



Chapter 6

Reconstruction from Multiple Views

Multiple View Geometry
Summer 2025

From Two Views to
Multiple Views

From Preimages to
Rank Constraints

Rank Constraints for
Points

The Multiple-view
Matrix

Relation to Epipolar
Constraints

Multiple-View
Reconstruction
Algorithms

Rank Constraints for
Lines

Summary

Prof. Daniel Cremers
Chair of Computer Vision and Artificial Intelligence
School of Computation, Information and Technology
Technical University of Munich



- ➊ From Two Views to Multiple Views
- ➋ From Preimages to Rank Constraints
- ➌ Rank Constraints for Points
- ➍ The Multiple-view Matrix
- ➎ Relation to Epipolar Constraints
- ➏ Multiple-View Reconstruction Algorithms
- ➐ Rank Constraints for Lines
- ➑ Summary

From Two Views to
Multiple Views

From Preimages to
Rank Constraints

Rank Constraints for
Points

The Multiple-view
Matrix

Relation to Epipolar
Constraints

Multiple-View
Reconstruction
Algorithms

Rank Constraints for
Lines

Summary

Multiple-View Geometry

In this section, we deal with the problem of 3D reconstruction given **multiple** views of a static scene, either obtained simultaneously, or sequentially from a moving camera.

The key idea is that the three-view scenario allows to obtain more measurements to infer the same number of 3D coordinates. For example, given two views of a single 3D point, we have four measurements (x- and y-coordinate in each view), while the three-view case provides 6 measurements per point correspondence. As a consequence, the estimation of motion and structure will generally be more constrained when reverting to additional views.

The three-view case has traditionally been addressed by the so-called **trifocal tensor** [Hartley '95, Vieville '93] which generalizes the fundamental matrix. This tensor – as the fundamental matrix – does not depend on the scene structure but rather on the inter-frame camera motion. It captures a **trilinear relationship** between three views of the same 3D point or line [Liu, Huang '86, Spetsakis, Aloimonos '87].



Trifocal Tensor versus Multiview Matrices

Traditionally the trilinear relations were captured by generalizing the concept of the Fundamental Matrix to that of a Trifocal Tensor. It was developed among others by [Liu and Huang '86], [Spetsakis, Aloimonos '87]. The use of tensors was promoted by [Vieville '93] and [Hartley '95]. Bilinear, trilinear and quadrilinear constraints were formulated in [Triggs '95]. This line of work is summarized in the books:

Faugeras and Luong, “The Geometry of Multiple Views”, 2001 and

Hartley and Zisserman, “Multiple View Geometry”, 2001, 2003.

In the following, however, we stick with a matrix notation for the multiview scenario. This approach makes use of matrices and rank constraints on these matrices to impose the constraints from multiple views. Such rank constraints were used by many authors, among others in [Triggs '95] and in [Heyden, Åström '97]. This line of work is summarized in the book

Ma, Soatto, Kosecka, Sastry, “An Invitation to 3D Vision”, 2004.





A **preimage of multiple images** of a point or a line is the (largest) set of 3D points that gives rise to the same set of multiple images of the point or the line.

For example, given the two images ℓ_1 and ℓ_2 of a line L , the preimage of these two images is the intersection of the planes P_1 and P_2 , i.e. exactly the 3D line $L = P_1 \cap P_2$.

In general, the preimage of multiple images of points and lines can be **defined by the intersection**:

$$\begin{aligned}\text{preimage}(\mathbf{x}_1, \dots, \mathbf{x}_m) &= \text{preimage}(\mathbf{x}_1) \cap \dots \cap \text{preimage}(\mathbf{x}_m), \\ \text{preimage}(\ell_1, \dots, \ell_m) &= \text{preimage}(\ell_1) \cap \dots \cap \text{preimage}(\ell_m).\end{aligned}$$

The above definition allows us to compute preimages for any set of image points or lines. The preimage of multiple image lines, for example, can be either an empty set, a point, a line or a plane, depending on whether or not they come from the same line in space.

Preimage and Coimage of Points and Lines



From Two Views to
Multiple Views

From Preimages to
Rank Constraints

Rank Constraints for
Points

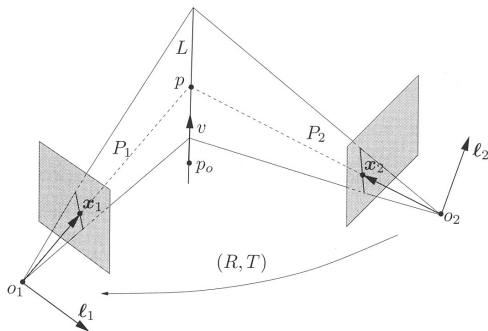
The Multiple-view
Matrix

Relation to Epipolar
Constraints

Multiple-View
Reconstruction
Algorithms

Rank Constraints for
Lines

Summary



Images of a point p on a line L :

- Preimages P_1 and P_2 of the image lines should intersect in the line L .
- Preimages of the two image points x_1 and x_2 should intersect in the point p .
- Normals ℓ_1 and ℓ_2 define the coimages of the line L .

