \in \mathcal{H}^{[a_1 a_2 \cdots a_d]}$: choosing $\varepsilon^{12\cdots d} = +1$ in some basis, all other components are ± 1 or 0. Under a change of basis the components of $\varepsilon^{a_1 \cdots a_d}$ are rescaled by the determinant of the transformation matrix. There is a corresponding dual tensor $\varepsilon_{a_1 a_2 \cdots a_d} \in \mathcal{H}_{[a_1 a_2 \cdots a_d]}$ with components ± 1 or 0 in the dual basis. $\varepsilon_{a_1 a_2 \cdots a_d}$ defines a volume element on $\mathcal{H}^a$, giving the volume of the hyper-parallelepiped formed by d vectors $\mathbf{x}^a_1, \ldots, \mathbf{x}^a_d$ as $\varepsilon_{a_1 a_2 \cdots a_d}\, \mathbf{x}^{a_1}_1 \cdots \mathbf{x}^{a_d}_d$. The determinant of a linear transformation $\mathbf{T}^a_b$ on $\mathcal{H}^a$ can be defined as $\frac{1}{d!}\varepsilon_{a_1 a_2 \cdots a_d}\, \mathbf{T}^{a_1}_{b_1} \cdots \mathbf{T}^{a_d}_{b_d}\, \varepsilon^{b_1 b_2 \cdots b_d}$, and this agrees with the determinant of the matrix of $\mathbf{T}^a_b$ in any coordinate basis. A contravariant antisymmetric k index tensor $\mathbf{T}^{[a_1 \cdots a_k]}$ has a covariant antisymmetric $d-k$ index **dual** $(*\mathbf{T})_{a_{k+1} \cdots a_d} \equiv \frac{1}{k!}\, \varepsilon_{a_{k+1} \cdots a_d b_1 \cdots b_k}\, \mathbf{T}^{b_1 \cdots b_k}$. Conversely $\mathbf{T}^{a_1 \cdots a_k} = \frac{1}{(d-k)!}\, (*\mathbf{T})_{b_{k+1} \cdots b_d}\, \varepsilon^{b_{k+1} \cdots b_d a_1 \cdots a_k}$. A tensor and its dual contain the same information and both have $\binom{d}{k}$ independent components.

A.2. Grassmann Coordinates

Antisymmetrization and duality are important in the theory of linear subspaces. Consider a set $\{\mathbf{v}^a_1, \ldots, \mathbf{v}^a_k\}$ of k independent vectors spanning a k dimensional subspace Σ of $\mathcal{H}^a$. Given some choice of basis the vectors can be viewed as column vectors and combined into a single $d \times k$ matrix. Any set $\{a_1, \ldots, a_k\}$ of k distinct rows of this matrix defines a $k \times k$ submatrix whose determinant is a $k \times k$ **minor** of the original matrix. Up to a constant scale factor these minors are exactly the components of the tensor $\Sigma^{a_1 \cdots a_k} \equiv \mathbf{v}^{[a_1}_1 \cdots \mathbf{v}^{a_k]}_k$. If the original vectors are independent the $d \times k$ matrix has rank k and at least one of the $k \times k$ minors (and hence the tensor $\Sigma^{a_1 \cdots a_k}$) will not vanish. Conversely, if the tensor vanishes the vectors are linearly dependent.

A vector $\mathbf{x}^a$ lies in the subspace Σ if and only if all of the $(k+1) \times (k+1)$ minors of the $d \times (k+1)$ matrix whose columns are $\mathbf{x}^a$ and the $\mathbf{v}^a_i$ vanish. In tensorial terms: $\mathbf{x}^a$ is an element of Σ if and only if $\Sigma^{[a_1 \cdots a_k}\, \mathbf{x}^{a]} = \mathbf{0}$. So no two distinct subspaces have the same $\Sigma^{a_1 \cdots a_k}$. Under a $k \times k$ linear redefinition $\mathbf{v}^a_i \to \tilde{\mathbf{v}}^a_i = \sum_j \Lambda_i{}^j \mathbf{v}^a_j$ of the spanning vectors, the $k \times k$ minors are simply a constant factor of $\mathrm{Det}(\Lambda_i{}^j)$ different from the old ones by the usual determinant of a product rule. So up to scale $\Sigma^{a_1 \cdots a_k}$ is independent of the set of vectors in Σ chosen to span it.

