Reconstruction from Projections Using Grassmann Tensors

Richard I. Hartley[1] and Fred Schaffalitzky[2]

[1] Australian National University and National ICT Australia, Canberra
[2] Australian National University, Canberra.

Abstract. In this paper a general method is given for reconstruction of a set of feature points in an arbitrary dimensional projective space from their projections into lower dimensional spaces. The method extends the methods applied in the well-studied problem of reconstruction of a set of scene points in $\mathcal{P}^3$ given their projections in a set of images. In this case, the bifocal, trifocal and quadrifocal tensors are used to carry out this computation. It is shown that similar methods will apply in a much more general context, and hence may be applied to projections from $\mathcal{P}^n$ to $\mathcal{P}^m$, which have been used in the analysis of dynamic scenes. For sufficiently many generic projections, reconstruction of the scene is shown to be unique up to projectivity, except in the case of projections onto one-dimensional image spaces (lines).

1 Introduction

The bifocal tensor (fundamental matrix), trifocal tensor and quadrifocal tensor have been much studied as a means of reconstucting a 3-dimensional scene from its projection in two, three or four images. It is well known that given sufficiently many point (or line) correspondences between the views, it is possible to compute the multiview tensor and subsequently extract from it the original projection matrices of the cameras, up to an unavoidable projective equivalence. There have been too many papers related to this to cite them all, and so we refer here only to the following papers: [3,2,4]. The methods previously given for extracting the projection matrices from the bifocal, trifocal and quadrifocal tensor have been quite different, and it was not clear that a general method exists.

In work involving the analysis of dynamic scenes, Wolf and Shashua ([11]) have considered projections from higher-dimensional projective spaces $\mathcal{P}^n \rightarrow \mathcal{P}^2$. They showed that such problems can also be studied in terms of tensors, and give some methods for working with these tensors. They do not, however give a general method for defining these tensors, or extracting the projection matrices afterwards. Neither do they consider projections into higher dimensional projective spaces.

At the other end of the scale, Quan and others ([7,1]) have studied projections between low-dimensional spaces, namely projections from $\mathcal{P}^2$ to $\mathcal{P}^1$, and solve the reconstruction problem using a trifocal tensor. Quan shows ([6]) that in this case, there are two possible reconstructions.

T. Pajdla and J. Matas (Eds.): ECCV 2004, LNCS 3021, pp. 363–375, 2004.

This paper unifies all this previous work by showing that reconstruction from projections of $\mathcal{P}^n$ into arbitrary dimensional projective spaces is always possible, and is almost always projectively unique. The method involves a generalization of the multiview tensors for projections $\mathcal{P}^3 \rightarrow \mathcal{P}^2$ (referred to subsequently as the "classical" tensors). The exception is the case where all the projections are onto one-dimensional projective spaces (lines). In this case, two reconstructions are always possible.

The reconstruction method described in this paper involves Grassmann tensors, which relate Grassmann coordinates[1] of linear subspaces in the image. The concept of Grassmann tensor was introduced by Triggs ([9,10])to unify the classical bifocal, trifocal and quadrifocal tensors. The same formalism was taken up by Heyden in a paper exploring the multilinear matching relations ([5]). Triggs paper does not put any restriction on the dimensions of projective spaces considered, though he seems mainly to be concerned with the classical projection $\mathcal{P}^3 \rightarrow \mathcal{P}^2$, and point correspondences. Nevertheless, in [9] he observes that relations exist involving Grassmann coordinates of higher-dimensional subspaces, though he does not pursue the subject. In this paper, we consider this topic in detail, defining a general class of tensors. The main result of this paper, however, is that from any of these tensors the projection matrices may be retrieved, up to projective equivalence. This result does not appear in the papers of Triggs or Heyden.

The Grassmann tensor. We consider a sequence of projections from $\mathcal{P}^n$ to $\mathcal{P}^{m_i}$, for $i = 1, \ldots, r$. Thus, we do not assume that the image space always has the same dimension. For each i, we select integers α_i satisfying $1 \leq \alpha_i \leq m_i$ and $\sum_i \alpha_i = n + 1$. These values represent the codimension of linear subspaces to be specified in each of the image spaces. Thus, when $\alpha_i = m_i$, the linear subspace is a point (dimension 0), and when $\alpha_i = 1$, the linear subspace is a codimension-1 hyperplane. A set of linear subspaces with these dimensions are said to *correspond* when there exists at least one point $\mathbf{X}$ in $\mathcal{P}^n$ that maps via each projection to a point in the given linear subspace in the corresponding image space.

For instance, in the case of the classical trifocal tensor, we say that $\mathbf{x} \leftrightarrow \mathbf{l}' \leftrightarrow \mathbf{l}''$ is a point-line-line correspondence if there exists a point $\mathbf{X}$ in $\mathcal{P}^3$ that maps to $\mathbf{x}$ in the first image, and to points on the lines $\mathbf{l}'$ and $\mathbf{l}''$ in the other two images. The corresponding point and lines satisfy a relationship $\sum_{i,j,k} x^i l'_j l''_k \mathcal{T}_i^{jk} = 0$ ([3]).

