STRUCTURE COMPUTATION FROM TWO IMAGES.
Thank you for the slides.
They come mostly from the following source.

Marc Pollefeys U. of North Carolina
3D reconstruction of cameras and structure

reconstruction problem:

given $x_i \leftrightarrow x'_i$, compute $P, P'$ and $X_i$

$$x_i = PX_i \quad x'_i = P'X'_i \quad \text{for all } i$$

without additional information possible up to projective ambiguity
outline of reconstruction

(i) Compute F from correspondences
(ii) Compute camera matrices from F
(iii) Compute 3D point for each pair of corresponding points

computation of F
use \( x'_i F x_i = 0 \) equations, linear in coeff. F
8 points (linear), 7 points (non-linear), 8+ (least-squares)

computation of camera matrices
use \( P = [I | 0] \quad P' = [[e']_x F + e'v^T | \lambda e'] \) general form

triangulation
compute intersection of two backprojected rays
Reconstruction ambiguity: \textit{calibrated cameras}

\textit{Similarity}

\[ x_i = PX_i = \left( PH_S^{-1} \right) H_S X_i \]

\[ PH_S^{-1} = K[R \mid t] \begin{bmatrix} R^T & -R^T t' \lambda \end{bmatrix} = K[RR^T \mid -RR^T t' + \lambda t] \]

\( H_S = \begin{bmatrix} R' & t' \\ 0^T \lambda \end{bmatrix} \)

\( \lambda \) - overall scaling

\textit{essential matrices}

\( H_S \) (4 x 4) matrix

\( \text{does not change K} \)

\( \text{does not change measured angles the epipoles remain the same} \)
Reconstruction ambiguity: projective

The angle between the points changes but the epipoles (intersection with the baseline) is not.

\[ x_i = PX_i = (PH_p^{-1})H_p X_i \]

The points unchanged. \( H_p \) is a (4 x 4) matrix.
Terminology

$x_i \leftrightarrow x'_i$

Original scene $X_i$

Projective, affine, similarity reconstruction
  = reconstruction that is identical to original up to projective, affine, similarity transformation

Literature: Metric and Euclidean reconstruction
  = similarity reconstruction
The projective reconstruction theorem

If a set of point correspondences in two views determine the fundamental matrix uniquely, then the scene and cameras may be reconstructed from these correspondences alone, and any two such reconstructions from these correspondences are projectively equivalent.

\[ x_i \leftrightarrow x_i' \quad (P_1, P_1', \{X_{1i}\}) = (P_2, P_2', \{X_{2i}\}) \] (same F)

was H not \(H^{-1}\) not important!

\[ P_2 = P_1 H^{-1} \quad P_2' = P_1' H^{-1} \quad X_2 = H X_1 \quad (\text{except: } F x_i = x_i' F = 0) \]

H is \((4 \times 4)\) nonsingular

theorem from last class

\[ P_2 (H X_{1i}) = P_1 H^{-1} H X_{1i} = P_1 X_{1i} = x_i = P_2 X_{2i} \]

\[ \Rightarrow \text{along same ray of } P_2, \]

two possibilities: \(X_{2i} = H X_{1i}\), or points along baseline

key result:

allows reconstruction from pair of uncalibrated images
projective reconstruction

the matrix F is computed followed by triangulation in 3D. These are views of the points in 2D.
Stratified reconstruction

(i) Projective reconstruction
(ii) Affine reconstruction
(iii) Metric reconstruction
Projective to affine

remember 2-D case

we need to locate the plane at infinity to have an affine reconstruction

remember the first homework!
Projective to affine

\[(P, P', \{X_i\})\]

\[\pi_\infty = (A, B, C, D)^T \leftrightarrow (0,0,0,1)^T\]

\[H^T\pi_\infty = (0,0,0,1)^T\]  

plane transformation in 3D

\[H = \begin{bmatrix} 1 & 0 \\ \pi_\infty \\ \end{bmatrix} \]

will not work if \(\pi_\infty, 4\) is zero then \(H\) is singular and Householder matrices are needed 2HZSec.A4.1.3

\(pi_\text{inf}\) is identified as the plane at infinity

\(pi_1 = H^{-T} pi_0\)

theorem says up to projective transformation, but projective with fixed \(\pi_\infty\) is affine transformation

\(H\) applied to all the points and the two cameras

can be sufficient depending on application, e.g. mid-point, centroid, parallelism
A4.1.2 Householder matrices and QR decomposition

For matrices of larger dimension, the QR decomposition is more efficiently carried out using Householder matrices. The symmetric matrix

\[ H_v = I - 2v v^T / v^T v \]  

has the property that \( H_v^T H_v = I \), and so \( H_v \) is orthogonal.