A subspace Σ can also be defined as the null space of a set of $d-k$ independent linear forms $\{\mathbf{u}^{k+1}_a, \ldots, \mathbf{u}^d_a\}$, *i.e.* as the set of $\mathbf{x}^a$ on which all of the $\mathbf{u}^i_a$ vanish: $\mathbf{u}^i_a\, \mathbf{x}^a = 0$. The $\mathbf{u}^i_a$ can be viewed as a $(d-k) \times d$ matrix of row vectors. Arguments analogous to those above show that the covariant antisymmetric $d-k$ index tensor $\Sigma_{a_{k+1} \cdots a_d} \equiv \mathbf{u}^{k+1}_{[a_{k+1}} \cdots \mathbf{u}^d_{a_d]}$ is independent (up to scale) of the $\{\mathbf{u}^i_a\}$ chosen to characterize Σ and defines Σ as the set of points for which $\Sigma_{a_{k+1} \cdots a_d} \mathbf{x}^{a_d} = \mathbf{0}$. We use the same symbol for $\Sigma_{a_{k+1} \cdots a_d}$ and $\Sigma^{a_1 \cdots a_k}$ because up to scale they turn out to be mutually dual: $\Sigma_{a_{k+1} \cdots a_d} \sim \frac{1}{k!}\, \varepsilon_{a_{k+1} \cdots a_d b_1 \cdots b_k}\, \Sigma^{b_1 \cdots b_k}$. In particular a hypersurface can be denoted either by $\mathbf{u}_a$ or by $\mathbf{u}^{[a_1 \cdots a_{d-1}]}$.

Hence, up to scale, $\Sigma^{a_1 \cdots a_k}$ and its dual $\Sigma_{a_{k+1} \cdots a_d}$ are intrinsic characteristics of the subspace Σ, independent of the bases chosen to span it and uniquely defined by and defining it. In this sense the antisymmetric tensors provide a sort of coordinate system on the space of linear subspaces of $\mathcal{H}^a$, called **Grassmann coordinates**.

Unfortunately, only very special antisymmetric tensors specify subspaces. The space of k dimensional linear subspaces of a d dimensional vector space is only $k\,(d-k)$ dimensional, whereas the antisymmetric k index tensors have $\binom{d}{k}$ independent components, so the Grassmann coordinates are massively redundant. The tensors that do define subspaces are called **simple** because they satisfy the following complex quadratic **Grassmann relations**:

$$\Sigma^{a_1 \cdots [a_k}\, \Sigma^{b_1 \cdots b_k]} = \mathbf{0}$$

or in terms of the dual

$$\Sigma_{a_{k+1} \cdots a_d}\, \Sigma^{a_d b_2 \cdots b_k} = \mathbf{0}$$

These relations obviously hold for any tensor of the form $\mathbf{v}_1^{[a_1} \cdots \mathbf{v}_k^{a_k]}$ because one of the vectors must appear twice in an antisymmetrization. What is less obvious is that they do not hold for any tensor that can not be written in this form.

Although their redundancy and the complexity of the Grassmann relations makes them rather inconvenient for numerical work, Grassmann coordinates are a powerful tool for the algebraization of geometric operations on subspaces. For example the union of two independent subspaces is just $\Sigma^{[a_1 \cdots a_k} \Gamma^{b_1 \cdots b_l]}$ and dually the intersection of two (minimally) intersecting subspaces is $\Sigma_{[a_1 \cdots a_k} \Gamma_{b_1 \cdots b_l]}$.

A.3. Projective Geometry

Given a $d+1$ dimensional vector space $\mathcal{H}^a$ with nonzero elements $\mathbf{x}^a$ and $\mathbf{y}^a$ ($a = 0, \ldots, d$), we will write $\mathbf{x}^a \sim \mathbf{y}^a$ and say that $\mathbf{x}^a$ and $\mathbf{y}^a$ are *equivalent up to scale* whenever there is a nonzero scalar λ such that $\mathbf{x}^a = \lambda \mathbf{y}^a$. The d dimensional **projective space** $\mathcal{P}^a$ is defined to be the set of nonzero elements of $\mathcal{H}^a$ under equivalence up to scale. When we write $\mathbf{x}^a \in \mathcal{P}^a$ we really mean the equivalence class $\{\lambda\, \mathbf{x}^a \mid \lambda \neq 0\}$ of $\mathbf{x}^a$ under $\sim$.