In the general case now being considering, there also exists a tensor relating the coordinates of a set of corresponding linear subspaces in the set of images. However, to assign coordinates to linear subspaces of arbitrary dimension, we need to use Grassmann coordinates (described later). Note that for points and lines in $\mathcal{P}^2$, the Grassmann coordinates are nothing more than the homogeneous coordinates of the point or line. It is only when we consider image spaces of higher dimension that the correct generalization in terms of Grassmann coordinates

[1] Grassman coordinates, used in this paper, are also called Plücker coordinates by some authors.

becomes apparent. In this case, the tensor relates the Grassmann coordinates of the corresponding linear subspaces in each image. The relationship is of the form

$$\sum_{\sigma_1,\sigma_2,\ldots\sigma_r} |S^1_{\sigma_1}||S^2_{\sigma_2}|\ldots|S^r_{\sigma_r}|\mathcal{A}_{\sim\sigma_1\sim\sigma_2\ldots\sim\sigma_r} = 0 \quad (1)$$

The notation $|S^i_{\sigma_i}|$ represents the σ_i-th Grassmann coordinate of the subspace S^i. Recall that S^i is a subspace of codimension α_i in $\mathcal{P}^{m_i}$. Consequently, the vector of Grassmann coordinates has dimension $C^{\alpha_i}_{m_i+1}$, and σ_i is an index into this Grassmann vector. The sum is over all combinations of Grassmann coordinates. The notation $\sim\sigma_i$ is to be read "not" σ_i. What this means is not made clear until later, but the reader may safely ignore the $\sim$, for it is only a notational convenience (or perhaps inconvenience). In fact $\mathcal{A}_{\sim\sigma_1\sim\sigma_2\ldots\sim\sigma_r}$ represents a tensor indexed by the indices σ_i.[2] We refer to $\mathcal{A}$ as a *Grassmann tensor.*

Computation of the Grassmann tensor. Given a correspondence between subspaces of codimension α_i in each $\mathcal{P}^{m_i}$, we obtain a single linear relationship between the elements of the Grassmann tensor. It is also possible to obtain relationships involving the tensor in the case of correspondences between subspaces of greater codimension than α_i. In the case of the trifocal tensor, a 3-point correspondence $\mathbf{x} \leftrightarrow \mathbf{x}' \leftrightarrow \mathbf{x}''$ leads to four linear relations. These are obtained by choosing any two lines passing through $\mathbf{x}'$ and any two lines passing through $\mathbf{x}''$. Each choice of lines leads to a point-line-line correspondence, from each of which one obtains a linear relation. This same idea allows us to derive linear relations for Grassmann tensors in higher dimension, given a correspondence betwen subspaces of higher codimension. The exact number of correspondences generated in this way is not explored in this paper, though it is well understood in the $\mathcal{P}^3 \rightarrow \mathcal{P}^2$ case. In any case, given sufficiently many correspondences the Grassmann tensor may be computed linearly.

For clarification, it should be pointed out that for a set of projections $\mathcal{P}^n \rightarrow \mathcal{P}^{m_i}$, there may be many different tensors, depending on the choice of the sequence of codimensions $(\alpha_i, \alpha_2, \ldots, \alpha_r)$. The only restrictions are that $1 \leq \alpha_i \leq m_i$ and $\sum_i \alpha_i = n + 1$. In the well-known case of the trifocal tensor, there are actually three different tensors depending on which of the three images is chosen to have the contravariant index. The three tensors have codimension sequences $(2, 1, 1)$, $(1, 2, 1)$ and $(1, 1, 2)$ respectively. In the general case, we call the sequence of codimensions $(\alpha_1, \alpha_2, \ldots \alpha_r)$ the *profile* of the corresponding tensor. Each such profile corresponds to a different tensor. If we are computing a tensor from point correspondences across several views, then it is necessary to choose in advance which profile to use, since any profile consistent with the dimensions of the image spaces can be used.

Extraction of projection matrices. Having computed a Grassmann tensor from a set of linear subspace correspondences, we now seek to extract the projection matrices. Ad-hoc techniques for computing the projections from multiview tensors have been proposed in the past, both for the standard case of $\mathcal{P}^3 \rightarrow \mathcal{P}^2$ as

[2] In some respects the $\sim$ sign is analogous to the use of upper and lower indices in the classical tensor notation.

well as for higher dimensional cases ([11]). We now give a general procedure for doing this, and show that (at least for generic projections) the projection matrices are determined uniquely by a Grassmann tensor up to projective equivalence, **except** in the case where each $m_i = 1$. In this latter case, there will always be two non-equivalent solution, and indeed this represents a basic ambiguity for projective reconstruction from projections onto lines. This ambiguity persists however many point correspondences are involved. The two projective reconstructions are related to each other by a Cremona transform, which is a non-linear transformation of $\mathcal{P}^n$ ([8]).

The method for computing the projection matrices given the tensor is related to the way a spanning set of vectors for a linear subspace is computed from the Grassmann coordinates. However, it is somewhat more involved. We make no claim that this method is optimum, or even robust in the presence of noise. In fact we are not concerned with noise at all. The present paper provides an existence and uniqueness proof for reconstruction rather than attempting to determine an optimal algorithm. As a deliberate choice, no experimental results will be reported.

1.1 Definition of the Tensors

A mapping from $\mathcal{P}^n$ to $\mathcal{P}^m$ is represented by a matrix of dimension $(m+1)\times(n+1)$, acting on homogeneous coordinates. We consider a set of r such mappings, where the i-th mapping is from $\mathcal{P}^n$ to $\mathcal{P}^{m_i}$. Thus the dimension of the image of this mapping may be different in each case. The matrix representing the i-th mapping will be denoted by $\mathtt{A}^i$

We introduce the concept of an *ordered partition* of $n+1$. This is an ordered tuple of non-negative integers $(\alpha_1, \alpha_2, \ldots, \alpha_r)$ that sum to $n+1$. We are interested in those partitions of n for which each α_i lies in the range 1 to m_i. We will show that for each such ordered partition, there exists an r-view tensor (where r is the length of the partition) relating the coordinates of matched codimension-α_i linear subspaces in r images.

Thus when $n = 3$ and each $m_i = 2$, the possible ordered partitions of $4 = 3+1$ are $(2,2)$, $(2,1,1)$, $(1,2,1)$, $(1,1,2)$ and $(1,1,1,1)$. These partitions correspond to the well-known multiview tensors for 2, 3 and 4 views. We see that there is a bi-focal tensor (the fundamental matrix) corresponding to the partition $(2,2)$, three trifocal tensors corresponding to the three partitions of length 3, and one quadrifocal tensor.

