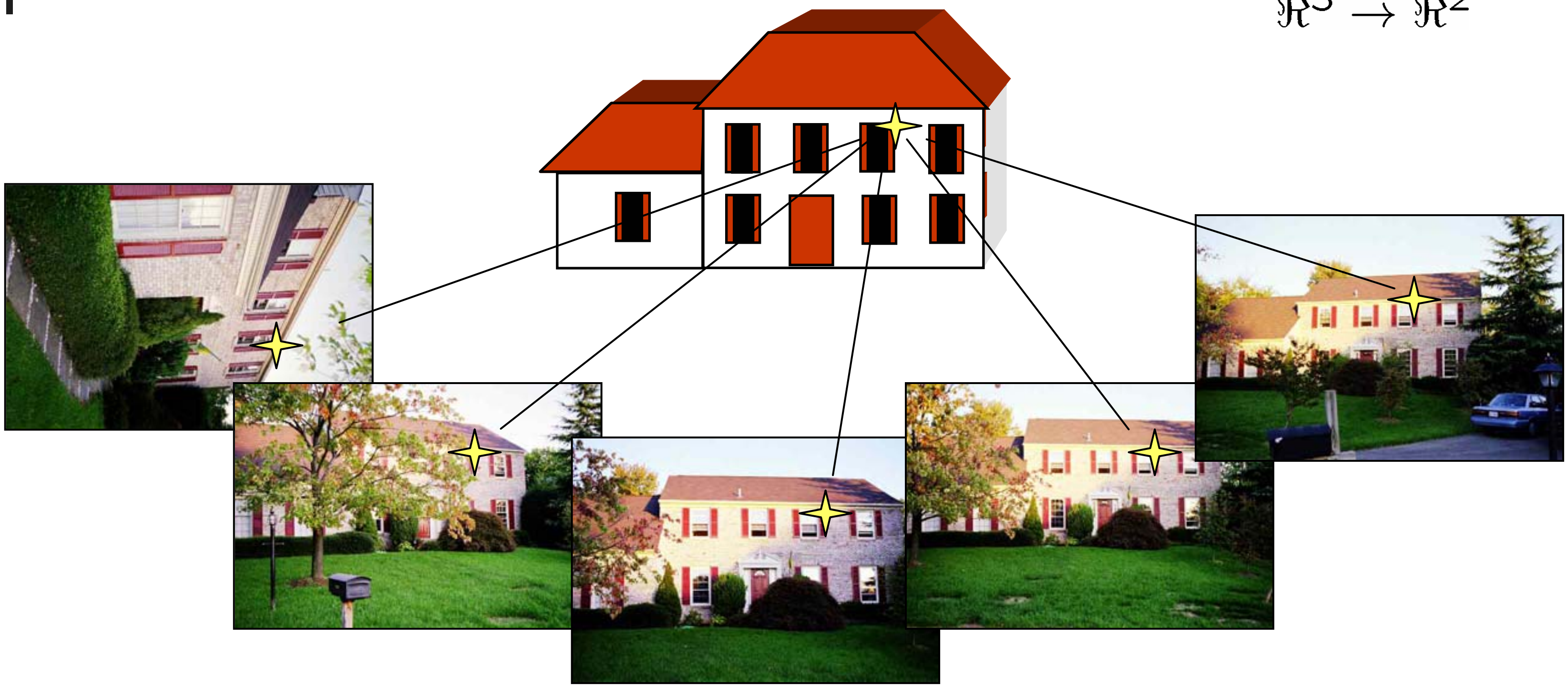
Multiple-view Reconstruction from Points and Lines

Perception & Decision Laboratory
Decision & Control Group, CSL
Image Formation & Processing Group, Beckman
Electrical & Computer Engineering Dept., UIUC
http://decision.csl.uiuc.edu/~yima

Problem formulation

Input: Corresponding images (of "features") in multiple images.
Output: Camera motion, camera calibration, object structure.

$\Re^3 \rightarrow \Re^2$



Projection model – point features

Homogeneous coordinates of a 3-D point p

$$\mathbf{X} = [X, Y, Z, W]^T \in \Re^4, \quad (W = 1)$$

Homogeneous coordinates of its 2-D image

$$\mathbf{x} = [x, y, z]^T \in \Re^3, \quad (z = 1)$$

Projection of a 3-D point to an image plane

$$\boxed{\lambda(t)\mathbf{x}(t) = \Pi(t)\mathbf{X}}$$

$$\lambda(t) \in \Re, \ \Pi(t) = [R(t), T(t)] \in \Re^{3\times 4}$$

$$R(t) \rightarrow K(t)R(t), \ T(t) \rightarrow K(t)T(t)$$

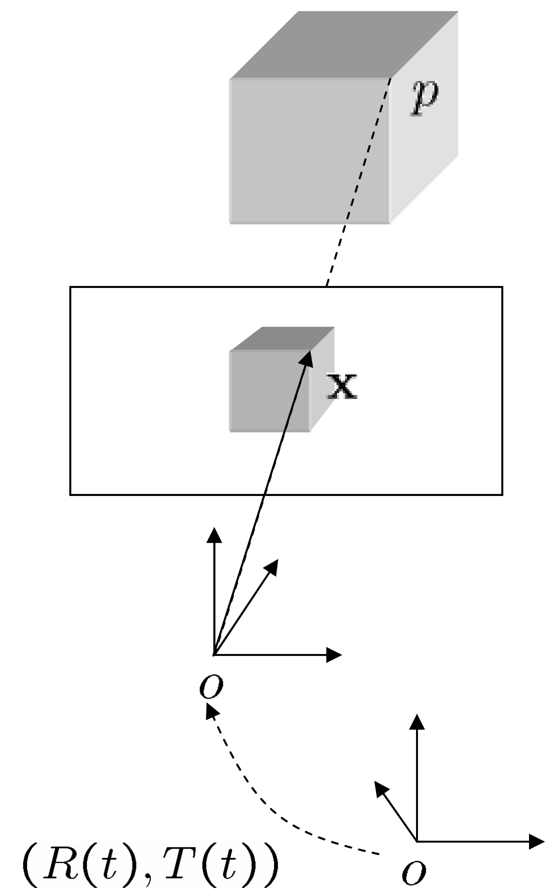


Image of a line feature

Homogeneous representation of a 3-D line L

$$\mathbf{X} = \mathbf{X}_o + \mu \mathbf{V}, \quad \mathbf{X}_o, \mathbf{V} \in \Re^4, \mu \in \Re$$