The span of any $k+1$ independent representatives $\{\mathbf{x}_0^a, \ldots, \mathbf{x}_k^a\}$ of points in $\mathcal{P}^a$ is a $k+1$ dimensional vector subspace of $\mathcal{H}^a$ that projects to a well-defined k dimensional projective subspace of $\mathcal{P}^a$ called the subspace **through** the points. Two independent points define a one dimensional projective subspace called a projective line, three points define a projective plane, and so forth. The vector subspaces of $\mathcal{H}^a$ support notions of subspace dimension, independence, identity, containment, intersection, and union (vector space sum or smallest containing subspace). All of these descend to the projective subspaces of $\mathcal{P}^a$. Similarly, linear mappings between vector spaces, kernels and images, injectivity and surjectivity, and so on all have their counterparts for projective mappings between projective spaces.

Tensors on $\mathcal{H}^a$ also descend to projective tensors defined up to scale on $\mathcal{P}^a$. Elements $\mathbf{u}_a$ of the projective version $\mathcal{P}_a$ of the dual space $\mathcal{H}_a$ define $d-1$ dimensional projective hyperplanes in $\mathcal{P}^a$ via $\mathbf{u}_a \mathbf{x}^a = 0$. The duality of $\mathcal{H}^a$ and $\mathcal{H}_a$ descends to a powerful duality principle between points and hyperplanes on $\mathcal{P}^a$ and $\mathcal{P}_a$.

More generally the antisymmetric $k+1$ index contravariant and $d-k$ index covariant Grassmann tensors on $\mathcal{H}^a$ define k dimensional projective subspaces of $\mathcal{P}^a$. For example given independent points $\mathbf{x}^a$, $\mathbf{y}^a$ and $\mathbf{z}^a$ of $\mathcal{P}^a$ the projective tensor $\mathbf{x}^{[a}\mathbf{y}^{b]}$ defines the line through $\mathbf{x}^a$ and $\mathbf{y}^a$ and $\mathbf{x}^{[a}\mathbf{y}^b\mathbf{z}^{c]}$ defines the plane through $\mathbf{x}^a$, $\mathbf{y}^a$ and $\mathbf{z}^a$. Similarly, in 3D a line can be represented dually as the intersection of two hyperplanes $\mathbf{u}_{[a}\mathbf{v}_{b]}$ while a point requires three $\mathbf{u}_{[a}\mathbf{v}_b\mathbf{w}_{c]}$. In 2D a single hyperplane $\mathbf{u}_a$ suffices for a line, and two are required for a point $\mathbf{u}_{[a}\mathbf{v}_{b]}$. Dualization gives back the contravariant representation, *e.g.* $\mathbf{x}^a = \mathbf{u}_b \mathbf{v}_c\, \varepsilon^{abc}$ are the coordinates of the intersection of the two lines $\mathbf{u}_a$ and $\mathbf{v}_a$ in 2D.

A d dimensional projective space can be thought of as a d dimensional affine space (*i.e.* a Euclidean space with points, lines, planes, and so on, but no origin or notion of absolute distance) with a number of **ideal** points added 'at infinity'. Choosing a basis for $\mathcal{H}^a$, any representative $\mathbf{x}^a$ of an element $\mathcal{P}^a$ with $x^0 \neq 0$ can be rescaled to the form $(1, x^1, \ldots, x^d)^\top$. This defines an inclusion of the affine space $(x^1, \ldots, x^d)$ in $\mathcal{P}^a$, but the $d-1$ dimensional projective subspace 'at infinity' of elements of $\mathcal{P}^a$ with $x^0 = 0$ is not represented. Under this inclusion affine subspaces (lines, planes, etc) become projective ones, and all of affine geometry can be transferred to projective space. However projective geometry is simpler than affine geometry because projective spaces are significantly more uniform than affine ones — there are far fewer special cases to consider. For example two distinct lines always meet exactly once in the projective plane, whereas in the affine plane they always meet *except* when they are parallel. Similarly, there are natural transformations that preserve projective structure (*i.e.* that map lines to lines, preserve intersections and so) that are quite complicated when expressed in affine space but very simple and natural in projective terms. The 3D→2D pinhole camera projection is one of these, hence the importance of projective geometry to computer vision.

Appendix B

Factorization of the Fundamental Matrix

This appendix proves two claims made in section 3.

