Multi-linear Systems for 3D-from-2D Interpretation

Lecture 4

Dynamic $P^k \rightarrow P^2$ Tensors

Amnon Shashua

Hebrew University of Jerusalem
Israel

Material We Will Cover Today

- The general problem of reconstruction from “line of sights”
- Two Applications for $P^4 \rightarrow P^2$
- An Application for $P^6 \rightarrow P^2$
- The geometry of $P^4 \rightarrow P^2$
- Indexing into “action” (an example).

Wolf, Shashua ICCV01, IJCV'02 (Honorable Mention, Marr Prize)
Levin,Wolf,Shashua CVPR'01

Wolf,Shashua ICCV'01: (translating planes)
Wolf,Shashua CVPR'01: (segmentation tensors)

not discussed here

Dynamic configurations:

- Constant Velocity
- Points in 3D
- Straight-line Motion (of points)
- General Camera Motion
- Trajectories Span a 3D, 2D or 1D space

QuickTime™ and a
Microsoft Video 1 decompressor
are needed to see this picture.

Dynamic configurations:

- Recover Camera motion from image observations
- How many points/views?

QuickTime™ and a
Microsoft Video 1 decompressor
are needed to see this picture.

As Seen in the Image Space

Dynamic configurations:
(multi-body)

- 2 Rigid Configurations
- Bodies moving in relative translation to each other

QuickTime™ and a
Microsoft Video 1 decompressor
are needed to see this picture.

Dynamic configurations:
(multi-body)

- Recover Camera motion from image observations
- How many points/views?
- How many “segmented” points are needed?

QuickTime™ and a
Microsoft Video 1 decompressor
are needed to see this picture.

As Seen in the Image Space

A General Scheme for Constructing Multi-view Tensors of Dynamic Scenes

$P^2 \rightarrow P^2$ projections of (static) planar configuration

$P^3 \rightarrow P^2$ projections of (static) 3D configuration

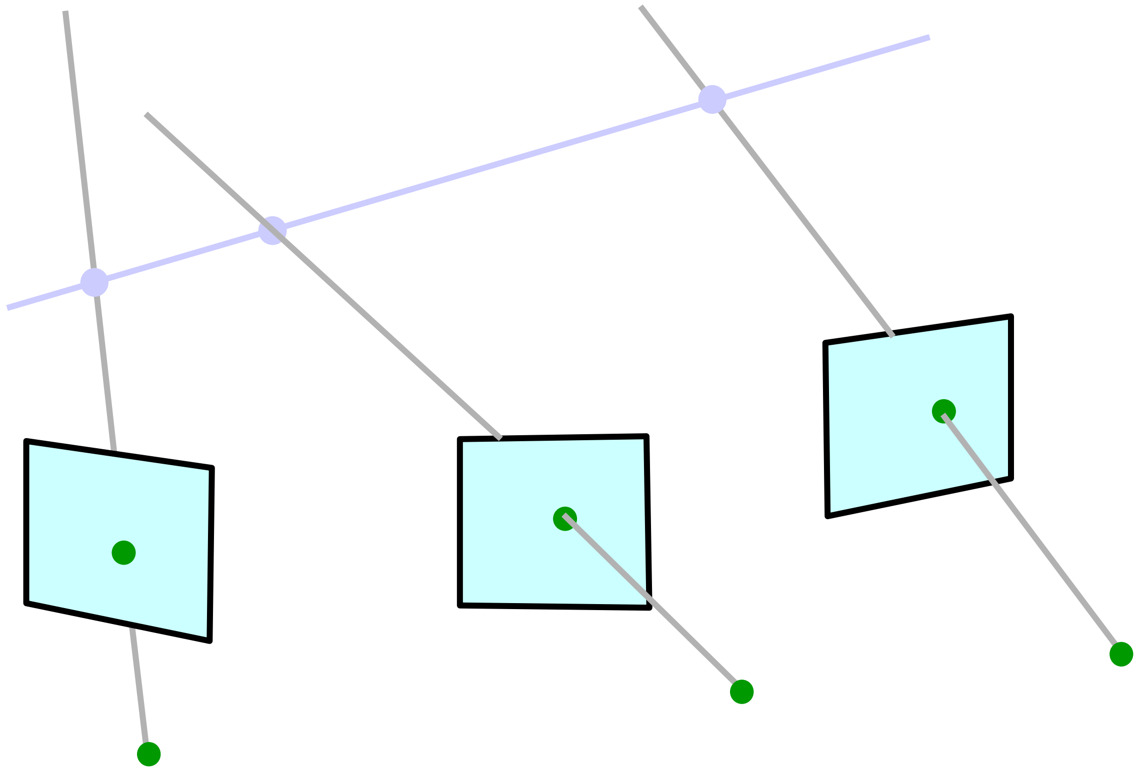
$P^k \rightarrow P^2$ projections of dynamic and multi-body scenes

$k = 3,4,5,6$

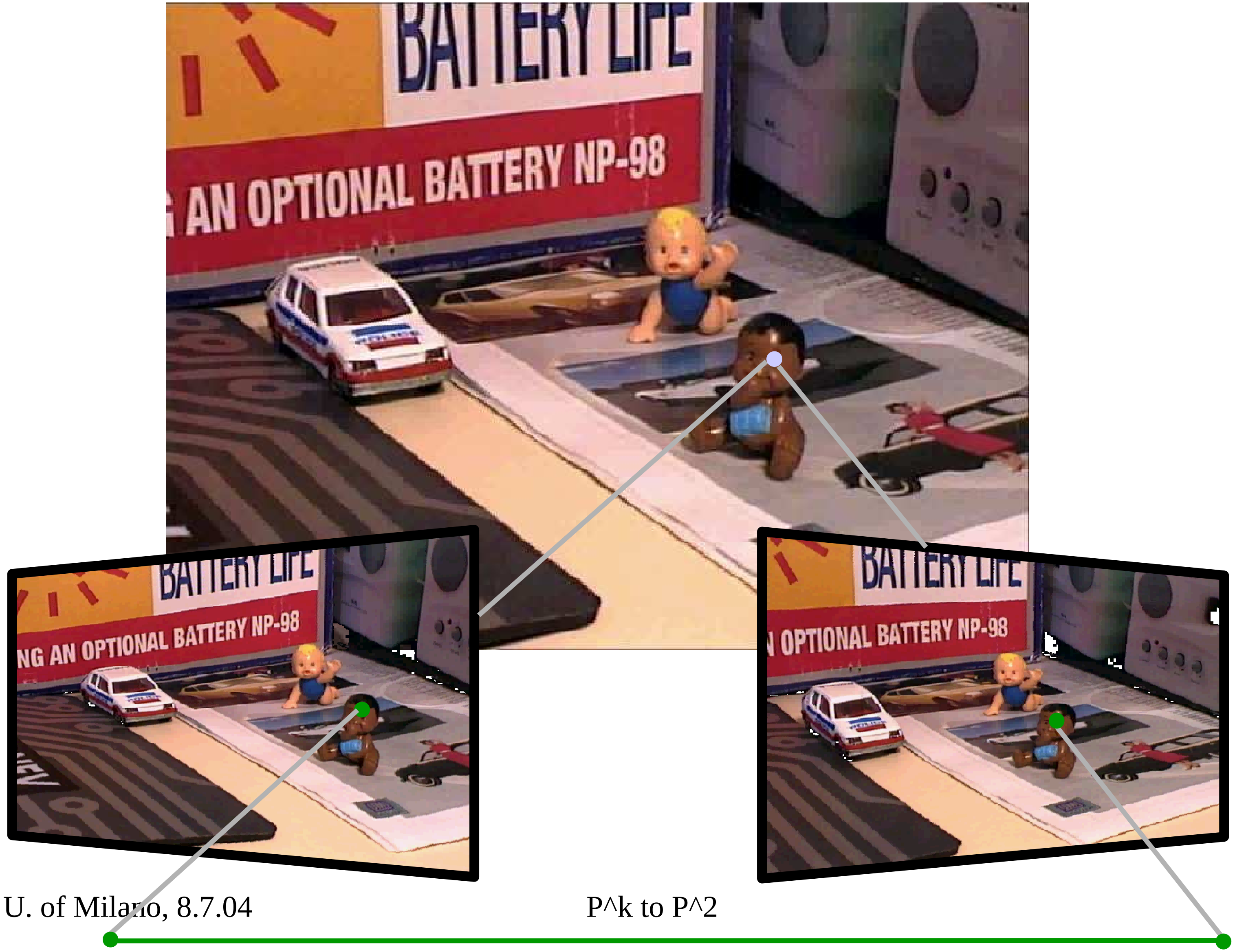
- We describe the multi-view constraints from these projection matrices.

- We show how to extract the structure and motion for each application.