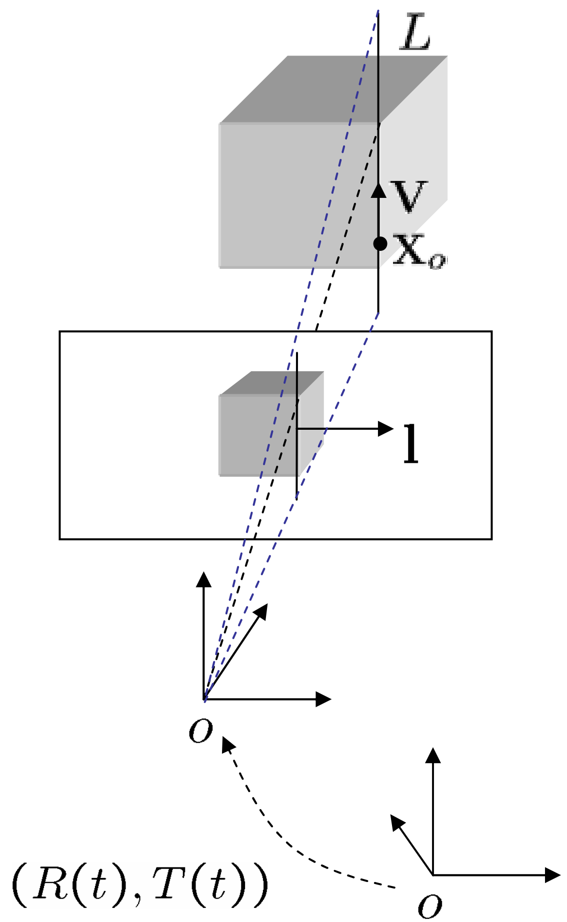
Homogeneous representation of its 2-D co-image

$$\mathbf{l} = [a, b, c]^T \in \Re^3$$

Projection of a 3-D line to an image plane

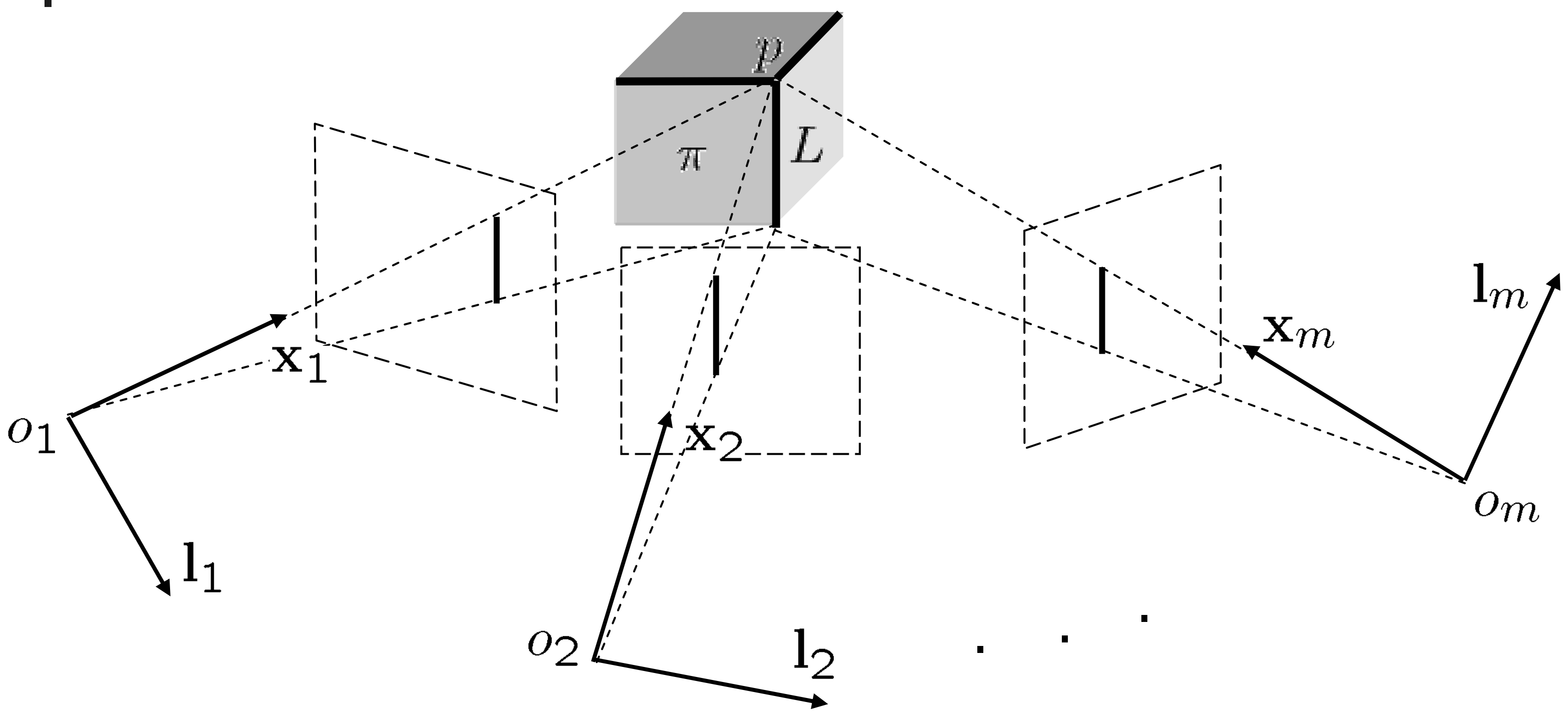
$$\boxed{\mathbf{l}(t)^T \mathbf{x}(t) = \mathbf{l}(t)^T \Pi(t) \mathbf{X} = 0}$$

$$\Pi(t) = [R(t), T(t)] \in \Re^{3\times4}$$



Incidence relations among features

Multiview constraints are nothing but incidence relations at play!