We will call the ordered partition corresponding to a given tensor the *profile* of the tensor. How the tensor with a given profile is defined will now be explained.

Given $d+1$ points spanning a linear subspace of some projective space, we assemble the points as the columns of a matrix $\mathtt{S}$. The linear subspace is simply the span of the columns of $\mathtt{S}$ and any point in this subspace can be written in the form $\mathtt{S}\mathbf{v}$ for some suitable vector $\mathbf{v}$. We may speak of the matrix $\mathtt{S}$ as *representing* the subspace. The condition for a point $\mathbf{X}$ in $\mathcal{P}^n$ to map into the subspace under a mapping represented by $\mathtt{A}$ is that $\mathtt{A}\mathbf{X} - \mathtt{S}\mathbf{v} = 0$ for some $\mathbf{v}$.

Now, choose a set of linear subspaces each of codimension α_i in its projective space $\mathcal{P}^{m_i}$ and let $\mathtt{S}^i$ be the matrix representing the subspace. Suppose that there exists a point $\mathbf{X}$ in $\mathcal{P}^n$ that maps under each projection (represented by $\mathtt{A}^i$) to a point lying in the subspace $\mathtt{S}^i$. It will be shown that this condition implies a single constraint on the set of projection matrices $\mathtt{A}^i$.

The fact that this same $\mathbf{X}$ projects into each of the subspaces may be written in one matrix equation as follows.

$$\begin{bmatrix} \mathtt{A}^1 & \mathtt{S}^1 & & & \\ \mathtt{A}^2 & & \mathtt{S}^2 & & \\ \vdots & & & \ddots & \\ \mathtt{A}^r & & & & \mathtt{S}^r \end{bmatrix} \begin{pmatrix} \mathbf{X} \\ -\mathbf{v}_1 \\ -\mathbf{v}_2 \\ \vdots \\ -\mathbf{v}_r \end{pmatrix} = \mathbf{0} \ . \tag{2}$$

Note that the matrix on the left is square. To check this: the number of rows is equal to $\sum_{i=1}^{r}(m_i+1)$, whereas the number of columns is equal to

$$(n+1) + \sum_{i=1}^{r}(m_i + 1 - \alpha_i) = \sum_{i=1}^{r}(m_i+1) \ \text{ since } \ \sum_{i=1}^{r}\alpha_i = (n+1) \ .$$

In order for a non-zero solution to this set of equations to exist, it is necessary that the determinant of the matrix be zero. If the coordinates of the subspaces (the matrices $\mathtt{S}^i$) are given, then this provides a single constraint on the entries of the matrices $\mathtt{A}^i$. To understand the form of this constraint, we need to expand out this determinant, and to do that, we shall need to use Grassmann coordinates.

Grassmann coordinates. Given a matrix $\mathtt{M}$ with q rows and p columns, where $p \leq q$, we define its *Grassmann coordinates* to be the sequence of determinants of all its $p \times p$ submatrices. It is a well known fact that the Grassmann coordinates of a matrix determine its column span. Alternatively, the Grassmann coordinates determine the matrix up to right-multiplication by a non-singular $p \times p$ matrix with unit determinant. Let σ represent a sequence of p distinct integers in the range 1 to q, in ascending order. Let $|\mathtt{M}_\sigma|$ represent the determinant of the matrix that consists of the rows of $\mathtt{M}$ specified by the sequence σ. Then the values $|\mathtt{M}_\sigma|$, as σ ranges over all such sequences, are the Grassmann coordinates of the matrix.

Now, given such a sequence σ indexing the rows of a matrix, let $\sim\sigma$ represent the sequence consisting of those integers not in σ. and define $\text{sign}(\sigma)$ to be $+1$ or -1 depending on whether the concatenated sequence $\sigma \sim \sigma$ is an even or odd permutation.[3] Thus, for example the sequence 125 has sign $+1$, since 12534 is an even permutation.

Given a square matrix divided into two blocks, for instance $[\mathtt{A}|\mathtt{B}]$, its determinant may be expressed in terms of the Grassmann coordinates of $\mathtt{A}$ and $\mathtt{B}$. In particular

$$|\mathtt{A}|\mathtt{B}| = \sum_{\sigma} \text{sign}(\sigma) \|\mathtt{A}_\sigma\| \ \|\mathtt{B}_{\sim\sigma}\|$$

[3] A permutation is called even or odd, according to whether it is the product of an even or odd number of pairwise swaps.

where the sum is over all ascending sequences σ of length equal to the number of columns of $\mathtt{A}$. The particular case where $\mathtt{A}$ consists of a single column is just the familiar cofactor expansion of the determinant.

Using this factorization, one may derive a precise formula for the determinant of the matrix on the left of (2), namely

$$\pm \sum_{\sigma_1,\sigma_2,\ldots,\sigma_r} \mathrm{sign}(\sigma_1)\ldots\mathrm{sign}(\sigma_r)\ |\mathtt{A}_{\sim\sigma_1\sim\sigma_2\ldots\sim\sigma_r}|\ |\mathtt{S}^1_{\sigma_1}||\mathtt{S}^2_{\sigma_2}|\ldots|\mathtt{S}^r_{\sigma_r}|\ . \quad (3)$$