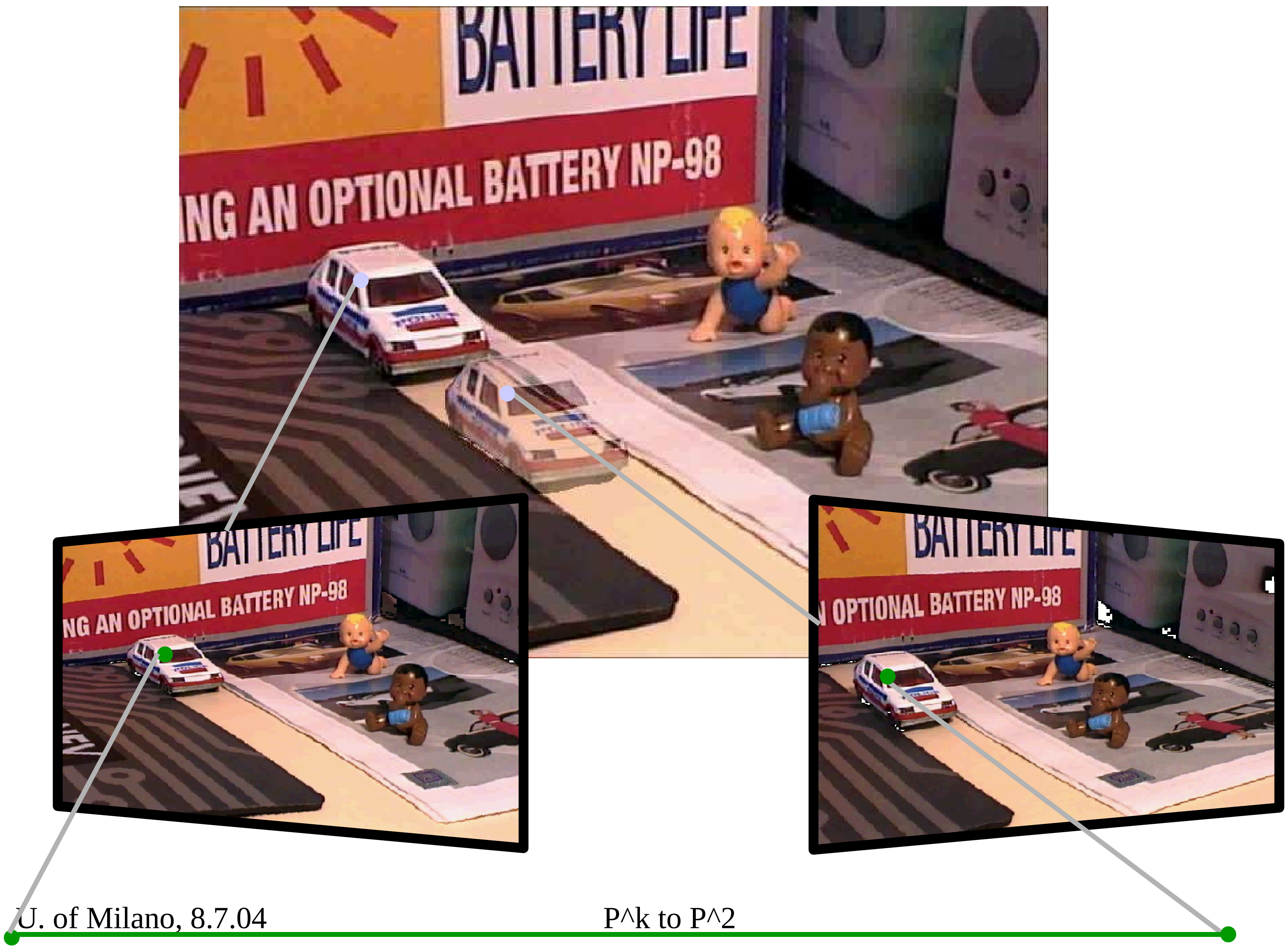
When Camera Trajectory is known



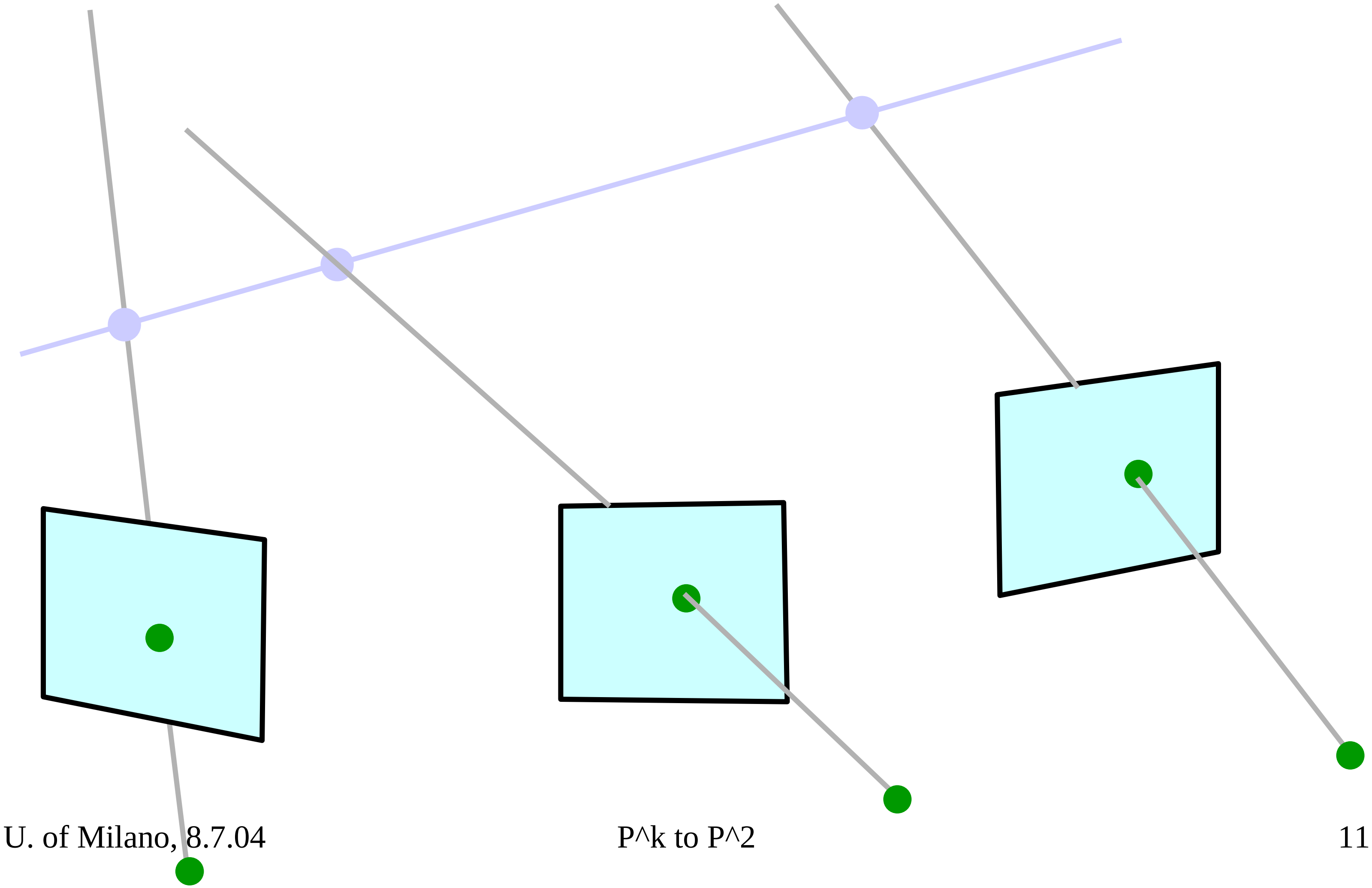
Triangulation



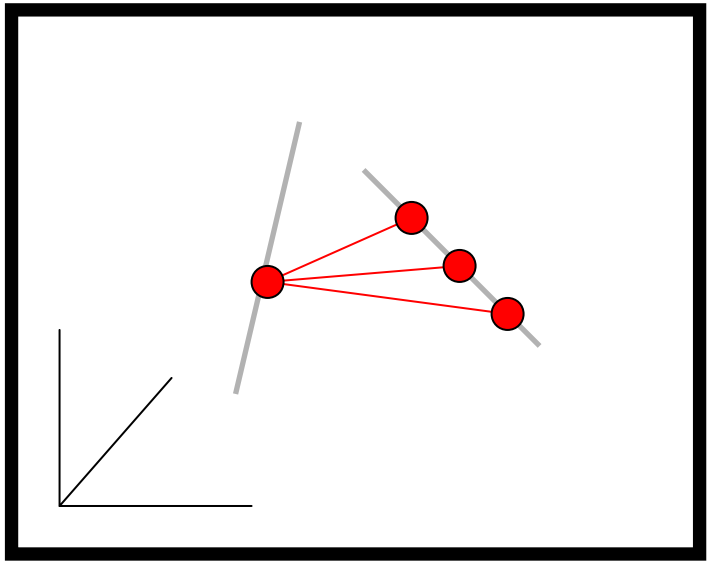
Triangulation?