Questions

- What are the basic relations among multiple images of a point/line?
- How many images do I need?
- When are those relations insufficient ?
- How can I use all the images to reconstruct camera pose and scene structure?
- How can I do the reconstruction if some features are occluded?

Traditional multifocal constraints

For m images of the same 3-D point p: $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$

$$\begin{bmatrix} \mathbf{x}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{x}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{x}_m \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{bmatrix} = \begin{bmatrix} \Pi_1 \\ \Pi_2 \\ \vdots \\ \Pi_m \end{bmatrix} \mathbf{X}$$

$$N_p \doteq \begin{bmatrix} \Pi_1 & \mathbf{x}_1 & 0 & \cdots & 0 \\ \Pi_2 & 0 & \mathbf{x}_2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \Pi_m & 0 & \cdots & 0 & \mathbf{x}_m \end{bmatrix} \in \Re^{3m \times (m+4)}.$$

$rank(N_p) \leq m + 3$ (leading to the conventional approach)

$\boxed{\det\left(N_{p(m+4)\times(m+4)}\right) = 0}$ $\Longrightarrow$ Multilinear constraints among 2, 3, 4-wise views

Rank conditions for point feature

WLOG, choose camera frame 1 as the reference

$$N_p = \begin{bmatrix} I & 0 & \mathbf{x}_1 & 0 & \cdots & 0 \\ R_2 & T_2 & 0 & \mathbf{x}_2 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ R_m & T_m & 0 & \cdots & 0 & \mathbf{x}_m \end{bmatrix} \in \Re^{3m\times(m+4)}, \quad M_p \doteq \begin{bmatrix} \widehat{\mathbf{x}_2} R_2 \mathbf{x}_1 & \widehat{\mathbf{x}_2} T_2 \\ \widehat{\mathbf{x}_3} R_3 \mathbf{x}_1 & \widehat{\mathbf{x}_3} T_3 \\ \vdots & \vdots \\ \widehat{\mathbf{x}_m} R_m \mathbf{x}_1 & \widehat{\mathbf{x}_m} T_m \end{bmatrix} \in \Re^{3(m-1)\times 2}$$

Multiple-View Matrix

Lemma [Rank Condition for Point Features]

$$rank(N_p) = (m+2) + rank(M_p)$$

Let $M_p = [M_1, M_2]$, then M_1 and M_2 are linearly dependent.

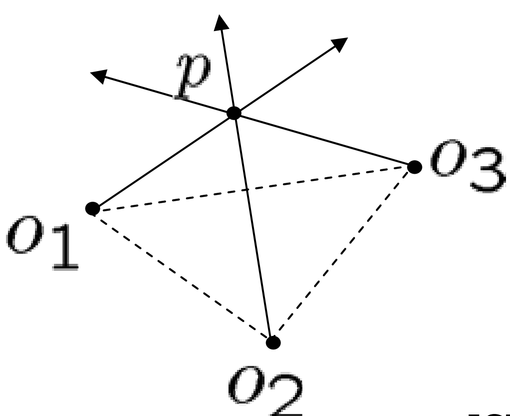
Multiple View Matrix for Point Features

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

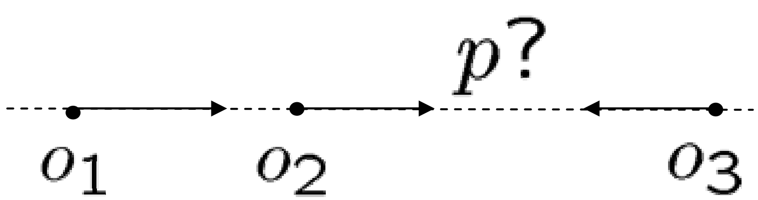
$$0 \leq \text{rank}(M_p) \leq 1$$

M_p encodes exactly the 3-D information missing in one image.

$\text{rank}(M_p) = 1$



$\text{rank}(M_p) = 0$



Rank conditions vs. multifocal constraints

$$rank \begin{bmatrix} \widehat{\mathbf{x}_2} R_2 \mathbf{x}_1 & \widehat{\mathbf{x}_2} T_2 \\ \widehat{\mathbf{x}_3} R_3 \mathbf{x}_1 & \widehat{\mathbf{x}_3} T_3 \\ \vdots & \vdots \\ \widehat{\mathbf{x}_m} R_m \mathbf{x}_1 & \widehat{\mathbf{x}_m} T_m \end{bmatrix} = 1 \;\Rightarrow\; \mathbf{x}_i^T \widehat{T_i} R_i \mathbf{x}_1 = 0$$

$$rank \begin{bmatrix} \widehat{\mathbf{x}_2} R_2 \mathbf{x}_1 & \widehat{\mathbf{x}_2} T_2 \\ \widehat{\mathbf{x}_3} R_3 \mathbf{x}_1 & \widehat{\mathbf{x}_3} T_3 \\ \vdots & \vdots \\ \widehat{\mathbf{x}_m} R_m \mathbf{x}_1 & \widehat{\mathbf{x}_m} T_m \end{bmatrix} = 1 \;\Rightarrow\; \widehat{\mathbf{x}_i} (R_i \mathbf{x}_1 T_j^T - T_i \mathbf{x}_1^T R_j^T) \widehat{\mathbf{x}_j} = 0$$

- These constraints are only necessary but NOT sufficient!
- However, there is NO further relationship among quadruple wise views. Quadrilinear constraints hence are redundant!