Preimage and Coimage of Points and Lines

For a **moving camera at time t** , let $\mathbf{x}(t)$ denote the image coordinates of a 3D point \mathbf{X} in homogeneous coordinates:

$$\lambda(t)\mathbf{x}(t) = K(t)\Pi_0 g(t)\mathbf{X},$$

where $\lambda(t)$ denotes the depth of the point, $K(t)$ denotes the intrinsic parameters, Π_0 the generic projection, and

$$g(t) = \begin{pmatrix} R(t) & T(t) \\ 0 & 1 \end{pmatrix} \in SE(3),$$

denotes the rigid body motion at time t .

Let us consider a **3D line L** in homogeneous coordinates:

$$L = \{\mathbf{X} \mid \mathbf{X} = \mathbf{X}_0 + \mu\mathbf{V}, \mu \in \mathbb{R}\} \subset \mathbb{R}^4,$$

where $\mathbf{X}_0 = [X_0, Y_0, Z_0, 1]^\top \in \mathbb{R}^4$ are the coordinates of the base point p_0 and $\mathbf{V} = [V_1, V_2, V_3, 0]^\top \in \mathbb{R}^4$ is a nonzero vector indicating the line direction.



From Two Views to
Multiple Views

From Preimages to
Rank Constraints

Rank Constraints for
Points

The Multiple-view
Matrix

Relation to Epipolar
Constraints

Multiple-View
Reconstruction
Algorithms

Rank Constraints for
Lines

Summary

Preimage and Coimage of Points and Lines

The **preimage of L** with respect to the image at time t is a plane P with normal $\ell(t)$, where $P = \text{span}(\hat{\ell})$. The vector $\ell(t)$ is orthogonal to all points $\mathbf{x}(t)$ of the line:

$$\ell(t)^\top \mathbf{x}(t) = \ell(t)^\top K(t) \Pi_0 g(t) \mathbf{X} = 0.$$

Assume we are given a **set of m images at times t_1, \dots, t_m** where

$$\lambda_i = \lambda(t_i), \mathbf{x}_i = \mathbf{x}(t_i), \ell_i = \ell(t_i), \Pi_i = K(t_i) \Pi_0 g(t_i).$$

With this notation, we can relate the **i -th image of a point p** to its world coordinates \mathbf{X} :

$$\lambda_i \mathbf{x}_i = \Pi_i \mathbf{X},$$

and the **i -th coimage of a line L** to its world coordinates $(\mathbf{X}_0, \mathbf{V})$:

$$\ell_i^\top \Pi_i \mathbf{X}_0 = \ell_i^\top \Pi_i \mathbf{V} = 0.$$



From Preimages to Rank Constraints



The above equations contain the **3D parameters of points and lines** as unknowns. As in the two-view case, we wish to **eliminate these unknowns** so as to obtain relationships between the 2D projections and the camera parameters.

In the two-view case an elimination of the 3D coordinates lead to the **epipolar constraint** for the essential matrix E or (in the uncalibrated case) the fundamental matrix F . The 3D coordinates (depth values λ_i associated with each point) could subsequently obtained from another constraint.

There exist **different ways to eliminate the 3D parameters** leading to different kinds of constraints which have been studied in Computer Vision.

A systematic elimination of these constraints will lead to a complete set of conditions.

Consider images of a 3D point \mathbf{X} seen in multiple views:

$$\mathcal{I}\vec{\lambda} \equiv \begin{pmatrix} \mathbf{x}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{x}_2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{x}_m \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix} = \begin{pmatrix} \Pi_1 \\ \Pi_2 \\ \vdots \\ \Pi_m \end{pmatrix} \mathbf{X} \equiv \Pi \mathbf{X},$$

which is of the form

$$\mathcal{I}\vec{\lambda} = \Pi \mathbf{X},$$

where $\vec{\lambda} \in \mathbb{R}^m$ is the **depth scale vector**, and $\Pi \in \mathbb{R}^{3m \times 4}$ the **multiple-view projection matrix** associated with the **image matrix** $\mathcal{I} \in \mathbb{R}^{3m \times m}$.

Note that apart from the 2D coordinates \mathcal{I} , everything else in the above equations is unknown. As in the two-view case, the goal is to decouple the above equation into constraints which allow to separately recover the camera displacements Π_i on one hand and the scene structure λ_i and \mathbf{X} on the other hand.



From Two Views to
Multiple Views

From Preimages to
Rank Constraints

Rank Constraints for
Points

The Multiple-view
Matrix

Relation to Epipolar
Constraints

Multiple-View
Reconstruction
Algorithms

Rank Constraints for
Lines

Summary

Point Features

Every column of \mathcal{I} lies in a four-dimensional space spanned by columns of the matrix Π . In order to have a solution to the above equation, the columns of \mathcal{I} and Π must therefore be linearly dependent. In other words, the matrix

$$N_p \equiv (\Pi, \mathcal{I}) = \begin{pmatrix} \Pi_1 & \mathbf{x}_1 & 0 & \cdots & 0 \\ \Pi_2 & 0 & \mathbf{x}_2 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Pi_m & 0 & 0 & \cdots & \mathbf{x}_m \end{pmatrix} \in \mathbb{R}^{3m \times (m+4)}$$

must have a nontrivial right null space. For $m \geq 2$ (i.e. $3m \geq m+4$), full rank would be $m+4$. Linear dependence of columns therefore implies the **rank constraint**:

$$\boxed{\text{rank}(N_p) \leq m+3.}$$

In fact, the vector $u \equiv (\mathbf{X}^\top, -\vec{\lambda}^\top)^\top \in \mathbb{R}^{m+4}$ is in the right null space, as $N_p u = 0$.