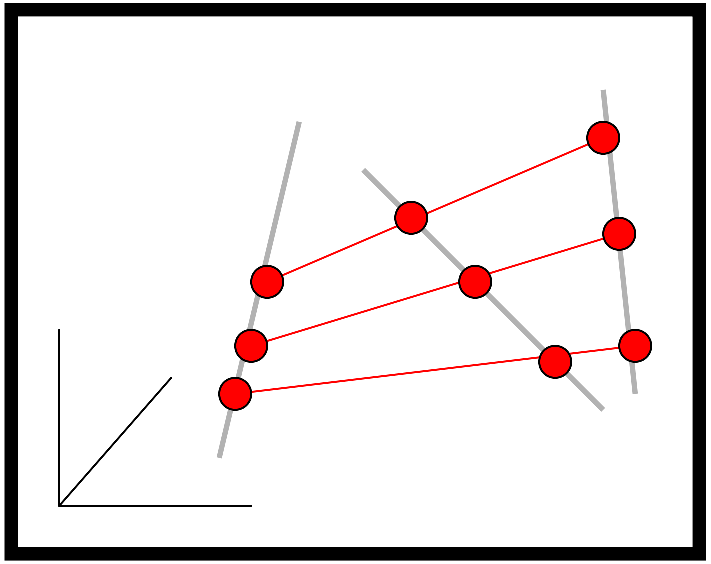


When Camera Trajectory is known



How Many Views for a Unique Line?

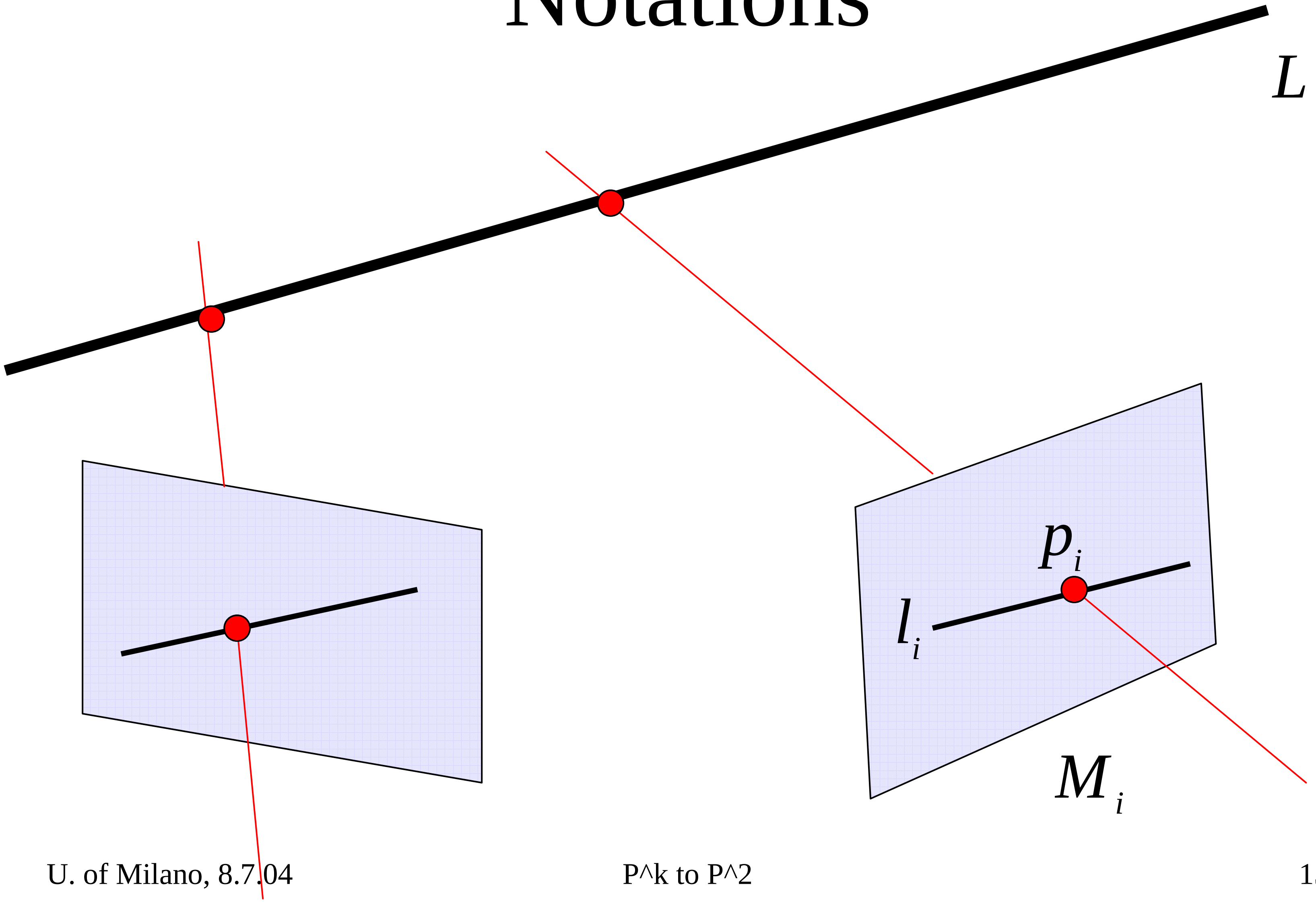




4 views - finite number of solutions (nonlinear)

5 views - unique solution

Notations



Plucker Coordinates

Given two 3D points

$$P_1 = \begin{pmatrix} X_1 & Y_1 & Z_1 & 1 \end{pmatrix}$$
$$P_2 = \begin{pmatrix} X_2 & Y_2 & Z_2 & 1 \end{pmatrix}$$

The 3D line is:

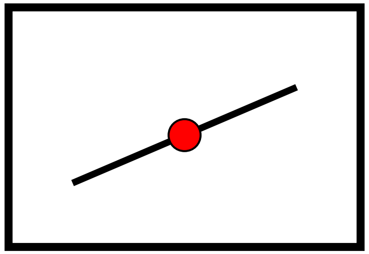
$$L = P_1 \wedge P_2$$
$$= \begin{pmatrix} L^{12} & L^{13} & L^{14} & L^{23} & L^{24} & L^{34} \end{pmatrix}$$

For example:

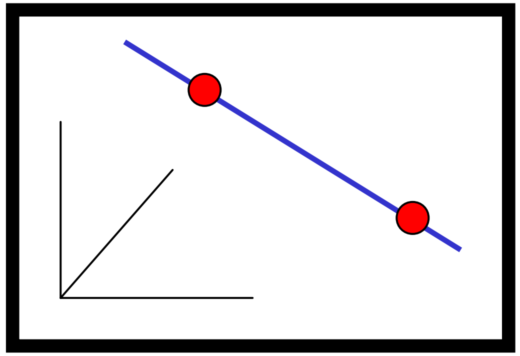
$$L^{12} = \det\begin{pmatrix} 1 & X_1 \\ 1 & X_2 \end{pmatrix}$$

$$\begin{pmatrix} 1 & X_1 & Y_1 & Z_1 \\ 1 & X_2 & Y_2 & Z_2 \end{pmatrix}$$

Linear Solution



$$p^T l = 0$$



$$l = (MP_1)\times(MP_2)$$

$$l = \tilde{M}L$$

$$p^T\left(\tilde{M}L\right) = 0$$

$$M = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix}_{3\times 4}$$

$$\tilde{M} = \begin{bmatrix} m_2 \wedge m_3 \\ m_3 \wedge m_1 \\ m_1 \wedge m_2 \end{bmatrix}_{3\times 6}$$

Dynamic $P^3 \rightarrow P^2$

What do we have so far:

- Reconstruction using 5 views with Known Camera Motion.
- Planar scene, 3 views, unknown camera motion.

Generalization to 3D-from-2D becomes unwieldy:
3x3x3x3x3x3 tensor with 729 observations per view!