Multiple-view structure and motion recovery

Given m images of n points: $(\mathbf{x}_1^j, \ldots, \mathbf{x}_i^j, \ldots, \mathbf{x}_m^j), \; j = 1, \ldots, n$

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

$$P_i \begin{bmatrix} \vec{T}_i \\ \vec{R}_i \end{bmatrix} = \begin{bmatrix} \alpha^1 \widehat{\mathbf{x}}_i^1 & \widehat{\mathbf{x}}_i^1 * {\mathbf{x}_1^1}^T \\ \alpha^2 \widehat{\mathbf{x}}_i^2 & \widehat{\mathbf{x}}_i^2 * {\mathbf{x}_1^2}^T \\ \vdots & \vdots \\ \alpha^n \widehat{\mathbf{x}}_i^n & \widehat{\mathbf{x}}_i^n * {\mathbf{x}_1^n}^T \end{bmatrix} \begin{bmatrix} \vec{T}_i \\ \vec{R}_i \end{bmatrix} = 0 \qquad \in \Re^{3n\times 1}$$

If $n \geq 6$, in general $rank(P_i) = 11$

Reconstruction Algorithm for Point Features

Given m images of n (>7) points

For the *jth* point

$$\begin{bmatrix} \widehat{\boldsymbol{x}_2^j} R_2 \boldsymbol{x}_1^j & \widehat{\boldsymbol{x}_2^j} T_2 \\ \widehat{\boldsymbol{x}_3^j} R_3 \boldsymbol{x}_1^j & \widehat{\boldsymbol{x}_3^j} T_3 \\ \vdots & \vdots \\ \widehat{\boldsymbol{x}_m^j} R_m \boldsymbol{x}_1^j & \widehat{\boldsymbol{x}_m^j} T_m \end{bmatrix} \begin{bmatrix} \lambda^j \\ 1 \end{bmatrix} = 0 \Rightarrow {\lambda^j}^s$$

SVD

Iteration

SVD

For the *ith* image

$$\begin{bmatrix} \lambda^1 {\boldsymbol{x}_1^1}^T \otimes \widehat{\boldsymbol{x}_i^1} & \widehat{\boldsymbol{x}_i^1} \\ \lambda^2 \boldsymbol{x}_1^2 \otimes \widehat{\boldsymbol{x}_i^2} & \widehat{\boldsymbol{x}_i^2} \\ \vdots & \vdots \\ \lambda^n \boldsymbol{x}_1^n \otimes \boldsymbol{x}_i^n & \widehat{\boldsymbol{x}_i^n} T_m \end{bmatrix} \begin{bmatrix} R_i^s \\ T_i^s \end{bmatrix} = 0 \Rightarrow (R_i^s, T_i^s)$$

SVD based 4-step algorithm for SFM

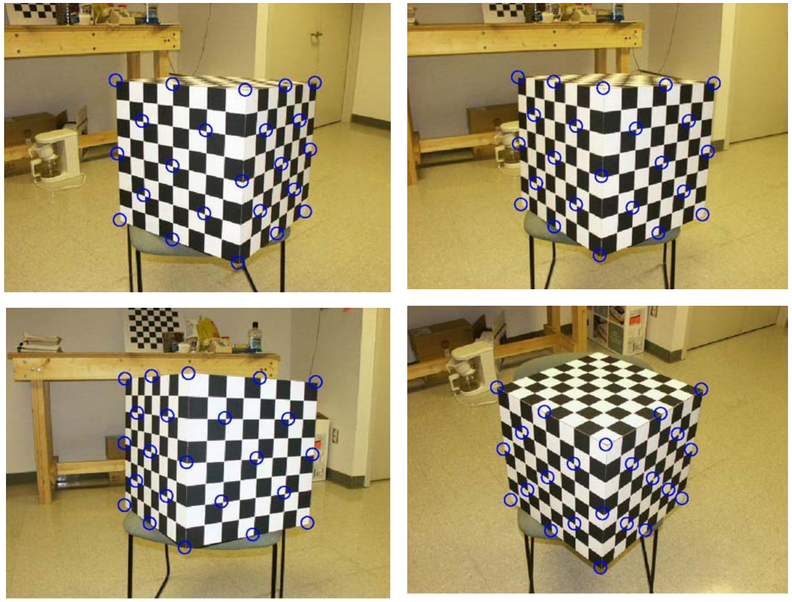
1. Initialization: (a) Set k=0; (b) Compute (R_2, T_2) using the eight point algorithm; (c) Compute $\alpha^j = \alpha^j_k$. Normalize s.t. $\alpha^1_k = 1$.

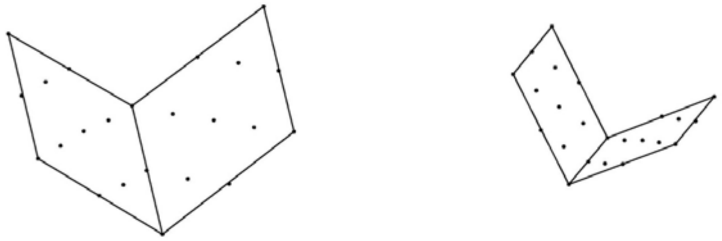
2. Compute $(\tilde{R}_i, \tilde{T}_i)$ as the null space of P_i, $i = 2, \ldots, m$.

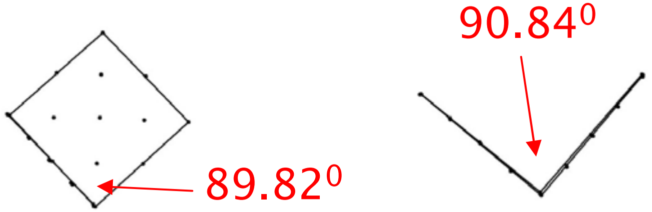
3. Compute the new $\alpha^j = \alpha^j_{k+1}$ as the null space of $M^j, j = 1, \ldots, n$. Normalize s.t. $\alpha^1_{k+1} = 1$.

4. If $||\alpha_k - \alpha_{k+1}|| > \epsilon$, then $k = k+1$ and goto 2. Else stop.

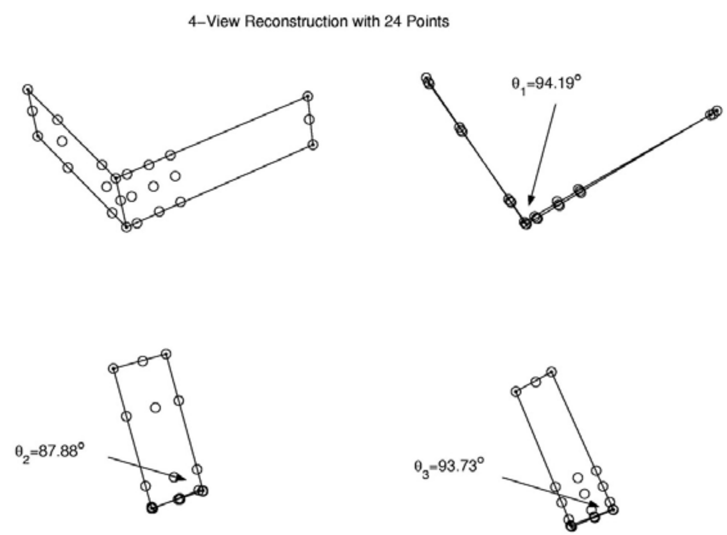
Reconstruction Algorithm for Point Features