(1) Given the factorization $\mathbf{F}_{AA'} = \mathbf{u}_A\, \mathbf{v}_{A'} - \mathbf{v}_A\, \mathbf{u}_{A'}$, *the most general redefinition of the* $\mathbf{u}$*'s and* $\mathbf{v}$*'s that leaves* $\mathbf{F}$ *unchanged up to scale is*

$$\begin{pmatrix} \mathbf{u}_A & \mathbf{u}_{A'} \\ \mathbf{v}_A & \mathbf{v}_{A'} \end{pmatrix} \longrightarrow \Lambda \begin{pmatrix} \mathbf{u}_A & \mathbf{u}_{A'} \\ \mathbf{v}_A & \mathbf{v}_{A'} \end{pmatrix} \begin{pmatrix} 1/\lambda & 0 \\ 0 & 1/\lambda' \end{pmatrix}$$

where Λ is an arbitrary nonsingular 2×2 matrix and $\{\lambda, \lambda'\}$ are arbitrary nonzero relative scale factors.

Since $\mathbf{u}_A$ and $\mathbf{v}_A$ are independent epipolar lines and there is only a two parameter family of these, any other choice $\tilde{\mathbf{u}}_A, \tilde{\mathbf{v}}_A$ must be a nonsingular linear combination of these two, and similarly for $\mathbf{u}_{A'}$ and $\mathbf{v}_{A'}$. Hence the only possibilities are:

$$\begin{pmatrix} \mathbf{u}_A \\ \mathbf{v}_A \end{pmatrix} \longrightarrow \Lambda \begin{pmatrix} \mathbf{u}_A \\ \mathbf{v}_A \end{pmatrix}, \quad \begin{pmatrix} \mathbf{u}_{A'} \\ \mathbf{v}_{A'} \end{pmatrix} \longrightarrow \Lambda' \begin{pmatrix} \mathbf{u}_{A'} \\ \mathbf{v}_{A'} \end{pmatrix}$$

for nonsingular 2×2 matrices Λ and Λ'. Then

$$\begin{aligned} \mathbf{F}_{AA'} &= \begin{pmatrix} \mathbf{u}_A & \mathbf{v}_A \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{u}_{A'} \\ \mathbf{v}_{A'} \end{pmatrix} \\ &\longrightarrow \begin{pmatrix} \mathbf{u}_A & \mathbf{v}_A \end{pmatrix} \Lambda^\top \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \Lambda' \begin{pmatrix} \mathbf{u}_{A'} \\ \mathbf{v}_{A'} \end{pmatrix} \end{aligned}$$

Since the covectors $\mathbf{u}_A, \mathbf{v}_A$ and $\mathbf{u}_{A'}, \mathbf{v}_{A'}$ are independent, for $\mathbf{F}$ to remain unchanged up to scale we must have

$$\Lambda^\top \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \Lambda' \sim \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Using the 2×2 matrix identity

$$\Lambda = -\mathrm{Det}(\Lambda) \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \Lambda^{-\top} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

we find that $\Lambda' \sim \Lambda$ up to scale. Defining λ'/λ to reflect the difference in scale, the result follows.

(2) Given any factorization $\mathbf{F}_{AA'} = \mathbf{u}_A\, \mathbf{v}_{A'} - \mathbf{v}_A\, \mathbf{u}_{A'}$ defining a 4D subspace $\mathcal{I}^\alpha$ of $\mathcal{H}^\alpha$ via

$$\begin{pmatrix} \mathbf{u}_A & \mathbf{u}_{A'} \\ \mathbf{v}_A & \mathbf{v}_{A'} \end{pmatrix} \begin{pmatrix} \mathbf{x}^A \\ \mathbf{x}^{A'} \end{pmatrix} = \mathbf{0}$$

and any pair $\{\mathbf{P}^A_a, \mathbf{P}^{A'}_a\}$ of rank 3 projection matrices with distinct centres of projection compatible with $\mathbf{F}_{AA'}$ in the sense that $\mathbf{F}_{AA'}\, \mathbf{P}^A_a \mathbf{P}^{A'}_b\, \mathbf{x}^a \mathbf{x}^b = 0$ for all $\mathbf{x}^a \in \mathcal{H}^a$, there is a fixed rescaling $\{\lambda, \lambda'\}$ that makes $\mathcal{I}^\alpha$ coincide with the image of $\mathcal{H}^a$ under the joint projection $(\lambda\, \mathbf{P}^A_a \;\; \lambda'\, \mathbf{P}^{A'}_a)^\top$.