Let \( e_1 \) be the vector \((1, 0, \ldots, 0)^T\), and let \( x \) be any vector. Let \( v = x \pm \|x\| e_1 \). One easily verifies that \( H_v x = \mp \|x\| e_1 \); thus \( H_v \) is an orthogonal matrix that transforms the vector \( x \) to a multiple of \( e_1 \). Geometrically \( H_v \) is a reflection in the plane perpendicular to \( v \), and \( v = x \pm \|x\| e_1 \) is a vector that bisects \( x \) and \( \mp \|x\| e_1 \). Thus reflection in the \( v \) direction takes \( x \) to \( \mp \|x\| e_1 \). For reasons of stability, the sign ambiguity in defining \( v \) should be resolved by setting

\[ v = x + \text{sign}(x_1) \|x\| e_1. \]  

If \( A \) is a matrix, \( x \) is the first column of \( A \), and \( v \) is defined by (A4.3), then forming the product \( H_v A \) will clear out the first column of the matrix, replacing the first column by \( (\|x\|, 0, 0, \ldots, 0)^T \). One continues left multiplication by orthogonal Householder matrices to clear out the below-diagonal part of the matrix \( A \). In this way, one finds that eventually \( QA = R \), where \( Q \) is a product of orthogonal matrices and \( R \) is an upper-triangular matrix. Therefore, one has \( A = Q^T R \). This is the QR decomposition of the matrix \( A \).

When multiplying by Householder matrices it is inefficient to form the Householder matrix explicitly. Multiplication by a vector \( a \) may be carried out most efficiently as

\[ H_v a = (I - 2v v^T / v^T v) a = a - 2v(v^T a) / v^T v \]  

and the same holds for multiplication by a matrix \( A \). For more about Householder matrices and the QR decomposition, the reader is referred to [Golub-89].

Note. In the QR or RQ decomposition, \( R \) refers to an upper-triangular matrix and \( Q \) refers to an orthogonal matrix. In the notation used elsewhere in this book, \( R \) refers usually to a rotation (hence orthogonal) matrix.

A4.2 Symmetric and skew-symmetric matrices

Symmetric and skew-symmetric matrices play an important role in this book. A matrix is called \textit{symmetric} if \( A^T = A \) and \textit{skew-symmetric} if \( A^T = -A \). The eigenvalue decompositions of these matrices are summarized in the following result.

Result A4.1. Eigenvalue decomposition.

(i) If \( A \) is a real symmetric matrix, then \( A \) can be decomposed as \( A = U D U^T \), where \( U \) is an orthogonal matrix and \( D \) is a real diagonal matrix. Thus, a real symmetric matrix has real eigenvalues, and the eigenvectors are orthogonal.

(ii) If \( S \) is real and skew-symmetric, then \( S = U B U^T \) where \( B \) is a block-diagonal
Translational motion

points at infinity are fixed for a pure translation
⇒ reconstruction of $x_i \leftrightarrow x_i$

affine reconstruction directly!!
three points 3D (and not four) are enough

a point first image and the same point in the second image IS on the plane at infinity

\[
F = [e]_x = [e']_x \quad P = [I | 0] \\
= [I | e'] \quad \text{no rotation}
\]
Scene constraints

Parallel lines
parallel lines intersect at infinity
reconstruction of corresponding vanishing point yields
point on plane at infinity

3 sets of parallel lines allow to uniquely determine $\pi_\infty$

remark: in presence of noise determining the intersection of parallel lines is a delicate problem (will see later)

remark: obtaining vanishing point in one image can be sufficient

3D point $X$ with vanishing points $v$ satisfies both equations
(point in the first a line in the second)

$[v]_{\mathcal{M}}(P X) = 0 = l'^* T (P' X)$
Scene constraints

- Using parallel lines: affine reconstruction is possible.
- Three sets of parallel lines.
The infinity homography

\[ P = [M \mid m] \quad P' = [M' \mid m'] \]

\[ H_\infty = M'M^{-1} \]

for points on plan at infinity \( X_\text{inf} \) the \( x'_\text{inf} = M'M^{-1}x_\text{inf} \)

unchanged under affine transformations

\[ P = [M \mid m] \begin{bmatrix} A & a \\ 0 & 1 \end{bmatrix} = [MA \mid Ma + m] \]

\[ H_\infty = M'AA^{-1}M^{-1} \]