For a more compact formulation of the above rank constraint, we introduce the matrix

$$\mathcal{I}^\perp \equiv \begin{pmatrix} \widehat{\mathbf{x}}_1 & 0 & \cdots & 0 \\ 0 & \widehat{\mathbf{x}}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \widehat{\mathbf{x}}_m \end{pmatrix} \in \mathbb{R}^{3m \times 3m},$$

which has the property of “annihilating” \mathcal{I} :

$$\mathcal{I}^\perp \mathcal{I} = 0,$$

we can premultiply the above equation to obtain

$$\mathcal{I}^\perp \Pi \mathbf{X} = 0.$$



From Two Views to
Multiple Views

From Preimages to
Rank Constraints

Rank Constraints for
Points

The Multiple-view
Matrix

Relation to Epipolar
Constraints

Multiple-View
Reconstruction
Algorithms

Rank Constraints for
Lines

Summary

Thus the vector \mathbf{X} is in the null space of the matrix

$$W_p \equiv \mathcal{I}^\perp \Pi = \begin{pmatrix} \widehat{\mathbf{x}}_1 \Pi_1 \\ \widehat{\mathbf{x}}_2 \Pi_2 \\ \vdots \\ \widehat{\mathbf{x}}_m \Pi_m \end{pmatrix} \in \mathbb{R}^{3m \times 4}.$$

To have a nontrivial solution, we must have

$$\text{rank}(W_p) \leq 3.$$

If all images \mathbf{x}_i are from a single 3D point \mathbf{X} , then the matrix W_p should only have a one-dimensional null space.
Given m images $\mathbf{x}_i \in \mathbb{R}^3$ of a point p with respect to m camera frames Π_i , we must have the rank condition

$$\text{rank}(W_p) = \text{rank}(N_p) - m \leq 3.$$



In the case of a **point** \mathbf{X} , we had the equation

$$W_p \mathbf{X} = 0, \quad \text{with } W_p = \begin{pmatrix} \widehat{\mathbf{x}}_1 \Pi_1 \\ \widehat{\mathbf{x}}_2 \Pi_2 \\ \vdots \\ \widehat{\mathbf{x}}_m \Pi_m \end{pmatrix} \in \mathbb{R}^{3m \times 4}.$$

Since all matrices $\widehat{\mathbf{x}}_i$ have rank 2, the number of independent rows in W_p is at most $2m$. These rows define a set of $2m$ planes. Since $W_p \mathbf{X} = 0$, the point \mathbf{X} is in the intersection of all these planes. In order for the $2m$ planes to have a unique intersection, we need to have $\text{rank}(W_p) = 3$.

Rank Constraints: Geometric Interpretation



From Two Views to
Multiple Views

From Preimages to
Rank Constraints

Rank Constraints for
Points

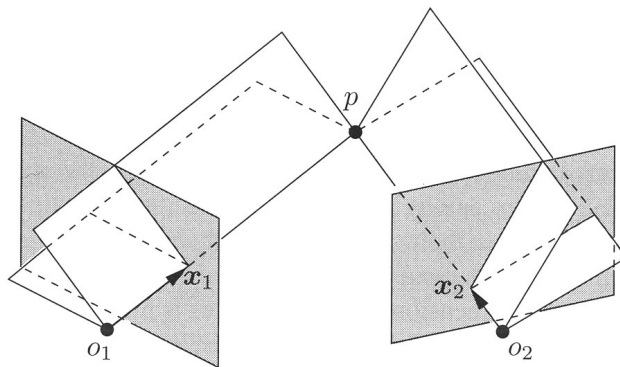
The Multiple-view
Matrix

Relation to Epipolar
Constraints

Multiple-View
Reconstruction
Algorithms

Rank Constraints for
Lines

Summary



Preimage of two image points.

The rows of the matrix W_p correspond to the normal vectors of four planes. The (nontrivial) rank constraint states that these four planes have to intersect in a single point.

The Multiple-view Matrix of a Point

In the following, the rank constraints will be rewritten in a more compact and transparent manner. Let us assume we have m images, the first of which is in world coordinates. Then we have projection matrices of the form

$$\Pi_1 = [I, 0], \Pi_2 = [R_2, T_2], \dots, \Pi_m = [R_m, T_m] \in \mathbb{R}^{3 \times 4},$$

which model the projection of a point \mathbf{X} into the individual images.

In general for uncalibrated cameras (i.e. $K_i \neq I$), R_i will not be an orthogonal rotation matrix but rather an arbitrary invertible matrix.