In this formula, each σ_i is an ordered sequence of integers in the range 1 to m_i+1, the length of the sequence being equal to the dimension of the subspace $\mathtt{S}^i$. Further, $|\mathtt{A}_{\sim\sigma_1\sim\sigma_2\ldots\sim\sigma_r}|$ is the determinant of the matrix obtained by selecting the rows indexed by $\sim\sigma_i$ (that is, omitting the rows indexed by σ_i) from each $\mathtt{A}^i$. The overall sign (whether + or −) does not concern us. The set of values

$$\mathcal{A}_{\sim\sigma_1\sim\sigma_2\ldots\sim\sigma_r} = \mathrm{sign}(\sigma_1)\ldots\mathrm{sign}(\sigma_r)\ |\mathtt{A}_{\sim\sigma_1\sim\sigma_2\ldots\sim\sigma_r}|$$

forms an r dimensional array whose elements are (up to sign) minors of the matrix $\mathtt{A}$ obtained by stacking the projection matrices $\mathtt{A}^i$. The only minors are ones corresponding to submatrices of $\mathtt{A}$, in which α_i rows are chosen from each $\mathtt{A}^i$. Recalling that the sequence $(\alpha_1,\ldots,\alpha_r)$ in which $\sum_{i=1}^r \alpha_i = n+1$ is called a *profile*, we will call the array $\mathcal{A}_{\sim\sigma_1\sim\sigma_2\ldots\sim\sigma_r}$ the *Grassmann tensor* corresponding to the profile $(\alpha_1,\ldots,\alpha_r)$.

The tensor $\mathcal{A}$ gives a linear relationship between the Grassmann coordinates of linear subspaces defined in each of the image spaces $\mathcal{P}^{m_i}$:

$$\sum_{\sigma_1,\sigma_2,\ldots,\sigma_r} \mathcal{A}_{\sim\sigma_1\sim\sigma_2\ldots\sim\sigma_r}\ |\mathtt{S}^1_{\sigma_1}||\mathtt{S}^2_{\sigma_2}|\ldots|\mathtt{S}^r_{\sigma_r}| = 0\ . \quad (4)$$

This relationship generalizes the classical bifocal and trifocal relations ([3]). The classical tensors involve relations between point and line coordinates in $\mathcal{P}^2$. However, the Grassmann coordinates of a single point (a 0-dimensional linear space) are simply the homogeneous coordinates of the point. Similarly, for a line in $\mathcal{P}^2$, the Grassmann coordinates are the same as the homogeneous coordinates, except for sign.

Given a change of coordinates in some of the image spaces $\mathcal{P}^{m_i}$, the tensor $\mathcal{A}$ does not in general transform strictly as a contravariant or covariant tensor. Rather, it transforms according to the inverse of the corresponding transformation of Grassmann coordinates induced by the change of coordinates. This map is the $m_i+1-\alpha_i$-fold exterior product mapping induced by the coordinate change.

2 Solving for the Projection Matrices

We now consider the problem of determining the projection matrices from a Grassmann tensor. As in the standard case of 3D reconstruction from uncalibrated image measurements, we can not expect to determine the projection matrices more exactly than up to projectivity. In addition, since the projection matrices are homogeneous objects, their overall scale is indeterminate. Thus, we make the following definition:

Definition 1. *Two sequences of projection matrices* $(\mathtt{A}^1, \ldots, \mathtt{A}^r)$ *and* $(\hat{\mathtt{A}}^1 \ldots, \hat{\mathtt{A}}^r)$ *are projectively equivalent if there exists an invertible matrix* $\mathtt{H}$ *as well as scalars* λ_i *such that* $\hat{\mathtt{A}}^i = \lambda_i \mathtt{A}^i \mathtt{H}$ *for all* i.

Now, let $\mathtt{A}$ be formed by stacking all the $\mathtt{A}^i$ on top of each other, resulting in a matrix of dimension $(\sum_{i=1}^r (m_i + 1)) \times (n + 1)$. We accordingly associate the matrices $\mathtt{A}^i$ with successive vertically stacked blocks of the matrix $\mathtt{A}$. Corresponding to definition 1, we may define an equivalence relation on matrices with this block structure, as follows.

Definition 2. *Two matrices* $\mathtt{A}$ *and* $\hat{\mathtt{A}}$ *made up of blocks* $\mathtt{A}^i$ *and* $\hat{\mathtt{A}}^i$ *of dimension* $(m_i + 1) \times (n + 1)$ *are* block projectively equivalent *if there exists an invertible* $(n+1)\times(n+1)$ *matrix* $\mathtt{H}$*, and scalar matrices* $\lambda_i \mathtt{I}_i$ *of dimension* $(m_i+1)\times(m_i+1)$ *such that*

$$\hat{\mathtt{A}} = \mathrm{diag}(\lambda_1 \mathtt{I}_1, \ldots, \lambda_r \mathtt{I}_r)\mathtt{A}\mathtt{H} \ .$$

It is easily seen that this definition is equivalent to the projective equivalence of the sequences of matrices $(\mathtt{A}^1, \ldots, \mathtt{A}^r)$ and $(\hat{\mathtt{A}}^1 \ldots, \hat{\mathtt{A}}^r)$ as stated in definition 1. It is evident that this is an equivalence relation on matrices with this given block structure.

2.1 Partitions and Determinants

Now, we assume that there are sufficiently many such projections that $\sum_{i=1}^r m_i \geq n$. Let $(\alpha_1, \ldots, \alpha_r)$ be an ordered partition of $(n + 1)$ with the property that $1 \leq \alpha_i \leq m_i$. We may form square matrices of dimension $(n+1)\times(n+1)$ by selecting exactly α_i rows from each matrix $\mathtt{A}^i$. We may then take the determinant of such a square matrix. Of course, we may select α_i rows from each $\mathtt{A}^i$ in many different ways – to be exact, there are $\prod_{i=1}^r C_{m_i+1}^{\alpha_i}$ ways of doing this, and that many such subdeterminants of $\mathtt{A}$ corresponding to the given partition.