If the compatibility condition holds for all $\mathbf{x}^a$, the symmetric part of the quadratic form $\mathbf{F}_{AA'}\, \mathbf{P}^A_a \mathbf{P}^{A'}_b$ must vanish. Expanding $\mathbf{F}$ and for clarity defining $\mathbf{u}_a \equiv \mathbf{u}_A\, \mathbf{P}^A_a$, $\mathbf{u}'_a \equiv \mathbf{u}_{A'}\, \mathbf{P}^{A'}_a$, $\mathbf{v}_a \equiv \mathbf{v}_A\, \mathbf{P}^A_a$, and $\mathbf{v}'_a \equiv \mathbf{v}_{A'}\, \mathbf{P}^{A'}_a$ we find:

$$\mathbf{u}_a\, \mathbf{v}'_b + \mathbf{v}'_a\, \mathbf{u}_b - \mathbf{v}_a\, \mathbf{u}'_b - \mathbf{u}'_a\, \mathbf{v}_b = \mathbf{0}$$

Since both projections have rank 3 none of the pulled back covectors $\mathbf{u}_a, \mathbf{u}'_a, \mathbf{v}_a, \mathbf{v}'_a$ vanish, and since the pairs $\mathbf{u}_A \not\sim \mathbf{v}_A$ and $\mathbf{u}_{A'} \not\sim \mathbf{v}_{A'}$ are independent, $\mathbf{u}_a \not\sim \mathbf{v}_a$ and $\mathbf{u}'_a \not\sim \mathbf{v}'_a$ are independent too. Contracting with any vector $\mathbf{x}^a$ orthogonal to both $\mathbf{u}_a$ and $\mathbf{u}'_a$ we find that

$$(\mathbf{v}'_a\, \mathbf{x}^a)\, \mathbf{u}_b - (\mathbf{v}_a\, \mathbf{x}^a)\, \mathbf{u}'_b = \mathbf{0}$$

Either there is some $\mathbf{x}^a$ for which one (and hence both) of the coefficients $\mathbf{v}_a\, \mathbf{x}^a$ and $\mathbf{v}'_a\, \mathbf{x}^a$ are nonzero — which implies that $\mathbf{u}_a \sim \mathbf{u}'_a$ — or both coefficients vanish for all such $\mathbf{x}^a$. But in this case we could conclude that $\mathbf{v}_a$ and $\mathbf{v}'_a$ were in $\mathrm{Span}(\mathbf{u}_a, \mathbf{u}'_a)$ and since $\mathbf{v}_a$ is independent of $\mathbf{u}_a$ and $\mathbf{v}'_a$ of $\mathbf{u}'_a$ that $\mathbf{v}_a \sim \mathbf{u}'_a$ and $\mathbf{v}'_a \sim \mathbf{u}_a$. Substituting back into $\mathbf{F}$ immediately shows that $\lambda\, \mathbf{u}_a \mathbf{u}_b - \lambda'\, \mathbf{v}_a \mathbf{v}_b = \mathbf{0}$ with nonzero λ and λ', and hence that $\mathbf{u}_a \sim \mathbf{v}_a$. So this branch is not possible and we can conclude that for some nonzero λ and λ', $\lambda\, \mathbf{u}_a + \lambda'\, \mathbf{u}'_a = \mathbf{0}$. Similarly, $\mu\, \mathbf{v}_a + \mu'\, \mathbf{v}'_a = \mathbf{0}$ for some nonzero μ and μ'. Substituting back into $\mathbf{F}$ gives $(\lambda/\lambda' - \mu/\mu')\, (\mathbf{u}_a \mathbf{v}_b + \mathbf{v}_a \mathbf{u}_b) = \mathbf{0}$, so up to scale $\{\mu, \mu'\} \sim \{\lambda, \lambda'\}$. The rescaling $\{\mathbf{P}^A_A,\, \mathbf{P}^{A'}_A\} \longrightarrow \{\lambda\, \mathbf{P}^A_A,\, \lambda'\, \mathbf{P}^{A'}_A\}$ then takes the projection of any $\mathbf{x}^a$ to a vector lying in $\mathcal{I}^\alpha$:

$$\begin{aligned} &\begin{pmatrix} \mathbf{u}_A & \mathbf{u}_{A'} \\ \mathbf{v}_A & \mathbf{v}_{A'} \end{pmatrix} \begin{pmatrix} \lambda\, \mathbf{P}^A_a \\ \lambda'\, \mathbf{P}^{A'}_a \end{pmatrix} \mathbf{x}^a \\ &\quad = \begin{pmatrix} \lambda\, \mathbf{u}_a + \lambda'\, \mathbf{u}'_a \\ \lambda\, \mathbf{v}_a + \lambda'\, \mathbf{v}'_a \end{pmatrix} \mathbf{x}^a = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \mathbf{x}^a = \mathbf{0} \end{aligned}$$

Notes

1. *Epipolarity:* $\mathbf{u}_A\, \mathbf{e}^A = 0 = \mathbf{v}_A\, \mathbf{e}^A$ follows from $\mathbf{0} = \mathbf{F}_{AA'}\, \mathbf{e}^A = (\mathbf{u}_A \mathbf{e}^A)\, \mathbf{v}_{A'} - (\mathbf{v}_A \mathbf{e}^A)\, \mathbf{u}_{A'}$, given the independence of $\mathbf{u}_{A'}$ and $\mathbf{v}_{A'}$ for rank 2 $\mathbf{F}$. *Correspondence:* For any $\mathbf{x}^A$ on $\mathbf{u}_A$, $\mathbf{u}_A\, \mathbf{x}^A = 0$ implies that $\mathbf{F}_{AA'}\, \mathbf{x}^A = -(\mathbf{v}_A \mathbf{x}^A)\, \mathbf{u}_{A'} \sim \mathbf{u}_{A'}$.

2. If $\mathbf{x}^{A_1}$ and $\mathbf{x}^{A_2}$ are *not* matching points, the transfer equations trace out an entire line of mutually inconsistent 'solutions' as $[A_2B_2]$ or C_2 vary. For fixed $\mathbf{x}^{A_1}$ and *any* line $\mathbf{l}_{A_2}$ there is a 'solution' $\mathbf{x}^{A_3}(\mathbf{x}^{A_1}, \mathbf{l}_{A_2}) \sim \mathbf{l}_{A_2}\, \mathbf{G}_{A_1}{}^{A_2A_3}\, \mathbf{x}^{A_1}$. This is just the intersection of the image 3 epipolar line of $\mathbf{x}^{A_1}$ with the image 3 epipolar line of the intersection of $\mathbf{l}_{A_2}$ and the image 2 epipolar line of $\mathbf{x}^{A_1}$, *i.e.* the transfer of the only point on $\mathbf{l}_{A_2}$ that *could* be a correct match. In general, as $\mathbf{l}_{A_2}$ traces out the the pencil of lines through $\mathbf{x}^{A_2}$ the corresponding 'solutions' $\mathbf{x}^{A_3}$ trace out the entire epipolar line of $\mathbf{x}^{A_1}$ in image 3. The line of 'solutions' collapses to a point only when $\mathbf{x}^{A_2}$ lies on the epipolar line of $\mathbf{x}^{A_1}$. For reliable transfer the line $\mathbf{l}_{A_2}$ should meet the epipolar line of $\mathbf{x}^{A_1}$ reasonably transversally and if possible should pass close to the image 3 epipole. This can be arranged by projecting the free index C_2 along (an approximation to) the image 3 epipole $\mathbf{e}_3^{A_2}$.

   Similarly, $\mathbf{x}^{A_3}$ could be predicted as the intersection of the epipolar lines of $\mathbf{x}^{A_1}$ and $\mathbf{x}^{A_2}$ in $\mathcal{P}^{A_3}$. This intersection always exists, but it is not structurally meaningful if $\mathbf{x}^{A_1}$ and $\mathbf{x}^{A_2}$ do not correspond. The moral is that it is dangerous to use only *some* of the available equations for transfer.
3. *Proof*: By the rank k condition the vector of minors does not vanish. Adding any $(k+1)^{st}$ row vector $\mathbf{v}$ to the system gives a $(k+1)\times(k+1)$ matrix. By the usual cofactor expansion, the determinant of this matrix is exactly the dot product of $\mathbf{v}$ with the vector of minors. The determinant vanishes when $\mathbf{v}$ is chosen to be any of the existing rows of the matrix, so the minor vector is orthogonal to each row.