Again, we define the matrix W_p as follows:

$$W_p \equiv \mathcal{I}^\perp \Pi = \begin{pmatrix} \widehat{\mathbf{x}}_1 \Pi_1 \\ \widehat{\mathbf{x}}_2 \Pi_2 \\ \vdots \\ \widehat{\mathbf{x}}_m \Pi_m \end{pmatrix} \in \mathbb{R}^{3m \times 4}.$$



The Multiple-view Matrix of a Point

The rank of the matrix W_p is not affected if we multiply by a full-rank matrix $D_p \in \mathbb{R}^{4 \times 5}$ as follows:

$$W_p D_p = \begin{pmatrix} \widehat{\mathbf{x}}_1 \Pi_1 \\ \widehat{\mathbf{x}}_2 \Pi_2 \\ \vdots \\ \widehat{\mathbf{x}}_m \Pi_m \end{pmatrix} \begin{pmatrix} \widehat{\mathbf{x}}_1 & \mathbf{x}_1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \widehat{\mathbf{x}}_1 \widehat{\mathbf{x}}_1 & 0 & 0 \\ \widehat{\mathbf{x}}_2 R_2 \widehat{\mathbf{x}}_1 & \widehat{\mathbf{x}}_2 R_2 \mathbf{x}_1 & \widehat{\mathbf{x}}_2 T_2 \\ \widehat{\mathbf{x}}_3 R_3 \widehat{\mathbf{x}}_1 & \widehat{\mathbf{x}}_3 R_3 \mathbf{x}_1 & \widehat{\mathbf{x}}_3 T_3 \\ \vdots & \vdots & \vdots \\ \widehat{\mathbf{x}}_m R_m \widehat{\mathbf{x}}_1 & \widehat{\mathbf{x}}_m R_m \mathbf{x}_1 & \widehat{\mathbf{x}}_m T_m \end{pmatrix}.$$

This means that $\text{rank}(W_p) \leq 3$ if and only if the submatrix

$$M_p \equiv \begin{pmatrix} \widehat{\mathbf{x}}_2 R_2 \mathbf{x}_1 & \widehat{\mathbf{x}}_2 T_2 \\ \widehat{\mathbf{x}}_3 R_3 \mathbf{x}_1 & \widehat{\mathbf{x}}_3 T_3 \\ \vdots & \vdots \\ \widehat{\mathbf{x}}_m R_m \mathbf{x}_1 & \widehat{\mathbf{x}}_m T_m \end{pmatrix} \in \mathbb{R}^{3(m-1) \times 2}$$

has $\text{rank}(M_p) \leq 1$.



[From Two Views to
Multiple Views](#)

[From Preimages to
Rank Constraints](#)

[Rank Constraints for
Points](#)

[The Multiple-view
Matrix](#)

[Relation to Epipolar
Constraints](#)

[Multiple-View
Reconstruction
Algorithms](#)

[Rank Constraints for
Lines](#)

[Summary](#)

The Multiple-view Matrix of a Point

The matrix

$$M_p \equiv \begin{pmatrix} \widehat{\mathbf{x}}_2 R_2 \mathbf{x}_1 & \widehat{\mathbf{x}}_2 T_2 \\ \widehat{\mathbf{x}}_3 R_3 \mathbf{x}_1 & \widehat{\mathbf{x}}_3 T_3 \\ \vdots & \vdots \\ \widehat{\mathbf{x}}_m R_m \mathbf{x}_1 & \widehat{\mathbf{x}}_m T_m \end{pmatrix} \in \mathbb{R}^{3(m-1) \times 2}$$

is called the **multiple-view matrix** associated with a point p . It involves both the image \mathbf{x}_1 in the first view and the coimages $\widehat{\mathbf{x}}_i$ in the remaining views.

In summary:

For multiple images of a point p the matrices N_p , W_p and M_p satisfy:

$$\text{rank}(M_p) = \text{rank}(W_p) - 2 = \text{rank}(N_p) - (m + 2) \leq 1.$$



From Two Views to
Multiple Views

From Preimages to
Rank Constraints

Rank Constraints for
Points

The Multiple-view
Matrix

Relation to Epipolar
Constraints

Multiple-View
Reconstruction
Algorithms

Rank Constraints for
Lines

Summary

Multiview Matrix: Geometric Interpretation

Let us look into the geometric information contained in the multiple-view matrix

$$M_p \equiv \begin{pmatrix} \widehat{\mathbf{x}}_2 R_2 \mathbf{x}_1 & \widehat{\mathbf{x}}_2 T_2 \\ \widehat{\mathbf{x}}_3 R_3 \mathbf{x}_1 & \widehat{\mathbf{x}}_3 T_3 \\ \vdots & \vdots \\ \widehat{\mathbf{x}}_m R_m \mathbf{x}_1 & \widehat{\mathbf{x}}_m T_m \end{pmatrix} \in \mathbb{R}^{3(m-1) \times 2}.$$

The constraint $\text{rank}(M_p) \leq 1$ implies that the two columns are linearly dependent. In fact we have

$\lambda_1 \widehat{\mathbf{x}}_i R_i \mathbf{x}_1 + \widehat{\mathbf{x}}_i T_i = 0, i = 2, \dots, m$ which yields

$$M_p \begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix} = 0.$$

Therefore the coefficient capturing the linear dependence is simply the **distance** λ_1 of the point p from the first camera center. In other words, the multiple-view matrix captures exactly the information about a point p that is missing from a single image, but encoded in multiple images.