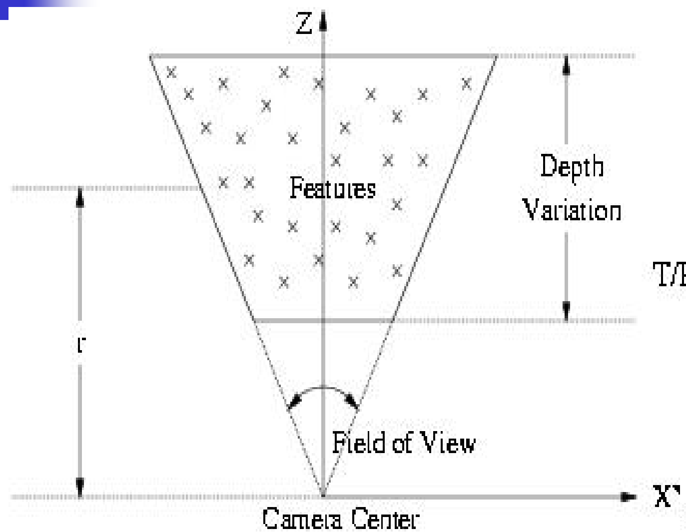




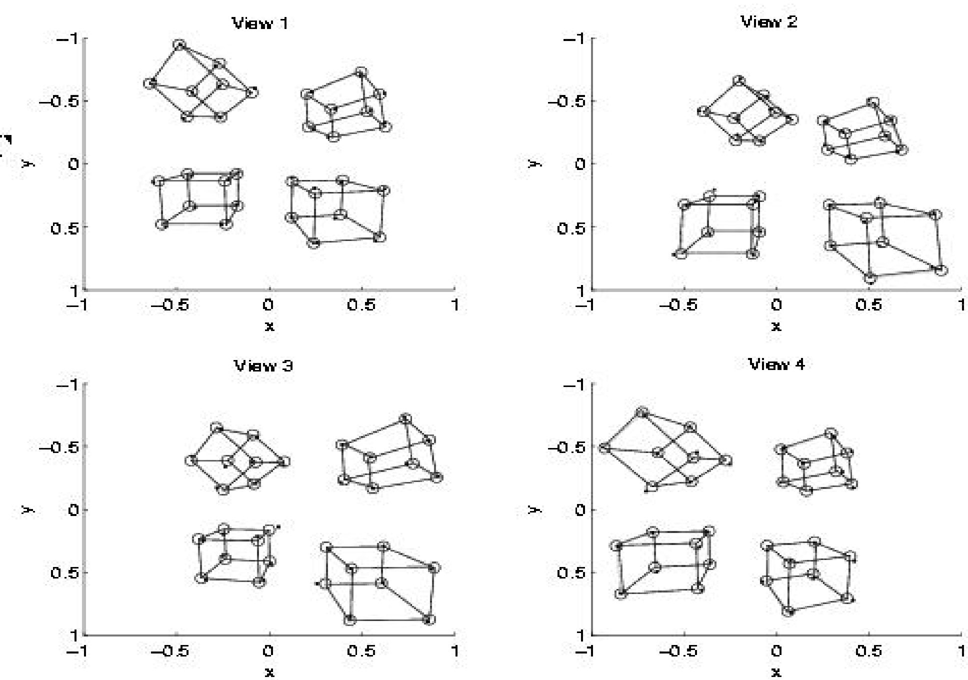
Reconstruction Algorithm for Point Features



Example: simulations



T/R ratio = $\|T\|/r\theta$



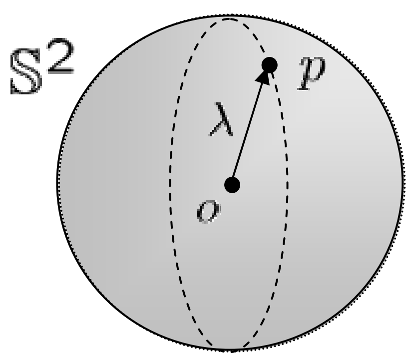
Multiple View Matrix for Line Features

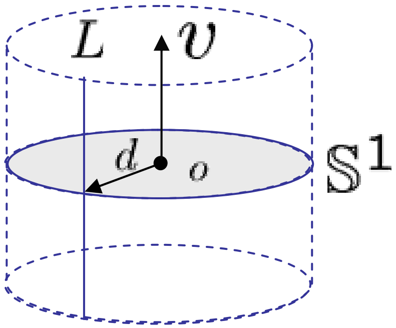
- Point Features
- Line Features

$$M_p = \begin{bmatrix} \widehat{\boldsymbol{x}_2} R_2 \boldsymbol{x}_1 & \widehat{\boldsymbol{x}_2} T_2 \\ \widehat{\boldsymbol{x}_3} R_3 \boldsymbol{x}_1 & \widehat{\boldsymbol{x}_3} T_3 \\ \vdots & \vdots \\ \widehat{\boldsymbol{x}_m} R_m \boldsymbol{x}_1 & \widehat{\boldsymbol{x}_m} T_m \end{bmatrix} \in \mathbb{R}^{3(m-1)\times 2}, \quad M_l = \begin{bmatrix} \boldsymbol{l}_2^T R_2 \widehat{\boldsymbol{l}_1} & \boldsymbol{l}_2^T T_2 \\ \boldsymbol{l}_3^T R_3 \widehat{\boldsymbol{l}_1} & \boldsymbol{l}_3^T T_3 \\ \vdots & \vdots \\ \boldsymbol{l}_m^T R_m \widehat{\boldsymbol{l}_1} & \boldsymbol{l}_m^T T_m \end{bmatrix} \in \mathbb{R}^{(m-1)\times 4}$$

$$0 \le rank(M_p) \le 1 \qquad 0 \le rank(M_l) \le 1$$

M encodes exactly the 3-D information missing in one image.