Before giving the main theorem, we state our assumption of genericity. All projections from $\mathcal{P}^n$ to $\mathcal{P}^m$ are assumed to be "generic", which means in effect that improbable special cases are ruled out. Any polynomial expression in the coordinates of the matrix representation of the projections, or related points may be assumed to be non-zero, unless it is always zero. Thus the results we prove will hold, except on a set of measure zero. We now state the main theorem of this part of the paper.

Theorem 1. *Let* $\mathtt{A}$ *be a generic matrix with blocks* $\mathtt{A}^i; i = 1, \ldots, r$ *of dimension* $(m_i + 1) \times (n + 1)$*, and let* $(\alpha_1, \ldots, \alpha_r)$ *be any fixed ordered partition of* $n + 1$*. If at least one* m_i *is greater than one, then the matrix* $\mathtt{A}$ *is determined up to block projective equivalence by the collection of all its minors, chosen with* α_i *rows from each* $\mathtt{A}^i$*. If all* $m_i = 1$*, then there are two equivalence classes of solutions.*

We refer to the partition $(\alpha_1, \ldots, \alpha_r)$ as the *profile* of the minors. Thus, the theorem states that the matrix $\mathtt{A}$ is determined up to projective equivalence by its collection of minors with a given fixed profile.

Proof. We would like to give a more leisurely proof, with examples, but are prevented by lack of space. The proof given below is complete, but telegraphic. Let $\mathtt{A}$ and $\hat{\mathtt{A}}$ be two matrices with corresponding blocks $\mathtt{A}^i$ and $\hat{\mathtt{A}}^i$, each of which gives rise to the same collection of minors. Our goal is to show that the two matrices are block projective-equivalent, which means that $\hat{\mathtt{A}}^i = \lambda_i \mathtt{A}^i \mathtt{H}$ for some choice of $\mathtt{H}$ and λ_i. The strategy of the proof is to apply a sequence of transformations to $\mathtt{A}$ (and to $\hat{\mathtt{A}}$), each transformation replacing $\mathtt{A}$ by a projectively equivalent matrix, until eventually $\mathtt{A}$ and $\hat{\mathtt{A}}$ become identical. This will demonstrate the projective equivalence of the original matrices.

By assumption, there exists at least one non-zero minor, and without loss of generality (by rearranging rows if required), this may be chosen to belong to the submatrix of $\mathtt{A}$ in which the *first* α_i rows are chosen from each $\mathtt{A}^i$. Let this submatrix be denoted by $\mathtt{G}$. Choosing $\mathtt{H} = \mathtt{G}^{-1}$, we may replace $\mathtt{A}$ by an equivalent matrix $\mathtt{AH}$ in which the matrix $\mathtt{G}$ is replaced by the identity matrix. Doing the same thing to $\hat{\mathtt{A}}$, we may assume that both $\mathtt{A}$ and $\hat{\mathtt{A}}$ have this simple form.

After this transformation, the form of the matrix $\mathtt{A}$ is somewhat simplified. The first α_i rows from each block are known, consisting of zeros, except for one unit element in each such row. We refer to these rows of $\mathtt{A}$ as the *reduced* rows. The elements of the remaining rows of $\mathtt{A}$ are still to be determined. We show that they can be determined (up to block projective equivalence) from other minors of the matrix.

We consider a finer block decomposition of the matrix $\mathtt{A}$ into blocks indexed by (i, j) where the block $\mathtt{A}^{ij}$ has dimension $(m_i + 1) \times \alpha_j$ as shown:

$$\begin{bmatrix} \mathtt{A}^{11} & \ldots & \mathtt{A}^{1r} \\ \vdots & \ddots & \vdots \\ \mathtt{A}^{r1} & \ldots & \mathtt{A}^{rr} \end{bmatrix} . \tag{5}$$

The first α_i rows of each such $\mathtt{A}^{ij}$ are reduced, so

$$\mathtt{A}^{ii} = \begin{bmatrix} \mathtt{I} \\ \mathtt{B}^{ii} \end{bmatrix} \quad \text{and} \quad \mathtt{A}^{ij} = \begin{bmatrix} \mathtt{0} \\ \mathtt{B}^{ij} \end{bmatrix} \text{ for } i \neq j$$

The reduced rows of $\mathtt{A}$ form an identity matrix, having unit determinant. Let $\mathtt{B}$ be the matrix obtained from $\mathtt{A}$ by removing the reduced rows. Then $\mathtt{B}$ has the same type of block structure as $\mathtt{A}$. We investigate the relationship between minors[4] of $\mathtt{A}$ and those of $\mathtt{B}$.

Consider a submatrix of $\mathtt{A}$ chosen according to a given profile $(\alpha_1, \ldots, \alpha_r)$. Some of the rows of this submatrix will be rows of $\mathtt{B}$, while other will be chosen

[4] For brevity, when we speak of the minors of $\mathtt{A}$, we mean those chosen according to the given profile, with α_i rows from each $\mathtt{A}^i$.

from among the reduced rows of $\mathtt{A}$. A reduced row is one in which there is one unit (1) entry and the rest are zero. In computing the determinant, we may strike out any reduced rows, as well as the columns containing the unit entries, resulting in a smaller matrix containing only elements belonging to rows from $\mathtt{B}$. The columns that remain are ones that did not have a 1 in any reduced row of the chosen submatrix. Here is the key observation: if a row is chosen from the i-th block of rows $[\mathtt{B}^{i1}\ldots\mathtt{B}^{ir}]$ then some reduced row from the same numbered blocks $[\mathtt{A}^{i1}\ldots\mathtt{A}^{ir}]$ must be absent. Such a row has its unit element in the block $\mathtt{A}^{ii}$, but this row is not present in the chosen submatrix. The corresponding column, belonging to the i-th block of columns, must therefore "survive" when rows and columns corresponding to the reduced rows are struck out. This means:

The minors of $\mathtt{A}$ *are in one-to-one correspondence with (and equal up to sign to) the minors of* $\mathtt{B}$ *chosen in the following manner:* β_i *rows are chosen from the i-th block of rows of* $[\mathtt{B}^{i1}\ldots\mathtt{B}^{ir}]$ *and* β_i *columns from the i-th block of columns, containing the blocks* $\mathtt{B}^{1i}\ldots\mathtt{B}^{ri}$*. Here the* β_i *are integers in the range* $0\leq\beta_i\leq\alpha_i$*.*

Such minors of $\mathtt{B}$ will be called "symmetrically chosen" minors. The minors of $\mathtt{A}$ and $\mathtt{B}$ are equal only up to sign, because of the order of the rows, but the sign correspondence is well determined, so that if one knows the values of the minors of $\mathtt{A}$, then the symmetrically chosen minors or $\mathtt{B}$ are also known. We will show that $\mathtt{B}$ is determined by its symmetrically chosen minors, and hence by the minors of $\mathtt{A}$. Therefore, knowing the minors of $\mathtt{A}$, we know $\mathtt{B}$ and hence $\mathtt{A}$, up to projective equivalence. This would complete the proof.

The exact truth is slightly more complicated. We define a different type of equivalence relation on block matrices of the form $\mathtt{B}=[\mathtt{B}^{ij}]$, where $i,j=1,\ldots,r$.

Definition 3. *Two matrices* $\mathtt{B}=[\mathtt{B}^{ij}]$ *and* $\hat{\mathtt{B}}=[\hat{\mathtt{B}}^{ij}]$ *will be called* bilinearly equivalent *if there exist non-zero scalars* λ_i *such that* $\hat{\mathtt{B}}^{ij}=\lambda_i\lambda_j^{-1}\mathtt{B}^{ij}$*.*

The truth is that the symmetrically chosen minors of $\mathtt{B}$ determine $\mathtt{B}$ up to bilinear equivalence. This is sufficient, however, since if $\mathtt{B}$ and $\hat{\mathtt{B}}$ are bilinearly equivalent, then the corresponding matrices $\mathtt{A}^i$ and $\hat{\mathtt{A}}^i$ are projectively equivalent, which is all we need. This is true because

$$\hat{\mathtt{A}}^i=\lambda_i\mathtt{A}^i\mathrm{diag}(\lambda_1^{-1}\mathtt{I}_{\alpha_1},\ldots,\lambda_r^{-1}\mathtt{I}_{\alpha_r})$$

follows from the fact that $\hat{\mathtt{B}}^{ij}=\lambda_i\lambda_j^{-1}\mathtt{B}^{ij}$.

The proof of Theorem 1 will be completed therefore by proving the following lemma.

Lemma 1. *A matrix* $\mathtt{B}=[\mathtt{B}^{ij}]$ *is determined up to bilinear equivalence by its collection of symmetrically chosen minors.*

In fact, it will be sufficient only to consider only 3×3 minors, or 2×2 minors in the two-view case. The proof will proceed in three steps.

1. The 1×1 minors determine the elements of the diagonal blocks $\mathtt{B}^{ii}$.
2. The 2×2 minors determine symmetrically opposite pairs of blocks $\mathtt{B}^{ij}$ and $\mathtt{B}^{ji}$ up to a pair of inverse scalar multiples.
3. The 3×3 minors determine consistent scale factors.

Step 1 - the 1×1 *minors.* The 1×1 symmetrically chosen minors are nothing other than the elements of the diagonal blocks $\mathtt{B}^{ii}$, and hence these 1×1 minors determine the diagonal blocks.

Step 2 - the 2×2 *minors.* A 2×2 symmetrically chosen minor will be of the form $[a\ b; c\ d]$, where a and d are from diagonal blocks $\mathtt{B}^{ii}$ and $\mathtt{B}^{jj}$, and hence are known from the previous step. Elements b and c are from the symmetrically opposite blocks $\mathtt{B}^{ij}$ and $\mathtt{B}^{ji}$. Since the determinant is $ad - bc$ with ad known, we may obtain the value of bc from the value of the 2×2 minor. In fact, by chosing the right minor, we can determine the product of any two elements chosen from symmetrically opposite blocks such as $\mathtt{B}^{ij}$ and $\mathtt{B}^{ji}$.

Let $\mathbf{b}$ be the vector consisting of all elements from the block $\mathtt{B}^{ij}$ and $\mathbf{c}$ be the vector of elements of $\mathtt{B}^{ji}$. Then the set of all products v_{rs} of elements from blocks $\mathtt{B}^{ij}$ and $\mathtt{B}^{ji}$ can determined and written $b_r c_s = v_{rs}$. This means that the values v_{rs} form a rank-1 matrix that factors as $\mathtt{V} = \mathbf{b}\mathbf{c}^\top$. The factorization is easily carried out using the Singular Value Decomposition, or some more simple method[5].

Solution of the equations $b_r c_s = w_{rs}$ is only possible up to an indeterminate scale factor. Thus, we may multiply each b_r by λ and c_s by λ^{-1} with the same result, but this is the only possible ambiguity. The result of this is that one may determine the blocks $\mathtt{B}^{ij}$ and $\mathtt{B}^{ji}$ of the matrix $\mathtt{B}$ up to multiplication by inverse scalar factors.