[From Two Views to
Multiple Views](#)[From Preimages to
Rank Constraints](#)[Rank Constraints for
Points](#)[The Multiple-view
Matrix](#)[Relation to Epipolar
Constraints](#)[Multiple-View
Reconstruction
Algorithms](#)[Rank Constraints for
Lines](#)[Summary](#)

For the multiple-view matrix

$$M_p \equiv \begin{pmatrix} \widehat{\mathbf{x}}_2 R_2 \mathbf{x}_1 & \widehat{\mathbf{x}}_2 T_2 \\ \widehat{\mathbf{x}}_3 R_3 \mathbf{x}_1 & \widehat{\mathbf{x}}_3 T_3 \\ \vdots & \vdots \\ \widehat{\mathbf{x}}_m R_m \mathbf{x}_1 & \widehat{\mathbf{x}}_m T_m \end{pmatrix} \in \mathbb{R}^{3(m-1) \times 2}.$$

to have $\text{rank}(M_p) = 1$, it is necessary that the pair of vectors $\widehat{\mathbf{x}}_i T_i$ and $\widehat{\mathbf{x}}_i R_i \mathbf{x}_1$ to be linearly dependent for all $i = 2, \dots, m$. This gives the **epipolar constraints**

$$\mathbf{x}_i^\top \widehat{T}_i R_i \mathbf{x}_1 = 0$$

between the first and the i -th image. (Proof see next slide)

Yet, we shall see that the multiview constraint provides more information than the pairwise epipolar constraints.

Relation to Epipolar Constraints

In the previous slide, we claimed that the linear dependence of $\hat{\mathbf{x}}_i T_i$ and $\hat{\mathbf{x}}_i R_i \mathbf{x}_1$ gives rise to the epipolar constraint $\mathbf{x}_i^\top \hat{T}_i R_i \mathbf{x}_1 = 0$. In the following, we shall give a proof of this statement which provides an intuitive geometric understanding of this relationship.

Assume the two vectors $\hat{\mathbf{x}}_i T_i$ and $\hat{\mathbf{x}}_i R_i \mathbf{x}_1$ are dependent, i.e. there is a scalar γ , such that

$$\hat{\mathbf{x}}_i T_i = \gamma \hat{\mathbf{x}}_i R_i \mathbf{x}_1.$$

Since $\hat{\mathbf{x}}_i T_i \equiv \mathbf{x}_i \times T_i$ is proportional to the normal on the plane spanned by \mathbf{x}_i and T_i , and $\hat{\mathbf{x}}_i R_i \mathbf{x}_1$ is proportional to the normal spanned by \mathbf{x}_i and $R_i \mathbf{x}_1$, the linear dependence is equivalent to saying that **the three vectors \mathbf{x}_i , T_i and $R_i \mathbf{x}_1$ are coplanar.**

This again is equivalent to saying that the vector \mathbf{x}_i is orthogonal to the normal on the plane spanned by the vectors T_i and $R_i \mathbf{x}_1$, i.e.

$$\mathbf{x}_i^\top (T_i \times R_i \mathbf{x}_1) = \mathbf{x}_i^\top \hat{T}_i R_i \mathbf{x}_1 = 0.$$



Analysis of the Multiple-view Constraint

For any nonzero vectors $a_i, b_i \in \mathbb{R}^3, i = 1, 2, \dots, n$, the matrix

$$\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_n & b_n \end{pmatrix} \in \mathbb{R}^{3n \times 2}$$

is rank-deficient if and only if $a_i b_j^\top - b_i a_j^\top = 0$ for all $i, j = 1, \dots, n$. We will not prove this statement. Applied to the rank constraint on M_p we get:

$$\hat{\mathbf{x}}_i R_i \mathbf{x}_1 (\hat{\mathbf{x}}_j T_j)^\top - \hat{\mathbf{x}}_i T_i (\hat{\mathbf{x}}_j R_j \mathbf{x}_1)^\top = 0,$$

which gives the **trilinear constraint**

$$\hat{\mathbf{x}}_i (T_i \mathbf{x}_1^\top R_j^\top - R_i \mathbf{x}_1 T_j^\top) \hat{\mathbf{x}}_j = 0.$$

This is a matrix equation giving $3 \times 3 = 9$ scalar trilinear equations, only four of which are linearly independent. One can show that – except for degeneracies – the **bilinear (epipolar) constraints** relating two views are already contained in these **trilinear constraints** obtained for the multiview scenario.



Uniqueness of the Preimage

The bilinear and trilinear constraints help to assure the uniqueness of the preimage of a point observed in three images. Let $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \in \mathbb{R}^3$ be the 2D coordinates in three camera frames with distinct optical centers.

Proposition 1: If the three images satisfy the **pairwise epipolar constraints**

$$\mathbf{x}_i^\top \widehat{T}_{ij} R_{ij} \mathbf{x}_j = 0, \quad i, j = 1, 2, 3,$$

then a unique preimage is determined except if the three lines associated to image points $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ are coplanar – see the examples on slides 24 and 25. Here T_{ij} and R_{ij} refer to the transition between frames i and j .

Proposition 2: If these vectors satisfy all **trilinear constraints**

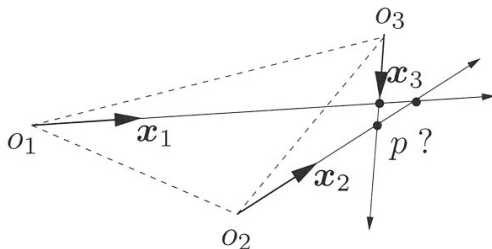
$$\widehat{\mathbf{x}}_j (T_{ji} \mathbf{x}_i^\top R_{ki}^\top - R_{ji} \mathbf{x}_i^\top T_{ki}^\top) \widehat{\mathbf{x}}_k = 0, \quad i, j, k = 1, 2, 3,$$

then a unique preimage is determined unless the three lines associated to image points $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ are colinear.