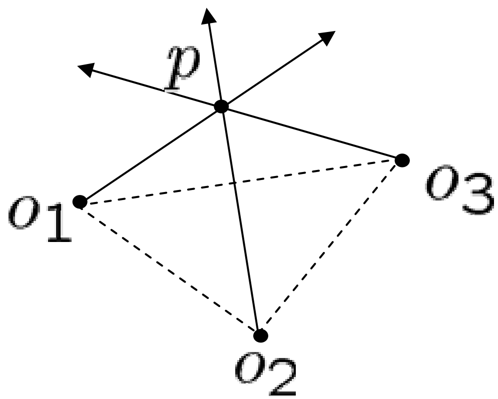




Multiple View Matrix for Line Features

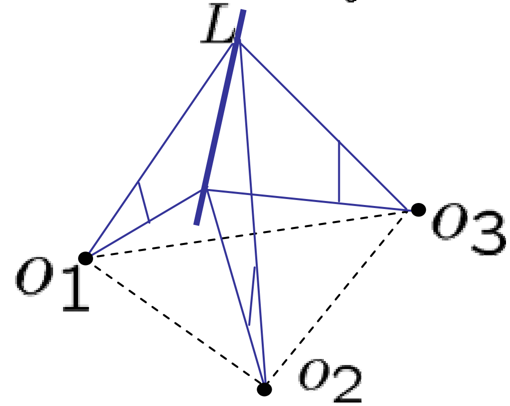
$$\text{rank}(M_p) = 1$$

$$\widehat{\boldsymbol{x}_i}(T_i \boldsymbol{x}_1^T R_j^T - R_i \boldsymbol{x}_1 T_j^T)\widehat{\boldsymbol{x}_j} = 0$$

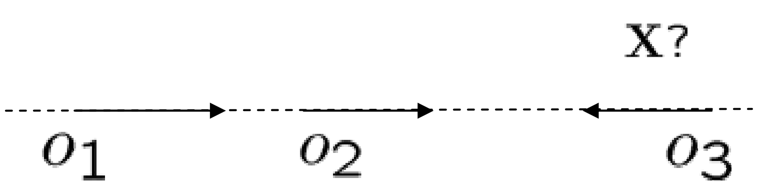


$$\text{rank}(M_l) = 1$$

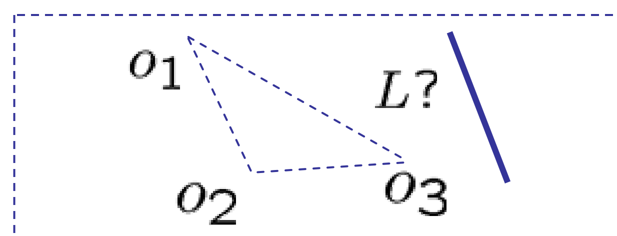
$$\boldsymbol{l}_j^T T_j \boldsymbol{l}_i^T R_i \widehat{\boldsymbol{l}_1} - \boldsymbol{l}_i^T T_i \boldsymbol{l}_j^T R_j \widehat{\boldsymbol{l}_1} = 0$$



$$\text{rank}(M_p) = 0$$



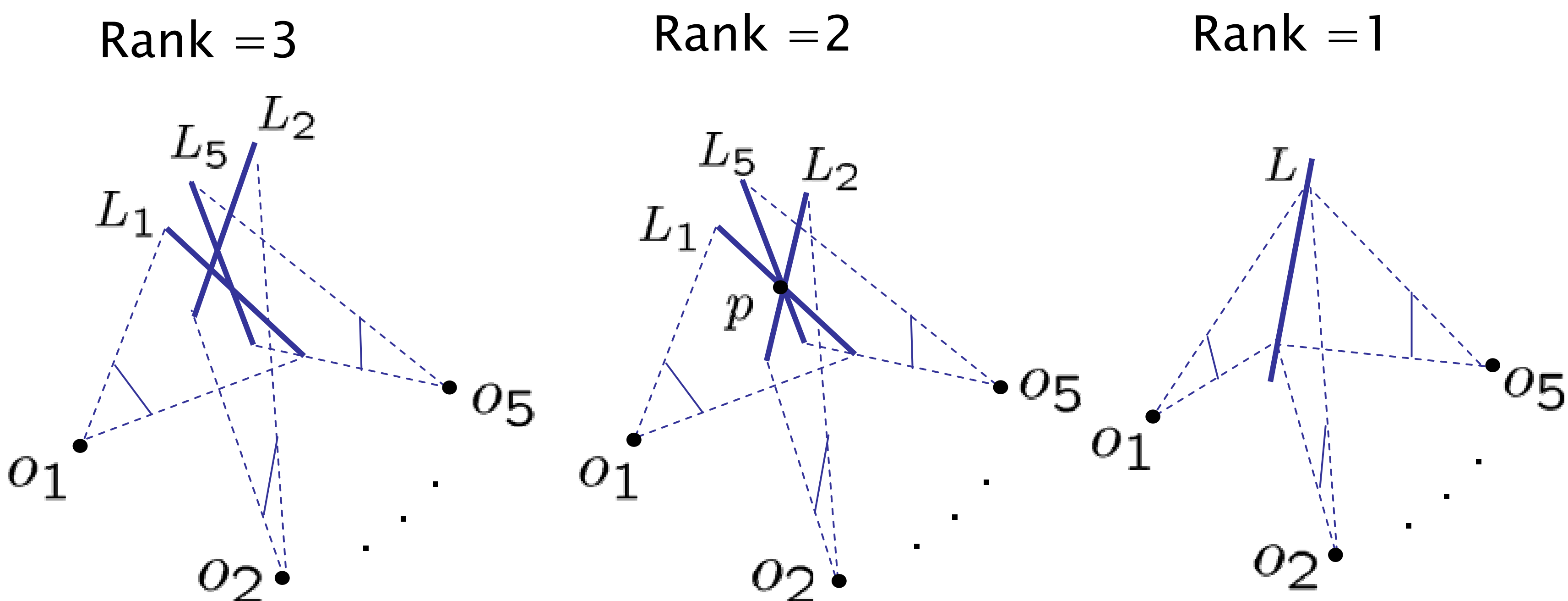
$$\text{rank}(M_l) = 0$$



Multiple View Matrix for Line Features

$$M_l = \begin{bmatrix} \boldsymbol{l}_2^T R_2 \widehat{\boldsymbol{l}_1} & \boldsymbol{l}_2^T T_2 \\ \boldsymbol{l}_3^T R_3 \widehat{\boldsymbol{l}_1} & \boldsymbol{l}_3^T T_3 \\ \boldsymbol{l}_4^T R_4 \widehat{\boldsymbol{l}_1} & \boldsymbol{l}_4^T T_4 \\ \boldsymbol{l}_5^T R_5 \widehat{\boldsymbol{l}_1} & \boldsymbol{l}_5^T T_5 \end{bmatrix} \in \mathbb{R}^{4\times 4}, \quad rank(M_l) = 3, 2, 1.$$