Let $\mathtt{B} = [\mathtt{B}^{ij}]$ and $\hat{\mathtt{B}} = [\hat{\mathtt{B}}^{ij}]$ be two sets of matrices having the same collection of symmetrically chosen minors. Our goal is to show that $\hat{\mathtt{B}}^{ij} = \lambda_i \lambda_j^{-1} \mathtt{B}^{ij}$ for all i, j. On the other hand, what we have shown so far is that there exist non-zero constants μ_{ij} with $\mu_{ii} = 1$ and $\mu_{ij} = \mu_{ji}^{-1}$ such that $\hat{\mathtt{B}}^{ij} = \mu_{ij} \mathtt{B}^{ij}$. It remains to show that μ_{ij} can be written as $\lambda_i \lambda_j^{-1}$. At this point, we modify the matrix $\mathtt{B}$ by multiply each block $\mathtt{B}^{ij}$ by $\mu_{1i} \mu_{1j}^{-1}$. This operation transforms $\mathtt{B}$ to another matrix that is bilinearly equivalent to it, and it is sufficient to prove that the new $\mathtt{B}$ thus obtained is equivalent to $\hat{\mathtt{B}}$. Note however that because of this modification to the matrix $\mathtt{B}$, the first row block and column block of $\mathtt{B}$ and $\hat{\mathtt{B}}$ are identical. Thus, in particular $\hat{\mathtt{B}}^{ij} = \mu_{ij} \mathtt{B}^{ij}$, and $\mu_{i1} = \mu_{1i} = 1$ for all i.

Step 3 - consistent choice of scale factors λ*.* The proof will now be completed by proving that $\mu_{ij} = 1$ for all i, j.

Consider allowable 3×3 subdeterminants of $\mathtt{B}$, in which one row is taken from the first row block, and one row each from each of two other row blocks. Corresponding columns are chosen from the corresponding blocks. The submatrix of $\mathtt{B}$ is

$$\begin{bmatrix} a & b & c \\ d & e & \mu f \\ g & \mu^{-1} h & k \end{bmatrix} \tag{6}$$

and the submatrix of $\hat{\mathtt{B}}$ is the same in which $\mu = 1$. Equating the two determinants gives an equation of the form $A\mu + B + C\mu^{-1} = A + B + C$ for constants

[5] The only thing that can go wrong here is that all w_{ij} are zero, but this a non-generic case.

A, B and C. Multiplying by μ gives a quadratic equation with two solutions: $\mu = 1$ and $\mu = C/A$. In terms of the entries of the matrix, the second solution is $\mu = (dhc)/(bfg)$. Thus, there are two possible solutions for μ. However, in most situations we may obtain a second equation for μ which will also have two solutions, but only the solution $\mu = 1$ will be common to both equations, and hence a solution to the complete system.

To see this, we need to make an assumption that the first projection matrix $\mathtt{A}^1 = [A^{11} \ \ldots \ A^{ir}]$ has more than two rows – its dimension is $m_1 + 1 \geq 3$. This is possible without loss of generality provided that there exists at least one projection matrix with at least three rows, for it may be chosen as the first. The number of rows chosen from $\mathtt{A}^1$ is α_1 which is in the range $1 \leq \alpha_1 \leq m_1$. Now, suppose that the rows and columns of (6) are chosen from the row and column blocks numbered 1, i and j. Thus, the entries of (6) are drawn from the block matrix

$$\begin{bmatrix} B^{11} & B^{1i} & B^{1j} \\ B^{i1} & B^{ii} & B^{ij} \\ B^{j1} & B^{ji} & B^{jj} \end{bmatrix} \quad . \tag{7}$$

Now, the dimension of the matrix B^{ij} is $(m_i + 1 - \alpha_i) \times \alpha_j$. Specifically, $\mathtt{B}^{i1}$ has dimension $(m_i + 1 - \alpha_i) \times \alpha_1$, and $\mathtt{B}^{1j}$ has dimension $(m_1 + 1 - \alpha_1) \times \alpha_j$. However, since $1 \leq \alpha_1 \leq m_1$ and $m_1 > 1$, it must follow that either $\alpha_1 > 1$ or $m_1 + 1 - \alpha_1 > 1$. Thus, either $\mathtt{B}^{i1}$ has at least two columns, or $\mathtt{B}^{1j}$ has at least two rows. In either case, there is more than one way of selecting rows and columns from (7) to obtain a submatrix of the form (6). Each such choice will give a different equation for μ. The solution $\mu = 1$ will be common to both equations whereas the second solution $\mu = (dhc)/(bfg)$ will be different for the two cases, since generically the entries of the matrix (6) will be different for the different choices of rows and columns.

This completes the proof, since we have shown that the only value of μ_{ij} that is consistent with the assumed equality of all the allowable minors of $\mathtt{B}$ and $\hat{\mathtt{B}}$ is that $\mu_{ij} = 1$ for all i, j. Hence, $\mathtt{B} = \hat{\mathtt{B}}$.

2.2 Two Solutions in Minimal Case

It was seen that the case where all projection matrices have only two rows is a special case in which we can not find two equations for each value μ_{ij}. In such a case it is possible that there will be two possible solutions for matrices with the same set of minors. We will investigate this further, and show that in this case, generically there are indeed two solutions.

Thus, let the projection matrices $\mathtt{A}^i$ each have dimension $2 \times (n+1)$, representing a projection from an n-dimensional projective space onto a projective line. In order for us to form square submatrices by choosing one row from each such $\mathtt{A}^i$ there must be exactly $n+1$ such projections. Thus, $\mathtt{A}$ has dimension $2(n+1) \times (n+1)$ and each $\alpha_i = 1$.

With the same argument as before, we may assume that the first rows of each of the $\mathtt{A}^i$ form an identity matrix of dimension $(n+1) \times (n+1)$. Deleting these rows, we obtain an $(n+1)\times(n+1)$ matrix $\mathtt{B}$. In this case, each $\mathtt{B}^{ij}$ consists of a single element. We consider symmetrically chosen submatrices of $\mathtt{B}$. In this case, the summetrically chosen submatrices are those for which the indices of the selected rows and columns are the same. Such a submatrix is chosen by selecting the rows and columns numbered by a sequence of indices $(i_1, i_2, \ldots, i_r)$. The key observation here is that the minors of $\mathtt{B}$ and its transpose $\mathtt{B}^\top$ corresponding to a given sequence of indices are the same, because the determinant of a matrix and its transform are the same. In other words, it is impossible to distinguish between the matrix $\mathtt{B}$ and its transpose on the basis of such symmetrically chosen minors.