We will not prove these statements.



Degeneracies for the Bilinear Constraints



In the above example, the point p lies in the plane spanned by the three optical centers which is also called the **trifocal plane**. In this case, all pairs of lines do intersect, yet it does not imply a unique 3D point p (a unique preimage). In practice this degenerate case arises rather seldom.

[From Two Views to
Multiple Views](#)

[From Preimages to
Rank Constraints](#)

[Rank Constraints for
Points](#)

[The Multiple-view
Matrix](#)

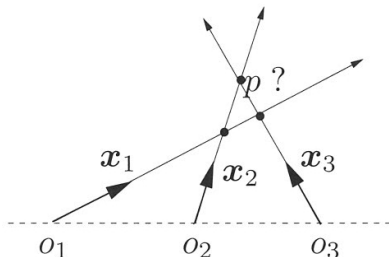
[Relation to Epipolar
Constraints](#)

[Multiple-View
Reconstruction
Algorithms](#)

[Rank Constraints for
Lines](#)

[Summary](#)

Degeneracies for the Bilinear Constraints



In the above example, the optical centers lie on a straight line (**rectilinear motion**). Again, all pairs of lines may intersect without there being a unique preimage p .

This case is frequent in applications when the camera moves in a straight line (e.g. a car on a highway). Then the epipolar constraints will not allow a unique reconstruction.

Fortunately, the **trilinear constraint assures a unique preimage** (unless p is also on the same line with the optical centers).



[From Two Views to
Multiple Views](#)

[From Preimages to
Rank Constraints](#)

[Rank Constraints for
Points](#)

[The Multiple-view
Matrix](#)

[Relation to Epipolar
Constraints](#)

[Multiple-View
Reconstruction
Algorithms](#)

[Rank Constraints for
Lines](#)

[Summary](#)

Uniqueness of the Preimage

Using the multiple-view matrix we obtain a more general and simpler characterization regarding the uniqueness of the preimage:

Given m vectors representing the m images of a point in m views, they correspond to the same point in the $3D$ space if the rank of the M_p matrix relative to any of the camera frames is one. If the rank is zero, the point is determined up to the line on which all the camera centers must lie.

In summary we get:

$\text{rank}(M_p) = 2$	\Rightarrow	no point correspondence & empty preimage
$\text{rank}(M_p) = 1$	\Rightarrow	point correspondence & unique preimage
$\text{rank}(M_p) = 0$	\Rightarrow	point correspondence & preimage not unique

With these constraints we could decide which features to match for establishing point correspondence over multiple frames.





The rank condition on the multiple-view matrix captures all the constraints among multiple images of a point. In principle, one could perform reconstruction by maximizing some global objective function subject to the rank condition. This would lead to a **nonlinear optimization problem analogous to the bundle adjustment in the two-view case.**

Alternatively, one can aim for a similar **separation of structure and motion** as done for the two-view case in the eight-point algorithm. Such an algorithm shall be detailed in the following. One should point out that this approach does not necessarily lead to a practical algorithm as the spectral approaches do not imply optimality in the context of noise and uncertainty.

From Two Views to
Multiple Views

From Preimages to
Rank Constraints

Rank Constraints for
Points

The Multiple-view
Matrix

Relation to Epipolar
Constraints

Multiple-View
Reconstruction
Algorithms

Rank Constraints for
Lines

Summary

Multiple-view Factorization of Point Features

Suppose we have m images $\mathbf{x}_1^j, \dots, \mathbf{x}_m^j$ of n points p^j and we want to estimate the unknown projection matrix Π .

The condition $\text{rank}(M_p) \leq 1$ states that the two columns of M_p are linearly dependent. For the j -th point p^j this implies

$$\begin{pmatrix} \widehat{\mathbf{x}}_2^j R_2 \mathbf{x}_1^j \\ \widehat{\mathbf{x}}_3^j R_3 \mathbf{x}_1^j \\ \vdots \\ \widehat{\mathbf{x}}_m^j R_m \mathbf{x}_1^j \end{pmatrix} + \alpha^j \begin{pmatrix} \widehat{\mathbf{x}}_2^j T_2 \\ \widehat{\mathbf{x}}_3^j T_3 \\ \vdots \\ \widehat{\mathbf{x}}_m^j T_m \end{pmatrix} = 0 \quad \in \mathbb{R}^{3(m-1) \times 1},$$

for some parameters $\alpha^j \in \mathbb{R}, j = 1, \dots, n$. Each row in the above equation can be obtained from $\lambda_i^j \mathbf{x}_i^j = \lambda_1^j R_i \mathbf{x}_1^j + T_i$, multiplying by $\widehat{\mathbf{x}}_1^j$:

$$\widehat{\mathbf{x}}_i^j R_i \mathbf{x}_1^j + \widehat{\mathbf{x}}_1^j T_i / \lambda_1^j = 0.$$

Therefore, $\alpha^j = 1/\lambda_1^j$ is nothing but the **inverse of the depth of point p^j** with respect to the first frame.