$\boldsymbol{l}_1, \boldsymbol{l}_2, \boldsymbol{l}_3, \boldsymbol{l}_4, \boldsymbol{l}_5$ each is an image of a (different) line in 3-D:



Universal Rank Constraint

- What if we have both point and line features?
  - Traditionally points and lines are treated separately
  - Therefore, joint incidence relations not exploited

- Can we express joint incidence relations for
  - Points passing through lines?
  - Families of intersecting lines?

Universal rank condition

Theorem [The Universal Rank Condition] for images of a point on a line:

$$M \doteq \begin{bmatrix} D_2^{\perp} R_2 D_1 & D_2^{\perp} T_2 \\ D_3^{\perp} R_3 D_1 & D_3^{\perp} T_3 \\ \vdots & \vdots \\ D_m^{\perp} R_m D_1 & D_m^{\perp} T_m \end{bmatrix}, \quad \text{where} \quad \begin{cases} D_i \doteq & \mathbf{x}_i \text{ or } \widehat{\mathbf{l}_i}, \\ D_i^{\perp} \doteq & \widehat{\mathbf{x}_i} \text{ or } \mathbf{l}_i^T. \end{cases}$$

1. If $D_1 = \widehat{\mathbf{l}_1}$ and $D_i^{\perp} = \widehat{\mathbf{x}_i}$ for some $i \geq 2$, then:

$$\boxed{1 \leq rank(M) \leq 2.}$$

-Multi-nonlinear constraints among 3, 4-wise images.

2. Otherwise:

$$\boxed{0 \leq rank(M) \leq 1.}$$

-Multi-linear constraints among 2, 3-wise images.

Instances with mixed features

Examples:

Case 1: a line reference

$$M_l \doteq \begin{bmatrix} \widehat{\mathbf{x}_2} R_2 \widehat{\mathbf{l}_1} & \widehat{\mathbf{x}_2} T_2 \\ \mathbf{l}_3^T R_3 \widehat{\mathbf{l}_1} & \mathbf{l}_3^T T_3 \\ \widehat{\mathbf{x}_4} R_4 \widehat{\mathbf{l}_1} & \widehat{\mathbf{x}_4} T_4 \\ \mathbf{l}_5^T R_5 \widehat{\mathbf{l}_1} & \mathbf{l}_5^T T_5 \end{bmatrix}$$

$$1 \leq rank(M_l) \leq 2$$

Case 2: a point reference

$$M_p \doteq \begin{bmatrix} \mathbf{l}_2^T R_2 \mathbf{x}_1 & \mathbf{l}_2^T T_2 \\ \widehat{\mathbf{x}_3} R_3 \mathbf{x}_1 & \widehat{\mathbf{x}_3} T_3 \\ \mathbf{l}_4^T R_4 \mathbf{x}_1 & \mathbf{l}_4^T T_4 \\ \mathbf{l}_5^T R_5 \mathbf{x}_1 & \mathbf{l}_5^T T_5 \end{bmatrix}.$$

$$0 \leq rank(M_p) \leq 1$$

- All previously known constraints are the theorem’s instances.
- Degenerate configurations if and only if a drop of rank.

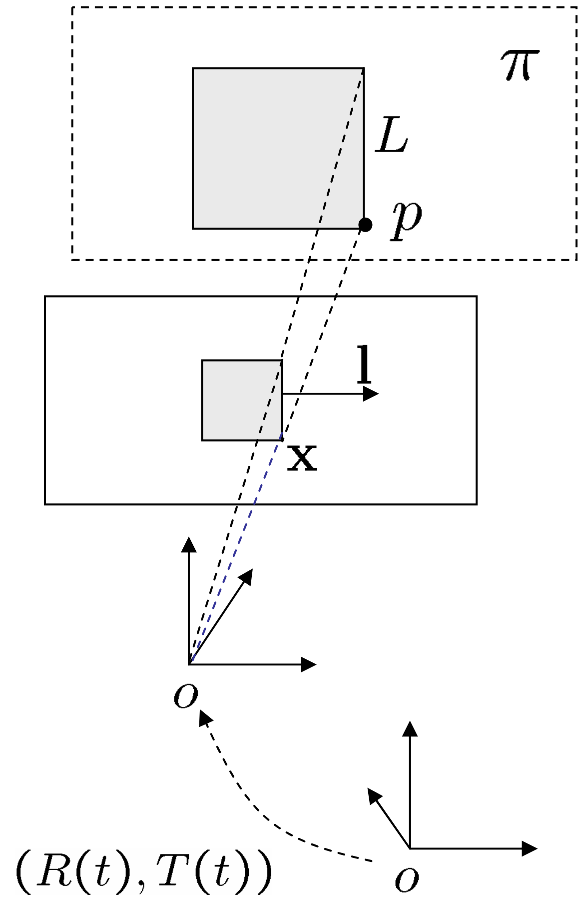
Generalization – restriction to a plane

Homogeneous representation of a 3-D plane π

$$aX + bY + cZ + d = 0.$$

$$\pi \mathbf{X} = 0, \quad \pi = [\pi^1, \pi^2] : \pi^1 \in \Re^3, \pi^2 \in \Re$$

$$M \doteq \begin{bmatrix} D_2^{\perp} R_2 D_1 & D_2^{\perp} T_2 \\ D_3^{\perp} R_3 D_1 & D_3^{\perp} T_3 \\ \vdots & \vdots \\ D_m^{\perp} R_m D_1 & D_m^{\perp} T_m \\ \pi^1 D_1 & \pi^2 \end{bmatrix}$$



Corollary [Coplanar Features]

Rank conditions on the new extended M remain exactly the same!