Referring this back to the original projection matrices we obtain two matrices $\mathtt{A}$ and $\hat{\mathtt{A}}$ which can not be distinguished by their minors with profile $(1,\ldots,1)$. We may write a specific example as follows: Let

$$\mathtt{A}^1 = \begin{bmatrix} 1 & 0 & 0 \\ a & b & c \end{bmatrix} \quad ; \quad \mathtt{A}^2 = \begin{bmatrix} 0 & 1 & 0 \\ d & e & f \end{bmatrix} \quad ; \quad \mathtt{A}^3 = \begin{bmatrix} 0 & 0 & 1 \\ g & h & j \end{bmatrix}$$

and

$$\hat{\mathtt{A}}^1 = \begin{bmatrix} 1 & 0 & 0 \\ a & d & g \end{bmatrix} \quad ; \quad \hat{\mathtt{A}}^2 = \begin{bmatrix} 0 & 1 & 0 \\ b & e & h \end{bmatrix} \quad ; \quad \hat{\mathtt{A}}^3 = \begin{bmatrix} 0 & 0 & 1 \\ c & f & j \end{bmatrix} .$$

The two matrices $\mathtt{A}$ and $\hat{\mathtt{A}}$ corresponding to these projection matrices can not be distinguished based on the eight minors formed by choosing one row from each $\mathtt{A}^i$, or respectively $\hat{\mathtt{A}}^i$. In terms of tensors, this means that the "trifocal tensors" corresponding to these triples of cameras are the same. Geometrically, this means that there are two possible reconstructions of a (planar) scene based on its projection onto three lines in the plane. The 1-dimensional trifocal tensor was studied by Quan and others in ([7,1]). The observation that there were two solutions in the case of projections from $\mathcal{P}^2$ to $\mathcal{P}^1$ was made in [6]. The ambiguity holds in higher dimensions also, as the above argument shows. Specifically, the tensor (collection of minors) corresponding to $n+1$ projections from $\mathcal{P}^n$ onto a projective line determines the set of projection matrices only up to 2-fold projective ambiguity. Consequently there are always two reconstructions possible from the projection of $\mathcal{P}^n$ onto $(n+1)$ projective lines.

Are there generically only two solutions? In the case where each $m_i = 1$, it may seem possible that there are more than two possibilities (up to bilinear equivalence) for the matrix $\mathtt{B}$ based on its set of minors. However, this is not possible. If we assume that two matrices $\mathtt{B}$ and $\hat{\mathtt{B}}$ have the same set of symmetric minors, then by carrying through the previous arguments, we find that $\mathtt{B}$ may be replaced by an equivalent matrix for which $\mathtt{B}$ and $\hat{\mathtt{B}}$ have the same first row. In addition, there exist constants μ_{ij} such that $\mu_{ij} = \mu_{ji}^{-1}$ and $\mathtt{B}^{ij} = \mu_{ij}\hat{\mathtt{B}}^{ij}$. By considering 3×3 symmetric minors containing the first row and column as in the proof of lemma 1 we obtain a single quadratic equation for each of the constants μ_{ij}. There are two choices. For each (i,j) we may choose $\mu_{ij} = 1$, or else we must choose the other non-unit solution at each position. It may be seen that once one value μ_{ij} with $i,j>1$ is chosen to equal 1, then they all must be. The details are tedious, and omitted. The solution in which μ_{ij} is taken to be the non-unit solution for each i, j may be verified to be equivalent (under bilinear equivalence) to the transposed solution in which $\mathtt{B} = \hat{\mathtt{B}}^\top$.

3 Conclusion

The classical multiview tensor extend to higher dimensions, and allow reconstruction of the scene from projections in any dimension. The solution is unique except in the case of projections onto lines. The work of Wolf and Shasha shows the importance of higher dimensional projections, and provides a potential application for this work, at least in proving the feasibility of a solution.

References

1. O. D. Faugeras, L. Quan, and P. Sturm. Self-calibration of a 1D projective camera and its application to the self-calibration of a 2D projective camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1179–1185, October 2000.
2. R. I. Hartley. Computation of the quadrifocal tensor. In *Proc. 5th European Conference on Computer Vision, Freiburg, Germany*, LNCS 1406, pages 20–35. Springer-Verlag, 1998.
3. R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
4. A. Heyden. Reduced multilinear constraints: Theory and experiments. *International Journal of Computer Vision*, 30(1):5–26, 1998.
5. A. Heyden. Tensorial properties of multilinear constraints. *Mathematical Methods in the Applied Sciences*, 23:169–202, 2000.
6. L. Quan. Two-way ambiguity in 2d projective reconstruction from three uncalibrated 1d images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(2):212–216, 2001.
7. L. Quan and T. Kanade. Affine structure from line correspondences with uncalibrated affine cameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(8):834–845, August 1997.
8. J. G. Semple and G. T. Kneebone. *Algebraic Projective Geometry*. Oxford University Press, 1979.
9. W. Triggs. The geometry of projective reconstruction i: Matching constraints and the joint image. *Unpublished: Available on Bill Triggs's web-site*, 1995.
10. W. Triggs. Matching constraints and the joint image. In E. Grimson, editor, *Proc. 5th International Conference on Computer Vision, Boston*, pages 338 – 343, Cambridge, MA, June 1995.
11. L. Wolf and A. Shashua. On projection matrices $\mathcal{P}^k \to \mathcal{P}^2, k = 3, \ldots, 6$, and their applications in computer vision. *International Journal of Computer Vision*, 48(1):53–67, 2002.