From Two Views to
Multiple Views

From Preimages to
Rank Constraints

Rank Constraints for
Points

The Multiple-view
Matrix

Relation to Epipolar
Constraints

Multiple-View
Reconstruction
Algorithms

Rank Constraints for
Lines

Summary

Motion Estimation from Known Structure



From Two Views to
Multiple Views

From Preimages to
Rank Constraints

Rank Constraints for
Points

The Multiple-view
Matrix

Relation to Epipolar
Constraints

Multiple-View
Reconstruction
Algorithms

Rank Constraints for
Lines

Summary

Assume we have the depth of the points and thus their inverse α^j (i.e. known structure). Then the above equation is linear in the camera motion parameters R_i and T_i . Using the stack notation $R_i^s = [r_{11}, r_{21}, r_{31}, r_{12}, r_{22}, r_{32}, r_{13}, r_{23}, r_{33}]^T \in \mathbb{R}^9$ and $T_i \in \mathbb{R}^3$, we have the linear equation system

$$P_i \begin{pmatrix} R_i^s \\ T_i \end{pmatrix} = \begin{pmatrix} \mathbf{x}_1^{1\top} \otimes \widehat{\mathbf{x}}_i^1 & \alpha^1 \widehat{\mathbf{x}}_i^1 \\ \mathbf{x}_1^{2\top} \otimes \widehat{\mathbf{x}}_i^2 & \alpha^2 \widehat{\mathbf{x}}_i^2 \\ \vdots & \vdots \\ \mathbf{x}_1^{n\top} \otimes \widehat{\mathbf{x}}_i^n & \alpha^n \widehat{\mathbf{x}}_i^n \end{pmatrix} \begin{pmatrix} R_i^s \\ T_i \end{pmatrix} = 0 \quad \in \mathbb{R}^{3n}.$$

One can show that the matrix $P_i \in \mathbb{R}^{3n \times 12}$ is of **rank 11 if more than $n = 6$ points in general position are given**. In that case the null space of P_i is onedimensional and the projection matrix $\Pi_i = (R_i, T_i)$ is given up to a scale factor. In practice one would use more than 6 points, obtain a full-rank matrix and compute the solution by a singular value decomposition (SVD).

Structure Estimation from Known Motion

In turn, if the camera motion $\Pi_i = (R_i, T_i), i = 1, \dots, m$ is known, we can estimate the structure (depth parameters $\alpha^j, j = 1, \dots, m$). The least squares solution for the above equation is given by:

$$\alpha^j = - \frac{\sum_{i=2}^m (\hat{\mathbf{x}}_i^j T_i)^\top \hat{\mathbf{x}}_i^j R_i \mathbf{x}_1^j}{\sum_{i=2}^m \|\hat{\mathbf{x}}_i^j T_i\|^2}, \quad j = 1, \dots, n.$$

In this way one can **iteratively estimate structure and motion**, estimating one while keeping the other fixed.

For **initialization** one could apply the **eight-point algorithm** to the first two images to obtain an estimate of the structure parameters α^j .

While the equation for Π_i makes use of the two frames 1 and i only, the structure parameter estimation takes into account all frames. This can be done either in **batch mode** or **recursively**.

As for the two-view case, such spectral approaches **do not guarantee optimality in the presence of noise and uncertainty**.





[From Two Views to
Multiple Views](#)

[From Preimages to
Rank Constraints](#)

[Rank Constraints for
Points](#)

[The Multiple-view
Matrix](#)

[Relation to Epipolar
Constraints](#)

[Multiple-View
Reconstruction
Algorithms](#)

[Rank Constraints for
Lines](#)

[Summary](#)

We can derive a similar rank constraint for lines. As we saw above, for the coimages $\ell_i, i = 1, \dots, m$ of a line L spanned by a base \mathbf{X}_0 and a direction \mathbf{V} we have:

$$\ell_i^\top \Pi_i \mathbf{X}_0 = \ell_i^\top \Pi_i \mathbf{V} = 0.$$

Therefore the matrix

$$\mathbf{W}_l \equiv \begin{pmatrix} \ell_1^\top \Pi_1 \\ \ell_2^\top \Pi_2 \\ \vdots \\ \ell_m^\top \Pi_m \end{pmatrix} \in \mathbb{R}^{m \times 4}$$

must satisfy the rank constraint

$\text{rank}(\mathbf{W}_l) \leq 2,$

since the null space of \mathbf{W}_l contains the two vectors \mathbf{X}_0 and \mathbf{V} .



In the case of a line L in two views, we have the equation

$$\text{rank}(W_l) \leq 2, \quad \text{with } W_l = \begin{pmatrix} \ell_1^\top \Pi_1 \\ \ell_2^\top \Pi_2 \end{pmatrix} \in \mathbb{R}^{2 \times 4}.$$

Clearly, we already have $\text{rank}(W_l) \leq 2$, so there is no intrinsic constraint on two images of a line: The preimage of two image lines always contains a line. This is no longer true for three or more images of a line, then the above constraint really becomes meaningful.