➡ Additional Constraints are Required

$P_i + jV_i$ Position of the feature point at frame j

General Idea

Describe the dynamic $P^3 \rightarrow P^2$ applications

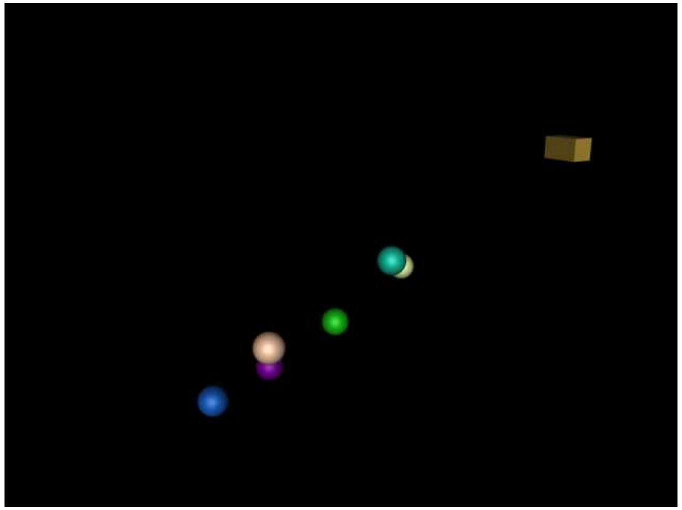
as a static $P^k \rightarrow P^2$ problem

- $P^6 \rightarrow P^2$ Multi-point, 3D, constant-velocity
- $P^5 \rightarrow P^2$ Multi-point, 3D, coplanar constant-velocity
- $P^4 \rightarrow P^2$ Multi-point, 3D, collinear constant-velocity

  Multi-point, 2D, constant-velocity

  Two-body, 3D, Segmentation (pure trans)
- $P^3 \rightarrow P^2$ Two-body, 2D, Segmentation (pure trans)

$P^4 \to P^2$ 3D segmentation tensor



$p_i \cong M_{3\times4} P_i$ if point P_i belongs to object 1

$\hat{p}_i \cong \hat{M}_{3\times4} P_i$ if point P_i belongs to object 2

Because the motion between the two objects is pure translation:

$$M = [A \quad v]$$

$$\hat{M} = [A \quad \hat{v}]$$

3D segmentation tensor

$$p_i \cong [A \quad v \quad \hat{v}]_{3\times 5} \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \\ 0 \end{pmatrix}$$

$\Rightarrow$

$$\hat{p}_i \cong [A \quad v \quad \hat{v}]_{3\times 5} \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 0 \\ 1 \end{pmatrix}$$

3D segmentation tensor

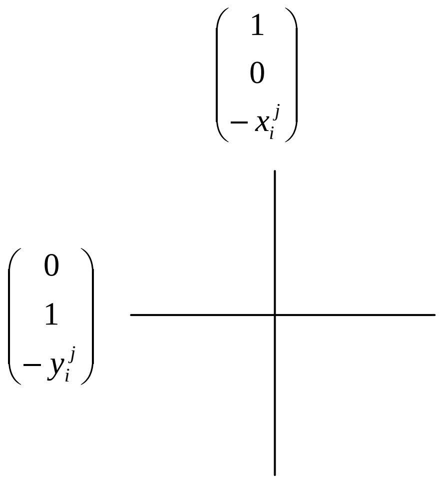
Let $\tilde{M}_j$ be the j'th projection matrix from $P^4 \to P^2$

$$p_i^j = \begin{pmatrix} x_i^j \\ y_i^j \\ 1 \end{pmatrix}$$ be the projection of $\tilde{P}_i \in P^4$ in the j'th view

$$\Rightarrow \quad p_i^j = \tilde{M}_j \tilde{P}_i$$

Note:

$$\begin{pmatrix} 1 \\ 0 \\ -x_i^j \end{pmatrix}^T \begin{pmatrix} x_i^j \\ y_i^j \\ 1 \end{pmatrix} = 0 \qquad \begin{pmatrix} 0 \\ 1 \\ -y_i^j \end{pmatrix}^T \begin{pmatrix} x_i^j \\ y_i^j \\ 1 \end{pmatrix} = 0$$



3D segmentation tensor

3x3x3 tensor (27 coefs)

Each triplet of matching
Points p,p',p''
Provides 2 constraints

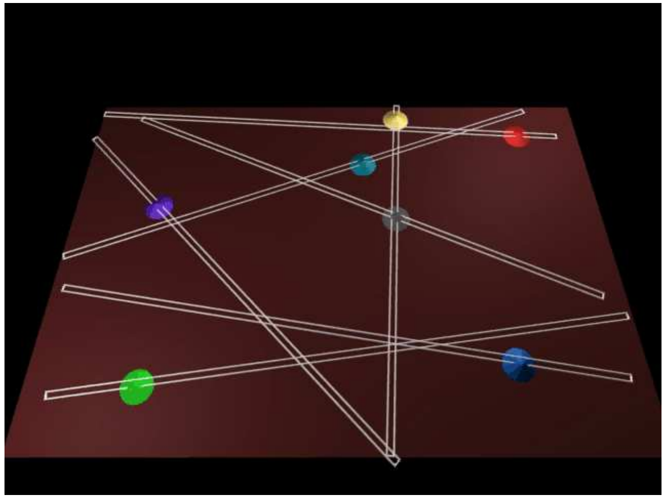
We need 13 points
(from both objects!)

$$\begin{bmatrix} \begin{pmatrix} 1 & 0 & -x_i^1 \end{pmatrix} \tilde{M}_1 \\ \begin{pmatrix} 0 & 1 & -y_i^1 \end{pmatrix} \tilde{M}_1 \\ \begin{pmatrix} 1 & 0 & -x_i^2 \end{pmatrix} \tilde{M}_2 \\ \begin{pmatrix} 0 & 1 & -y_i^2 \end{pmatrix} \tilde{M}_2 \\ l_3^T \tilde{M}_3 \end{bmatrix} \tilde{P}_i = 0$$

Once the tensor is recovered, 4 segmented points are sufficient to segment the entire scene

Another Application of $P^4 \rightarrow P^2$: 2D Dynamic Motion under const. velocity

Recall from Homography Tensors lecture:
26 triplets are required for general motion along straight line paths !



Points moving under const. velocity

Let:

$H_j,\ j=0,1,2$ three homography matrices

$p_j=(x_j,y_j,1)^\top \quad j=0,1,2$ projections of $(X,Y,1)$

Note: because of the const. Velocity constraint, the reconstruction could be up to 3D affine
So we are allowed to set the 3rd coordinate to be 1.

$$p_j \cong H_j \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix} + jH_j \begin{pmatrix} dX \\ dY \\ 0 \end{pmatrix} = \tilde{H}_j \begin{pmatrix} X \\ Y \\ 1 \\ dX \\ dY \end{pmatrix}$$

$$\tilde{H}_j = [H_j,\ jH_j^*]_{3\times 5}$$

13 matching triplets would be sufficient !

Multi-point, 3D, Constant Velocity

Let $P_i = (X_i, Y_i, Z_i, 1) \quad i = 1,\ldots,n$ be a 3D point configuration where each point moving independently along a direction

$$V_i = (dX_i, dY_i, dZ_i, 0)$$

At time $j = 0,\ldots,m$ the position of each point is $P_i + jV_i$

Let M_j be the 3×4 projection matrix

$$p_i^j \cong M_j(P_i + jV_i)$$

$$\begin{pmatrix} x_i^j \\ y_i^j \\ 1 \end{pmatrix} \cong [M_j \quad jM_j^*] \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \\ dX_i \\ dY_i \\ dZ_i \end{pmatrix} = \tilde{M}_j \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \\ dX_i \\ dY_i \\ dZ_i \end{pmatrix}$$

$$P^6 \rightarrow P^2$$

Multi-point, 3D, constant velocity

3x3x3x3 tensor (81 coefs)

Each quadruple of matching
Points p,p',p'',p'''
Provides 2 constraints

We need 40 points

$$\begin{bmatrix} \begin{pmatrix} 1 & 0 & -x_i^1 \end{pmatrix} \tilde{M}_1 \\ \begin{pmatrix} 0 & 1 & -y_i^1 \end{pmatrix} \tilde{M}_1 \\ \begin{pmatrix} 1 & 0 & -x_i^2 \end{pmatrix} \tilde{M}_2 \\ \begin{pmatrix} 0 & 1 & -y_i^2 \end{pmatrix} \tilde{M}_2 \\ \begin{pmatrix} 1 & 0 & -x_i^3 \end{pmatrix} \tilde{M}_3 \\ \begin{pmatrix} 0 & 1 & -y_i^3 \end{pmatrix} \tilde{M}_3 \\ l_4^T \tilde{M}_4 \end{bmatrix} \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \\ dX_i \\ dY_i \\ dZ_i \end{pmatrix} = 0$$

List of tensors

$\mathcal{P}^k$	Tensor Name	Size	# points.
$\mathcal{P}^3$	2D segmentation tensor	3^2	8
$\mathcal{P}^4$	2D constant velocity tensor	3^3	13
$\mathcal{P}^4$	3D segmentation tensor	3^3	13
$\mathcal{P}^4$	3D constant collincar vclocity	3^3	13
$\mathcal{P}^5$	3D constant coplanar velocity	3^3	26
$\mathcal{P}^6$	3D constant velocity tensor	3^4	40

Mathematical Background

The term *extensor* (cf. Barnabei,Brini,Rota 1985) is used to describe the linear space spanned by a collection of points.

Extensor of step 1 is a point
Extensor of step 2 is a line
Extensor of step 3 is a plane
Extensor of step k in P^k is a hyperplane

In P^n, the union ("join") of extensors of step k_1 and k_2 where $k_1 + k_2 \leq n+1$ is an extensor of step $k_1 + k_2$

The intersection ("meet") of extensors of step k_1 and k_2 is an extensor of step $k_1 + k_2 - (n+1)$

Mathematical Background

In P^n, the union ("join") of extensors of step k_1 and k_2 where $k_1 + k_2 \leq n+1$ is an extensor of step $k_1 + k_2$

The intersection ("meet") of extensors of step k_1 and k_2 is an extensor of step $k_1 + k_2 - (n+1)$

The following statements follow immediately:

1. The center of projection of a $P^k \rightarrow P^2$ projection is an extensor of step $k-2$

   Recall, the center of projection is the null space of the $3 \times (k+1)$ projection matrix.