affine reconstruction

\[ P = [I \mid 0] \quad P' = [H_\infty \mid e'] \]

if \( H_\text{inf} \) is already obtained

from the cameras \( P = [I \mid 0] \quad P' = [M' \mid e'] \) than \( H_\text{inf} = M' \)
Affine to metric

identify absolute conic \(\Omega_\infty\) on the plane at infinity

transform so that \(\Omega_\infty : X^2 + Y^2 + Z^2 = 0\), on \(\pi_\infty\)

then projective transformation relating original and reconstruction is a similarity transformation

in practice, find image of \(\Omega_\infty\)
image \(\omega_\infty\) back-projects to cone that intersects \(\pi_\infty\) in \(\Omega_\infty\)

note that image is independent of particular reconstruction
Affine to metric

given \( P = [M \mid m] \) \( \omega \) image of the absolute conic (IAC) to be known

possible transformation from affine to metric is

\[
H = \begin{bmatrix}
A^{-1} & 0 \\
0 & 1 \\
\end{bmatrix}
\]

\[
AA^T = \left( M^T \omega M \right)^{-1}
\]

(cholesky factorisation)

proof:

new calibration matrix, \( P^* \) also changes, \( F \) does not

\[
P_M = PH^{-1} = [MA \mid m]
\]

\[
\omega^* = M_M M_M^T = MAA^T M^T = \Omega^* = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}
\]

must identify the IAC in order to do metric reconstruction
Orthogonality

vanishing points corresponding to orthogonal directions

\[ v_1^T \omega v_2 = 0 \]

dir = \omega

vanishing line and vanishing point corresponding to plane and normal direction

\[ 1 = \omega v \]

e.g. vanishing point for the vertical direction and the vanishing line from the horizontal ground plane
known internal parameters

\[
\omega = K^{-T}K^{-1}
\]

rectangular pixels
\[
s = 0 \quad \omega_{12} = \omega_{21} = 0
\]

square pixels
\[
\alpha_x = \alpha_y \quad \omega_{11} = \omega_{22}
\]

see also Table 18.1 (2ed. Table 19.1) for an expression of omega

then three unknowns only
3. Same camera for all images

same intrinsics $\Rightarrow$ same image of the absolute conic

e.g. moving cameras

given sufficient images there is in general only one conic that projects to the same image in all images, i.e. the absolute conic

This approach is called self-calibration, see later.

Transfer of IAC: $\omega' = H_{\infty}^{-T} \omega H_{\infty}^{-1}$

Transformation of a conic

Know the plane at infinity therefore know $H_{\text{inf}}$

Omega $\sim$ Omega’ symmetric linear equations with five d.o.f. in omega and solves only for four!

Other constraints also needed.
metric (similarity) reconstruction

with the image of absolute conic, orthogonality also is taken into account the two images has square pixels, transferred with H_inf.

the aspect ratios are also respected

texture mapped piecewise planar models built from the wireframes
Direct metric reconstruction using $\omega$

if you know omega-s in at least two views you can go from projective to metric

**approach 1**

$$\omega = K^{-T}K^{-1} \Rightarrow K$$

calibrated reconstruction

but essential matrix only one of the four solutions is right

**approach 2**

compute projective reconstruction $P$ and $P'$

back-project $\omega$ from both images $\Rightarrow$ absolute conic $\Omega_{\infty}$

intersection defines $\Omega_{\infty}$ and its support plane $\pi_{\infty}$

(in general two solutions)

since the intersection not perfect
Direct reconstruction using ground truth

use control points $X_{Ei}$ with known coordinates
to go from projective to metric

points with known 3D location
in Euclidean world frame

from $x_i \leftrightarrow x'_i$ we obtain $X_i$ up to a homography
of 15 d.o.f

we need $n \geq 5$ (no four coplanar) points

$X_{Ei} = HX_i$

other way...

$x_i = PH^{-1}X_{Ei}$

(2 lin. eq. in $H^{-1}$ per view,
if $X_i$ visible in both view only $x_i$, $x'_i$
only 3 eqn. because of the coplanarity

once $H$ is computed $P$ and $P'$ can be
transformed into Euclidean matrices
Objective
Given two uncalibrated images compute \((P_M, P'_M, \{X_{Mi}\})\) (i.e. within similarity of original scene and cameras)

Algorithm
(i) Compute projective reconstruction \((P, P', \{X_i\})\)
   (a) Compute \(F\) from \(x_i \leftrightarrow x'_i\)
   (b) Compute \(P, P'\) from \(F\)
   (c) Triangulate \(X_i\) from \(x_i \leftrightarrow x'_i\)
(ii) Rectify reconstruction from projective to metric