From Two Views to
Multiple Views

From Preimages to
Rank Constraints

Rank Constraints for
Points

The Multiple-view
Matrix

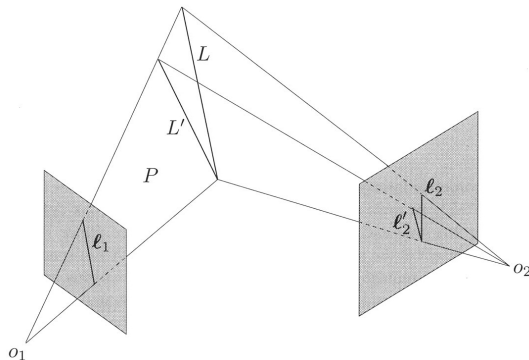
Relation to Epipolar
Constraints

Multiple-View
Reconstruction
Algorithms

Rank Constraints for
Lines

Summary

Rank Constraints: Geometric Interpretation



Preimage of two image lines.

For the case of a line observed from two images, the rank constraint is always fulfilled. Geometrically this states that the two preimages of each line always intersect in some 3D line – in this example either L or L' .



From Two Views to
Multiple Views

From Preimages to
Rank Constraints

Rank Constraints for
Points

The Multiple-view
Matrix

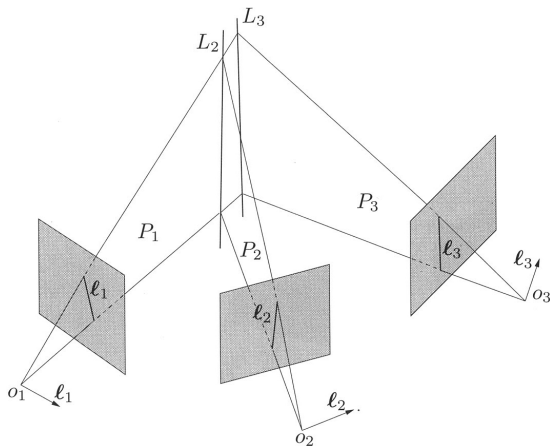
Relation to Epipolar
Constraints

Multiple-View
Reconstruction
Algorithms

Rank Constraints for
Lines

Summary

Uniqueness of the Preimage



In the case of three or more images, the rank constraint is not always fulfilled. In this example, it is not fulfilled. There is no preimage: The lines L_2 and L_3 don't coincide.

From Two Views to
Multiple Views

From Preimages to
Rank Constraints

Rank Constraints for
Points

The Multiple-view
Matrix

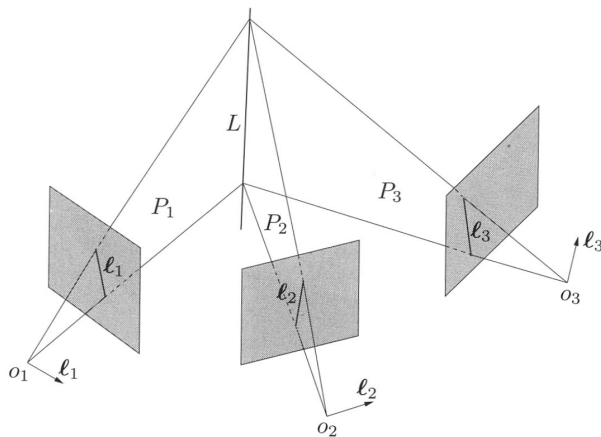
Relation to Epipolar
Constraints

Multiple-View
Reconstruction
Algorithms

Rank Constraints for
Lines

Summary

Uniqueness of the Preimage



Uniqueness of the preimage: The lines L_2 and L_3 coincide.

From Two Views to
Multiple Views

From Preimages to
Rank Constraints

Rank Constraints for
Points

The Multiple-view
Matrix

Relation to Epipolar
Constraints

Multiple-View
Reconstruction
Algorithms

Rank Constraints for
Lines

Summary



It is shown in [Ma, Soatto, Kosecka, Sastry, 2002](#) that the rank constraints on W_l can be further compressed into a multiview matrix $M_l \in \mathbb{R}^{(m-1) \times 4}$.

Overall we have the following cases:

$\text{rank}(M_l) = 2 \Rightarrow$ no line correspondence, no preimage

$\text{rank}(M_l) = 1 \Rightarrow$ line correspondence & unique preimage

$\text{rank}(M_l) = 0 \Rightarrow$ line correspondence & preimage not unique

[From Two Views to
Multiple Views](#)

[From Preimages to
Rank Constraints](#)

[Rank Constraints for
Points](#)

[The Multiple-view
Matrix](#)

[Relation to Epipolar
Constraints](#)

[Multiple-View
Reconstruction
Algorithms](#)

[Rank Constraints for
Lines](#)

[Summary](#)

Summary

One can generalize the two-view scenario to that of **simultaneously considering $m \geq 2$ images** of a scene. The intrinsic constraints among multiple images of a point or a line can be expressed in terms of **rank conditions** on the matrix N , W or M .

The relationship among these rank conditions is as follows:

	(Pre)image	coimage	Jointly
Point	$\text{rank}(N_p) \leq m + 3$	$\text{rank}(W_p) \leq 3$	$\text{rank}(M_p) \leq 1$
Line	$\text{rank}(N_l) \leq 2m + 2$	$\text{rank}(W_l) \leq 2$	$\text{rank}(M_l) \leq 1$

These rank conditions capture the relationships among corresponding geometric primitives in multiple images. They impose the **existence of unique preimages** (up to degenerate cases). Moreover, they give rise to natural **factorization-based algorithms for multiview recovery of 3D structure and motion** (i.e. generalizations of the eight-point algorithm).