2. The line of sight (image ray) joins the COP and a point (on the image plane):

   $P^3 \rightarrow P^2$ image ray is a line, extensor of step 1+1=2

   $P^4 \rightarrow P^2$ image ray is a plane, extensor of step 2+1=3

   $P^6 \rightarrow P^2$ image ray is an extensor of step 4+1=5

Mathematical Background

In P^n, the union ("join") of extensors of step k_1 and k_2 where $k_1 + k_2 \leq n+1$ is an extensor of step $k_1 + k_2$

The intersection ("meet") of extensors of step k_1 and k_2 is an extensor of step $k_1 + k_2 - (n+1)$

The following statements follow immediately:

3. The intersection of two lines of sight (triangulation) is the meet of the two extensors.

$P^3 \to P^2$	the intersection is generally empty 2+2-4=0
$P^4 \to P^2$	intersection is a point 3+3-5=1 (no constraint from two views!)
$P^6 \to P^2$	intersection is a plane 5+5-7=3

→ 4 views are required for a constraint.
(intersection of 3 rays: 3+5-7=1)

Mathematical Background

In P^n, the union ("join") of extensors of step k_1 and k_2 where $k_1 + k_2 \leq n+1$ is an extensor of step $k_1 + k_2$

The intersection ("meet") of extensors of step k_1 and k_2 is an extensor of step $k_1 + k_2 - (n+1)$

The following statements follow immediately:

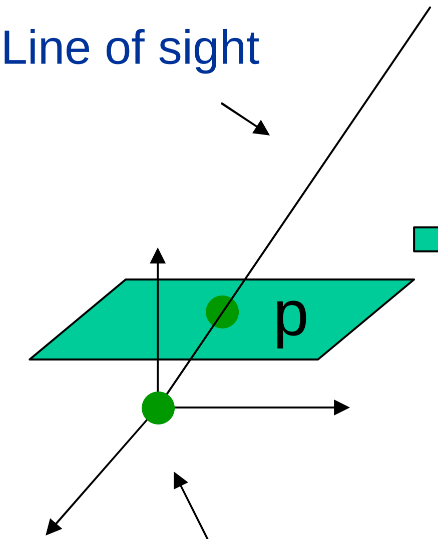
4. The epipole is the intersection between the join of the two COPs and an image plane.

$P^3 \rightarrow P^2$	the epipole is a point (1+1) +3 - 4 =1
$P^4 \rightarrow P^2$	the epipole is a line (2+2) + 3 -5=2
$P^6 \rightarrow P^2$	the epipole is not defined because 4+4 > 7 (the join of the two COPs is larger than the space dimension.

The Geometry of $P^4 \rightarrow P^2$ Projection Matrices

$P^3 \rightarrow P^2$

Line of sight



Center of projection

$$null(M_{3\times 4})$$

$P^4 \rightarrow P^2$

- Center of Projection is a Line $null(\tilde{M}_{3\times 5})$
- Line of sight is a plane
- Two planes intersect (always) at a point

$$3+3-5=1$$

⇨ Three views are necessary for a multilinear constraint

- A plane is the intersection of two hyperplanes

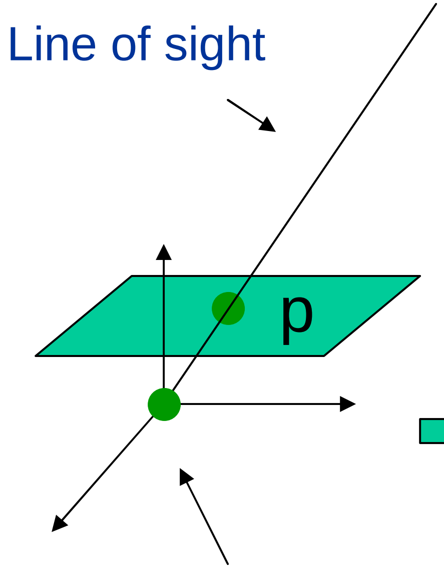
$$4+4-5=3$$

$$C_\pi = \begin{bmatrix} \pi_1 \\ \pi_2 \end{bmatrix}_{2\times 5} \Rightarrow C_\pi P = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \forall P \in \pi$$

The Geometry of $P^6 \rightarrow P^2$ Projection Matrices

$P^6 \rightarrow P^2$

$P^3 \rightarrow P^2$



$null(M_{3\times 4})$

- Center of Projection is a 4D-space $null(\tilde{M}_{3\times 7})$
- Line of sight is a 5D-space
- Two line-of-sights intersect in a plane

$$5+5-7=3$$

- A line-of-sight and a plane intersect at a point

$$5+3-7=1$$

⇨ Four views are necessary for a multilinear constraint

The Geometry of $P^4 \rightarrow P^2$

The coefficients of the determinant expansion
Live in a mixed tensor of the type

$$A_{ij}^{k}$$

The constraint: $p^i p'^j l''_k A_{ij}^{k} = 0$

➡ A triplet of matching points provides 2 constraints

$$\begin{bmatrix} \begin{pmatrix} 1 & 0 & -x_i^1 \end{pmatrix} \tilde{M}_1 \\ \begin{pmatrix} 0 & 1 & -y_i^1 \end{pmatrix} \tilde{M}_1 \\ \begin{pmatrix} 1 & 0 & -x_i^2 \end{pmatrix} \tilde{M}_2 \\ \begin{pmatrix} 0 & 1 & -y_i^2 \end{pmatrix} \tilde{M}_2 \\ l_3^T \tilde{M}_3 \end{bmatrix}_{5\times 5} \tilde{P}_i = 0$$

$$p^i p'^j A_{ij}^{k} \cong p''^k$$

↓ plane ↓ plane

intersection of two planes: 3+3-5=1 is a point (projected onto p'')

Goal: given A_{ij}^{k} (recovered from 13 matching triplets) we wish to find the three 3x5 projection matrices, and then to recover the 3D-2D underlying matrices.

The Geometry of $P^4 \to P^2$

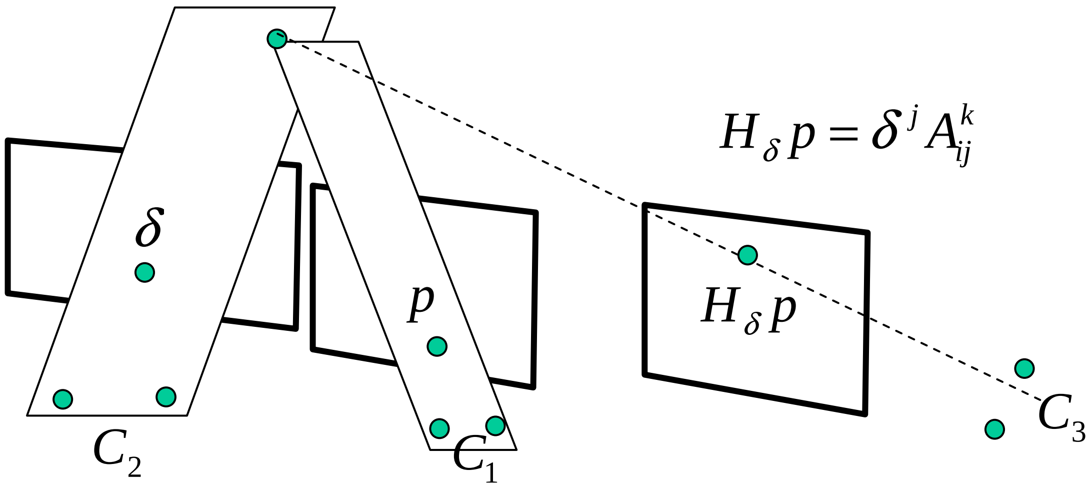
Homography slices:

$$\delta^j A_{ij}^k \cong H_i^k$$

a homography matrix $1 \to 3$ induced by the plane $C_2 \vee \delta$

$$\delta^i A_{ij}^k \cong E_j^k$$

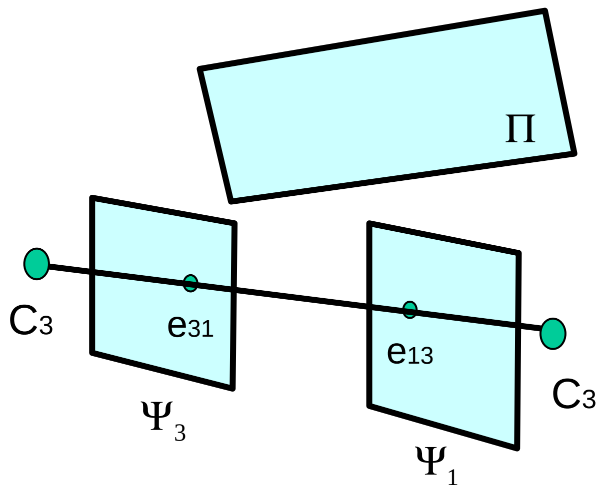
a homography matrix $2 \to 3$ induced by the plane $C_1 \vee \delta$