Generalization – restriction to a plane

Given that a point and line features lie on a plane π in 3-D space:

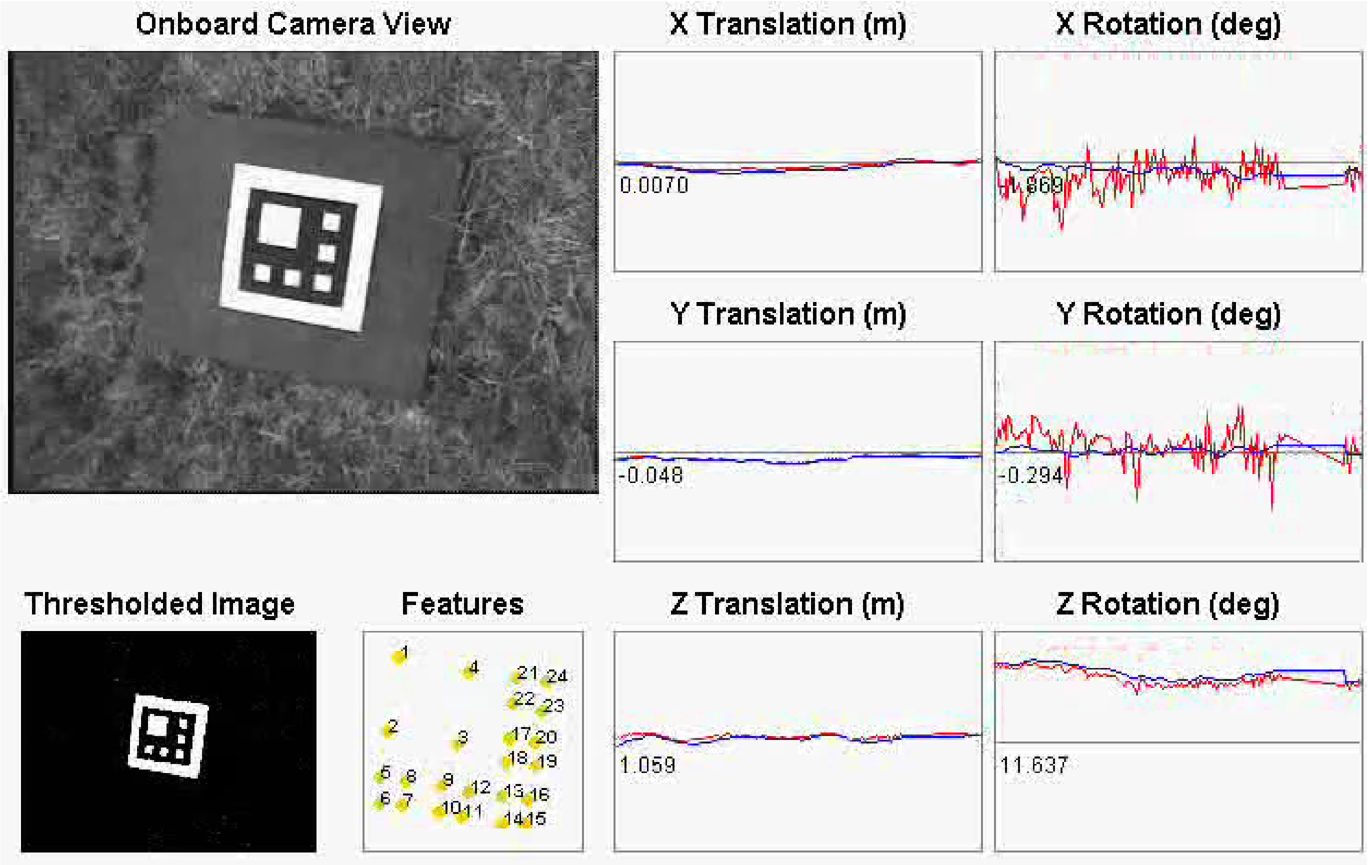
$$M_p = \begin{bmatrix} \widehat{\mathbf{x}_2} R_2 \mathbf{x}_1 & \widehat{\mathbf{x}_2} T_2 \\ \widehat{\mathbf{x}_3} R_3 \mathbf{x}_1 & \widehat{\mathbf{x}_3} T_3 \\ \vdots & \vdots \\ \widehat{\mathbf{x}_m} R_m \mathbf{x}_1 & \widehat{\mathbf{x}_m} T_m \\ \pi^1 \mathbf{x}_1 & \pi^2 \end{bmatrix} \in \Re^{(3m-2)\times 2}, \quad M_l = \begin{bmatrix} \mathbf{l}_2^T R_2 \widehat{\mathbf{l}_1} & \mathbf{l}_2^T T_2 \\ \mathbf{l}_3^T R_3 \widehat{\mathbf{l}_1} & \mathbf{l}_3^T T_3 \\ \vdots & \vdots \\ \mathbf{l}_m^T R_m \widehat{\mathbf{l}_1} & \mathbf{l}_m^T T_m \\ \pi^1 \widehat{\mathbf{l}_1} & \pi^2 \end{bmatrix} \in \Re^{m\times 4}$$

$$0 \le rank(M_p) \le 1 \qquad\qquad 0 \le rank(M_l) \le 1$$

In addition to previous constraints, it simultaneously gives homography:

$$\boxed{\widehat{\mathbf{x}_i}(R_i \pi^2 - T_i \pi^1)\mathbf{x}_1 = 0} \qquad \boxed{\mathbf{l}_i^T (R_i \pi^2 - T_i \pi^1)\widehat{\mathbf{l}_1} = 0}$$

Example Vision-based Landing of a Helicopter



Video courtesy of O. Shakernia and C. Sharp

Summary

- Incidence relations <=> rank conditions
- Rank conditions => multiple-view factorization
- Rank conditions implies all multi-focal constraints
- Rank conditions for points, lines, planes, and (symmetric) structures.
- Rank conditions holds for both calibrated and uncalibrated case – later additional constraints self-calibration