References

1. D. Adams. *The Hitchhikers Guide to the Galaxy*. Pan Books, 1979.
2. S. Carlsson. Multiple image invariance using the double algebra. In J. Mundy, A. Zissermann, and D. Forsyth, editors, *Applications of Invariance in Computer Vision*, volume 825 of *Lecture Notes in Computer Science*. Springer-Verlag, 1994.
3. O. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? In G. Sandini, editor, *European Conference on Computer Vision*, Santa Margherita Ligure, Italy, May 1992. Springer-Verlag.
4. O. Faugeras. *Three-dimensional computer vision: a geometric viewpoint*. MIT Press, 1993.
5. O. Faugeras, Q.-T. Luong, and S. J. Maybank. Camera self calibration: Theory and experiments. In G. Sandini, editor, *European Conference on Computer Vision*, Santa Margherita Ligure, Italy, May 1992. Springer-Verlag.
6. O. Faugeras and S. J. Maybank. A theory of self calibration of a moving camera. *Int. J. Computer Vision*, 8(2):123–151, 1992.
7. O. Faugeras and B. Mourrain. On the geometry and algebra of the point and line correspondences between n images. In E. Grimson, editor, *IEEE Int. Conf. Computer Vision*, Cambridge, MA, June 1995.
8. R. Hartley. Lines and points in three views – an integrated approach. In *Proc. Image Understanding Workshop*, Monterey, California, November 1994.
9. R. Hartley, R. Gupta, and T. Chang. Stereo from uncalibrated cameras. In *Proc. Conf. Computer Vision & Pattern Recognition*, pages 761–4, Urbana-Champaign, Illinois, 1992.
10. W. V. D. Hodge and D. Pedoe. *Methods of Algebraic Geometry*, volume 1. Cambridge University Press, 1947.
11. S. Laveau and O. Faugeras. 3d scene representation as a collection of images and fundamental matrices. Technical Report RR-2205, INRIA, Sophia Antipolis, France, February 1994.
12. H. C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293:133–5, 1981.
13. Q.-T. Luong, R. Deriche, O. Faugeras, and T. Papadopoulo. On determining the fundamental matrix: Analysis of different methods and experimental results. Technical Report RR-1894, INRIA, Sophia Antipolis, France, 1993.
14. Q.-T. Luong and T. Viéville. Canonic representations for the geometries of multiple projective views. Technical Report UCB/CSD-93-772, Dept. Electrical Engineering and Computer Science, Berkeley, California, 1993.
15. S. J. Maybank. *Theory of Reconstruction from Image Motion*, volume 28 of *Series in Information Sciences*. Springer-Verlag, 1993.
16. P. F. McLauchlan and D. W. Murray. A unifying framework for structure and motion recovery from image sequences. In E. Grimson, editor, *IEEE Int. Conf. Computer Vision*, Cambridge, MA, June 1995.
17. R. Mohr, B. Boufama, and P. Brand. Accurate relative positioning from multiple images. Technical Report RT-102, LIFIA, INRIA Rhône-Alpes, Grenoble, France, 1993. Submitted to *Journal of Artificial Intelligence*.
18. R. Penrose and W. Rindler. *Spinors and space-time. Vol. 1, Two-spinor calculus and relativistic fields*. Cambridge University Press, 1984.
19. A. Shashua. Algebraic functions for recognition. *IEEE Trans. Pattern Analysis & Machine Intelligence*, 1994.
20. B. Triggs. The Grassmannian geometry of projective reconstruction. In E. Grimson, editor, *IEEE Int. Conf. Computer Vision*, Cambridge, MA, June 1995.
21. B. Triggs. Least squares estimation in projective spaces. To appear, 1995.
22. M. Werman and A. Shashua. Elimination: An approach to the study of 3d-from-2d. 1994.

Bill Triggs grew up near a beach in sunny New Zealand. That's probably why he's such a nice balanced person. Caught practising general relativity without a doctorate, he was transported to post-industrial England, taught punting and fine mathematics and eventually sentenced to six years computational geometry in Oxford University Robotics Group's infamous 'Old Death Row'. Surviving only by an intense personal devotion to applied frisbee dynamics, he fled at the first opportunity and went to earth in a scenic technology park in the French Alpes, masquerading as a humble roboticist. He is known to like tensor calculus, ice cream and running to the top of very steep hills, but is believed to be harmless unless armed with a bicycle...