The Geometry of $P^4 \rightarrow P^2$

Recovering Epipoles from Homography slices: $e_{ij} = \tilde{M}_i(null(\tilde{M}_j))$

$P^3 \rightarrow P^2$



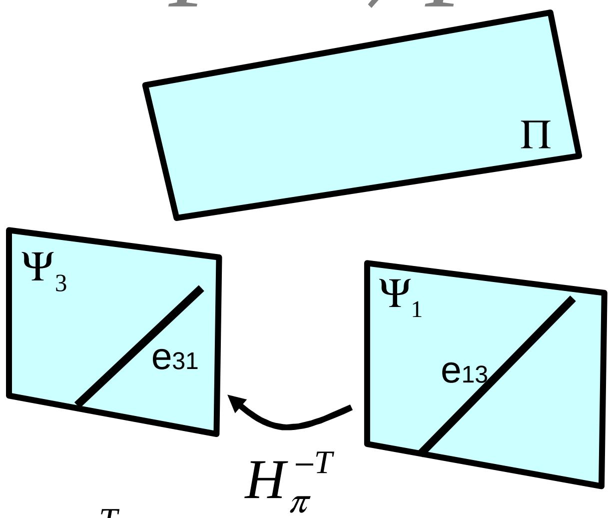
$H_\pi e_{13} \cong e_{31}$

$\Rightarrow (H_{\pi_1} - \lambda H_{\pi_2})e_{13} \cong 0$

$\Rightarrow e_{13}$ is a generelized

eigen - vector of (H_{π_1}, H_{π_2})

$P^4 \rightarrow P^2$



$H_\pi^{-T} e_{13} \cong e_{31}$

$\Rightarrow e_{13}$ is a generelized

eigen - vector of $(H_{\pi_1}^{-T}, H_{\pi_2}^{-T})$

(slices of the form : $\delta^j A_{ij}^k$)

The Geometry of $P^4 \to P^2$

The canonical form of projection matrices:

Claim: There exist a canonical projective frame such that:

$$p \cong [I_{3x3} \quad 0_3 \quad 0_3] \begin{pmatrix} p \\ \eta \\ \mu \end{pmatrix} \qquad p' = [H_\pi \quad e_1 \quad e_2] \begin{pmatrix} p \\ \eta \\ \mu \end{pmatrix}$$

Where H_π is a homography matrix from views 1 to 2, and e_1, e_2 are two points on the epipole (a line) on view 2.

The family of homography matrices between views 1 to 2 has the general Form with 7 degrees of freedom:

$$\lambda H_\pi + e_1 n_1^T + e_2 n_2^T$$

The Geometry of $P^4 \to P^2$

The canonical form of projection matrices:

Proof:

Let $\tilde{M}_1, \tilde{M}_2$ be the two 3x5 projections matrices, a point P in space and matching points p,p' satisfying:

$$p \cong \tilde{M}_1 P \qquad p' \cong \tilde{M}_2 P$$

Let C,C' be two points spanning the center-of-projection of the first camera: $\tilde{M}_1 C = 0$, $\tilde{M}_1 C' = 0$

Let a 5x5 projective change of basis W be defined: $W = [U, C, C']$ where U is some 5x3 matrix chosen such that $\tilde{M}_1 U = I_{3\times 3}$

$$\Rightarrow \quad \tilde{M}_1 W = [I_{3\times 3}; 0_{3\times 2}]$$

$$\tilde{M}_2 W = \tilde{M}_2 [U, C, C'] = [\tilde{M}_2 U, e_1, e_2]$$

Our goal will be to create the 5x3 matrix U such that $\tilde{M}_2 U$ is a homography matrix.

The Geometry of $P^4 \rightarrow P^2$

The canonical form of projection matrices:

Proof:

Consider the following construction of U: $U = \begin{bmatrix} \tilde{M}_1 \\ C_\pi \end{bmatrix}^{-1}_{1-3}$ (cols 1-3 from the 5x5 matrix)

Note: $\tilde{M}_1 U = I_{3\times 3}$

We have: $$\begin{bmatrix} \tilde{M}_1 \\ C_\pi \end{bmatrix} P = \begin{pmatrix} \tilde{M}_1 P \\ C_\pi P \end{pmatrix} \cong \begin{pmatrix} p \\ 0 \\ 0 \end{pmatrix} \qquad \forall P \in \pi$$

$$\Rightarrow \quad P = \begin{bmatrix} \tilde{M}_1 \\ C_\pi \end{bmatrix}^{-1} \begin{pmatrix} p \\ 0 \\ 0 \end{pmatrix}$$

$$(\tilde{M}_2 U) p = \tilde{M}_2 \begin{bmatrix} \tilde{M}_1 \\ C_\pi \end{bmatrix}^{-1} \begin{pmatrix} p \\ 0 \\ 0 \end{pmatrix} = \tilde{M}_2 P \cong p' \qquad \forall P \in \pi$$

We have thus shown that $\tilde{M}_2 U p \cong p'$ for all matching points arising from the plane π

Comparison With $P^3 \to P^2$

$P^3 \to P^2$:

$$\tilde{M}_1 = [I_{3x3} \quad 0_3]$$

$$\tilde{M}_2 = [H \quad e]$$

A basis is determined by 5 points:

H determines 3 points.

C_1 determines 1 point.

Scales between H, e: adds 1.

$P^4 \to P^2$:

$$\tilde{M}_1 = [I_{3x3} \quad 0_3 \quad 0_3]$$

$$\tilde{M}_2 = [H_{12} \quad e_1 \quad e_2]$$

A basis is determined by 6 points:

H_{12} determines 3 points.

C_1 determines 2 point (a line).

Scales between H_{12}, e_1, e_2: adds 1.

Note: the third camera matrix $\tilde{M}_3$ can be recovered from $\tilde{M}_1, \tilde{M}_2$ and the tensor A_{ij}^k

Performing the Segmentation

- Points from each object are located on a different hyper-plane.
- Points on the first object are on a hyper-plane: $\begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \\ 0 \end{pmatrix}$
- Four points are sufficient to recover this hyper-plane.

Take 4 matching points $p_i,\ p'_i,\ p''_i \quad i = 1,2,3,4$ which are known to come from one of the objects and reconstruct the points $\tilde{P}_i \quad i = 1,2,3,4$

Solve for the hyperplane $[\tilde{P}_1, \tilde{P}_2, \tilde{P}_3, \tilde{P}_1]\pi = 0$

Given a new matching triplet p, p', p'' reconstruct $\tilde{P}$ and if $\tilde{P}^\top \pi = 0$ then the point belongs to object 1.

Recovery of structure

- Done in two stages:
  - Reconstruction of 3x(k+1) camera matrices per view in P^k
  - Reconstruction of the constituent 3x4 camera matrices in 3D.
- We get Affine 3D structure

Experiments

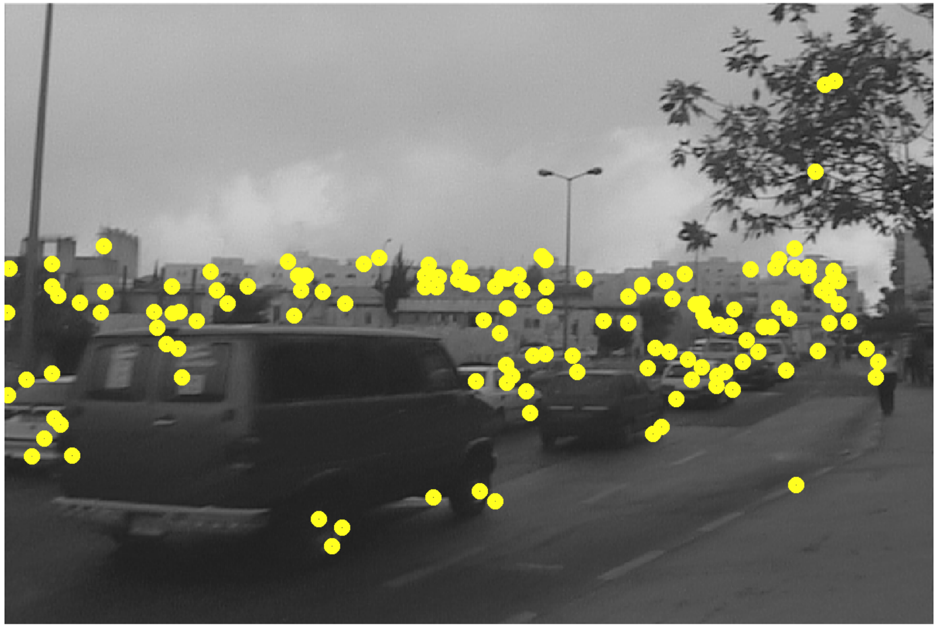
QuickTime™ and a
Microsoft Video 1 decompressor
are needed to see this picture.

Traditional solution

- Pick 7 points from each object.
- For each object compute the trifocal tensor separately.
- Does not require relative translation.

Computing the tensor

- We use all points to compute tensor.



segmenting

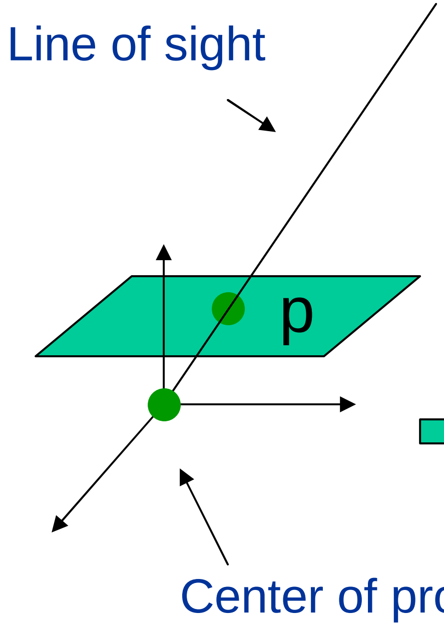
- Only 4 known points from one object are needed. (instead of 7)

Results

The Geometry of $P^6 \rightarrow P^2$ Projection Matrices

$$P^6 \rightarrow P^2$$

$$P^3 \rightarrow P^2$$



$$null(M_{3\times 4})$$

- Center of Projection is a 4D-space $null(\tilde{M}_{3\times 7})$
- Line of sight is a 5D-space
- Two line-of-sights intersect in a plane

$$5+5-7=3$$

- A line-of-sight and a plane intersect at a point

$$5+3-7=1$$

⇨ Four views are necessary for a multilinear constraint

Joint Epipole

- There is no such thing as epipole: the join of the two camera centers is larger then the whole space.
- We use *joint epipole:* the projection of the intersection of two camera centers.
- C_{12}^{3} :the projection of the intersection of the first two camera centers onto the third view.

Reconstruction of projection matrices from the tensor

$$\tilde{M}_1 \cong \begin{bmatrix} I_{3\times3} & 0_{3\times3} & 0_{3\times1} \end{bmatrix}$$
$$\tilde{M}_2 \cong \begin{bmatrix} 0_{3\times3} & I_{3\times3} & 0_{3\times1} \end{bmatrix}$$
$$\tilde{M}_3 \cong \begin{bmatrix} H_{13} & H_{23} & c^3_{21} \end{bmatrix}$$

$\tilde{M}_4$ can be determined linearly from the tensor and $\tilde{M}_1$, $\tilde{M}_2$, M_3

where:

H_{13} is a homography $1 \rightarrow 3$ of a plane contained in C_2

H_{23} is a homography $2 \rightarrow 3$ of a plane contained in C_1

c^3_{21} is the joint epipole of C_1, C_2 in view 3.

Constant Velocity Experiments

QuickTime™ and a
Microsoft Video 1 decompressor
are needed to see this picture.

The scene is 3D, The velocities are coplanar:
A situations for a $P^5 \rightarrow P^2$ tensor

Computing the tensor

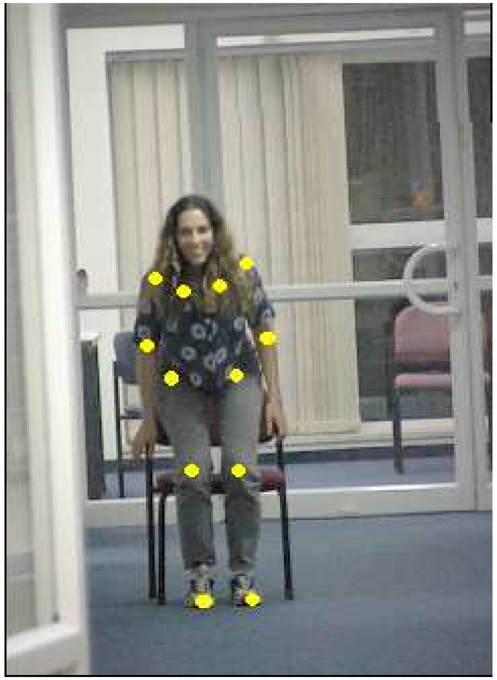
- We use all tracked points across three images to compute the tensor.

Results: 3-D velocity projected to first view

- Arrow's base – point in first view
- Arrow's tip – 3-D point in time = 2 projected to the first view

Indexing into "Action"

QuickTime™ and a
Microsoft Video 1 decompressor
are needed to see this picture.



Training Sequence

Single View:

is this a view of a person sitting?

Algebraic Polynomials for Indexing into Action

$$f(p_1,\ldots,p_k, j) = 0$$

If the collection of image points at time j are a projection of the dynamic configuration.

Example: Affine Camera Model

$$\begin{pmatrix} x_{ij} \\ y_{ij} \end{pmatrix} = \begin{bmatrix} a_j \\ b_j \end{bmatrix}_{2\times 4} \begin{pmatrix} P_i + jV_i \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} x_{ij} \\ y_{ij} \end{pmatrix} = \begin{bmatrix} a_j \\ b_j \end{bmatrix}_{2\times 4} \begin{pmatrix} P_i + jV_i \\ 1 \end{pmatrix} \Rightarrow$$

$$\begin{pmatrix} P_i^T + jV_i^T & 1 & x_{ij} \end{pmatrix} \begin{pmatrix} a_j \\ -1 \end{pmatrix} = 0$$

$$\begin{pmatrix} P_i^T + jV_i^T & 1 & y_{ij} \end{pmatrix} \begin{pmatrix} b_j \\ -1 \end{pmatrix} = 0$$

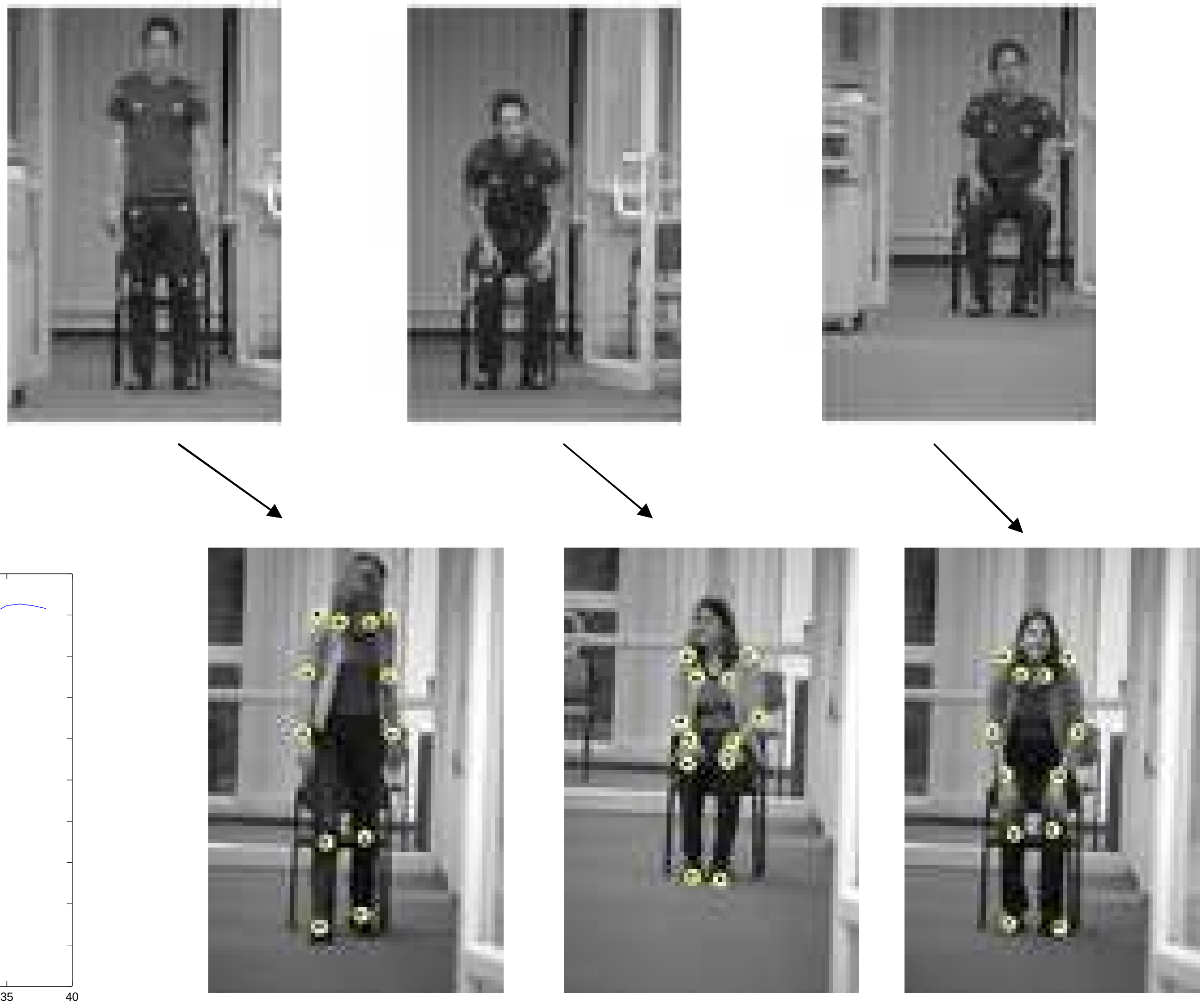
$$\begin{vmatrix} P_1^T + jV_1^T & 1 & 0 \\ P_2^T + jV_2^T & 1 & 1 \\ P_3^T + jV_3^T & 1 & 0 \\ P_4^T + jV_4^T & 1 & x_{4j} \\ P_5^T + jV_5^T & 1 & x_{5j} \end{vmatrix} = 0 \qquad \begin{vmatrix} P_1^T + jV_1^T & 1 & 0 \\ P_2^T + jV_2^T & 1 & 0 \\ P_3^T + jV_3^T & 1 & 1 \\ P_4^T + jV_4^T & 1 & y_{4j} \\ P_5^T + jV_5^T & 1 & y_{5j} \end{vmatrix} = 0$$

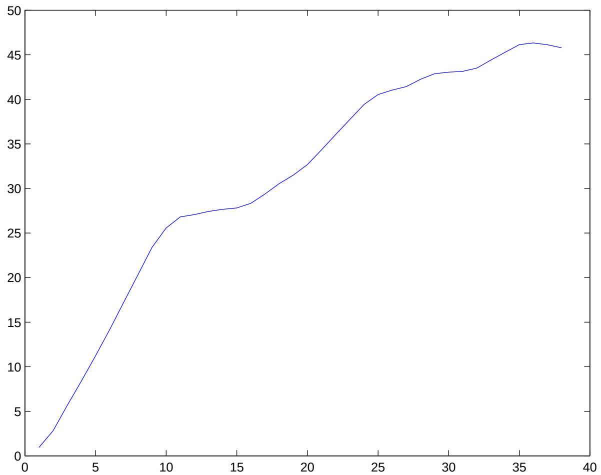
$$f_1(p_4, p_5, j) = 0 \qquad f_2(p_4, p_5, j) = 0$$

The coefficients of these polynomials depend on shape P1,…,P5 and action V1,…,V5

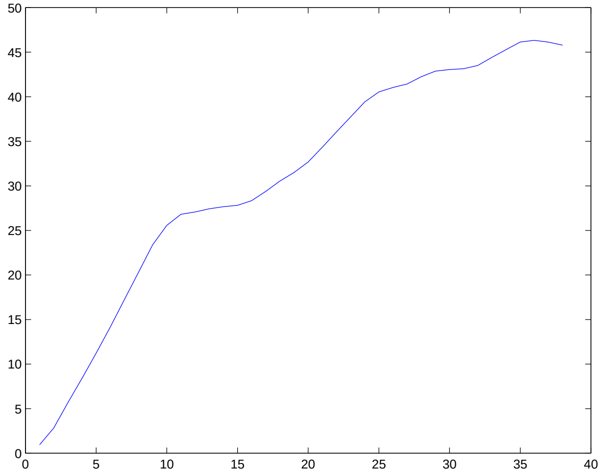
These coefficients can be recovered using 8 views.

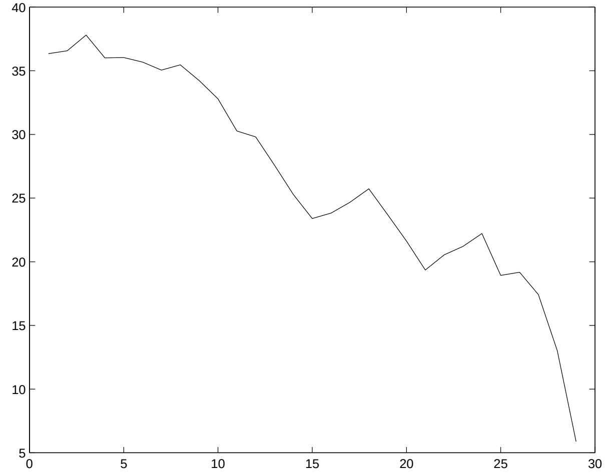
Model Sequence



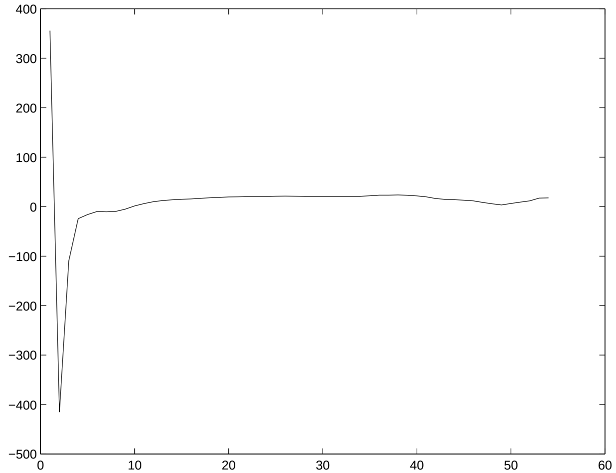


“time” increases monotonically for motions of the same “class”



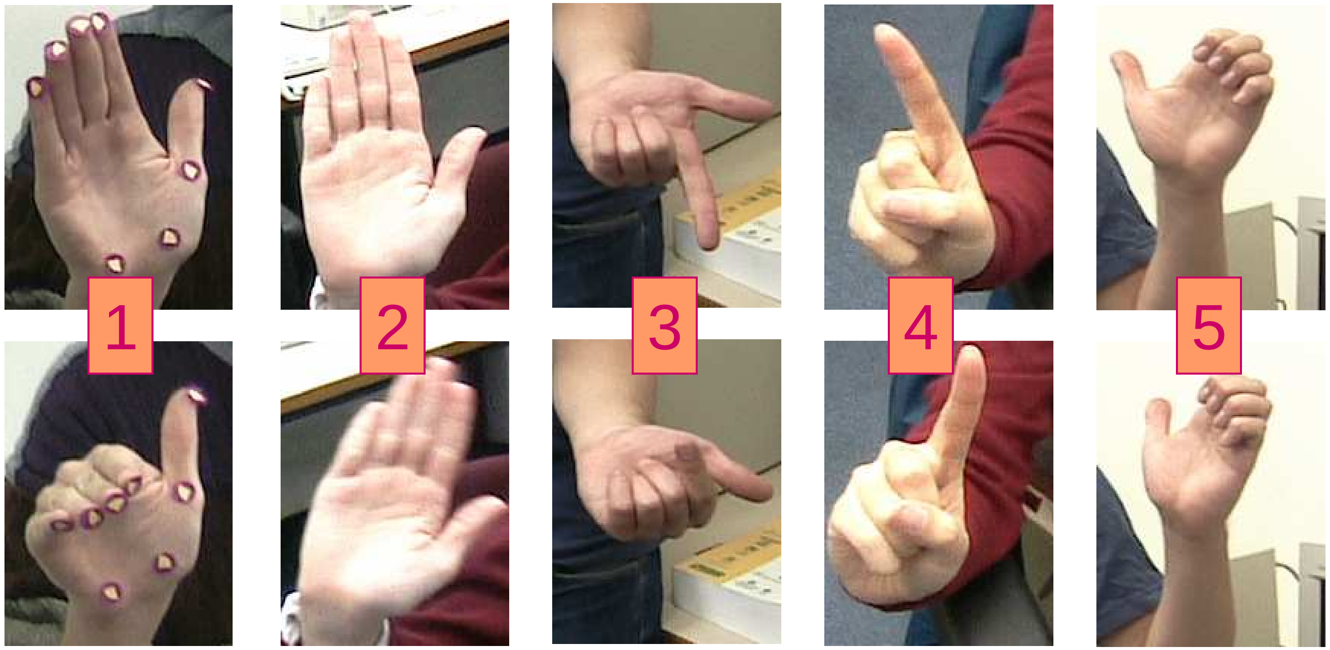


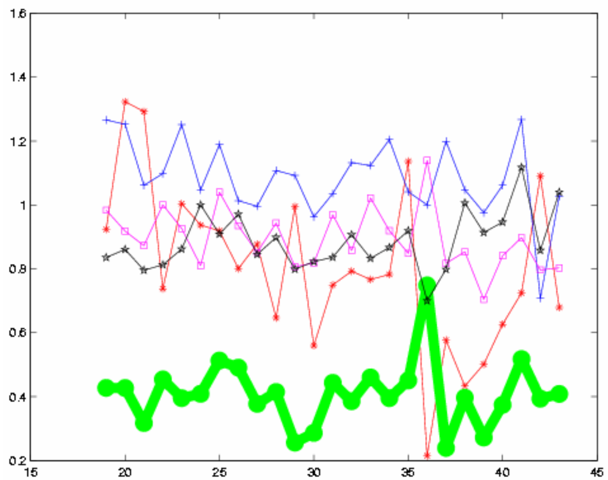
change of j for a "sitting" motion **change of j for a "standing-up" motion**



change of j for a "walking" motion using "sitting" polynomials

$$P_i = O_i + \cos(\lambda j)U_i + \sin(\lambda j)V_i$$





Residuals of 5 tensors on images of Action 1

The End