**Direct method:** compute \(H\) from control points \(X_{Ei} = HX_i\)

\[
P_M = PH^{-1} \quad P'_M = P'H^{-1} \quad X_{Mi} = HX_i
\]

**Stratified method:**
(a) **Affine reconstruction:** compute \(\pi_\infty\)

\[
H = \begin{bmatrix} 1 & 0 \\ \pi_\infty^T & 1 \end{bmatrix}
\]

(b) **Metric reconstruction:** compute \(\text{iAC} \ \omega\)

\[
H = \begin{bmatrix} A^{-1} & 0 \\ 0 & 1 \end{bmatrix}
\]

\[
AA^T = (M^T \omega M)^{-1}
\]

* \(P = [I \mid 0] \quad P' = [M \mid e']\)
  * three vanishing points + \(F\)
  => \(H_{\inf} \rightarrow M\) (see later)
  multiply all with the \(H_1\)
  * find \(\omega\) => metric
  multiply all with the \(H_2\)
<table>
<thead>
<tr>
<th>Image information provided</th>
<th>View relations and projective objects</th>
<th>3-space objects</th>
<th>reconstruction ambiguity</th>
</tr>
</thead>
<tbody>
<tr>
<td>point correspondences</td>
<td>F</td>
<td>π∞</td>
<td>projective</td>
</tr>
<tr>
<td>point correspondences</td>
<td>F,H∞</td>
<td>π∞</td>
<td>affine</td>
</tr>
<tr>
<td>including vanishing points</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Points correspondences</td>
<td>F,H∞</td>
<td>π∞</td>
<td>metric</td>
</tr>
<tr>
<td>and internal camera</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>calibration</td>
<td>ω,ω′</td>
<td>Ω∞</td>
<td></td>
</tr>
</tbody>
</table>
Point reconstruction

\[ x = PX \]

\[ x' = P' X \]

The two points \( x \) and \( x' \) are not corresponding.
linear triangulation

\[ x = PX \quad x' = P'X \]

**find X in 3D**

\[ x \times P^{-1}X = 0 \]

\[ A = \begin{bmatrix} \quad x p_{3T} - p_{1T} \\ y p_{3T} - p_{2T} \\ x' p_{3T} - p'_{1T} \\ y' p_{3T} - p'_{2T} \end{bmatrix} \]

**homogeneous**

\[ \| X \| = 1 \]

**inhomogeneous**

\[ (X, Y, Z, 1) \]

solution \( HX \) is not the same as \( X \)

\[ (AH^{-1})(HX) = e \]

\[ \text{AFFINE transformation} \]

\( X = (X, Y, Z, 1) \) is preserved \( HX = (X', Y', Z', 1) \)

inhomogeneous method affine invariant but homogeneous method \( \| X \| = 1 \) NOT.
geometric error

\[ d(x, \hat{x})^2 + d(x', \hat{x}')^2 \] subject to \[ \hat{x}^T F \hat{x} = 0 \]

or equivalently subject to \[ \hat{x} = P \hat{X} \] and \[ \hat{x}' = P' \hat{X} \]

possibility to compute using LM (for 2 or more points)

or directly (for 2 points)

Sampson approximation
2HZSec.12.4
Optimal 3D point in epipolar plane

Given an epipolar plane, find best 3D point for \((x_1, x_2)\)

- Select closest points \((x'_1, x'_2)\) on epipolar lines
- Obtain 3D point through exact triangulation
- Guarantees minimal reprojection error (given this epipolar plane)
Reconstruction uncertainty

consider angle between rays

depends on the angle between the two images
forward motion gives poor reconstruction
Some experiments performed with these two images
Reconstruction with two, upper is with the midpoint on the common perpendicular, lower with the optimal polynomial method. Both have large errors if the noise is large.
Line reconstruction

\[ \pi = P^T l \quad \pi' = P'^T l' \] the planes defined by a line and the camera matrix

\[ L = \begin{bmatrix} l^T P \\ l'^T P' \end{bmatrix} \]

doesn’t work for epipolar plane

intersection of more than two planes
\( A \) is \( n \) planes, one per row
\( A = U D V^T \) \( \rightarrow \) Two largest singular values correspond to the line \( L \)

\( l \) and \( l' \) are NOT epipolar lines

the two planes intersect in \( L \)
Computing vanishing points.

Maximum likelihood estimation determining a set of lines intersect in a single point. Levenberg-Marquardt iterations!
Scene planes and homographies

The 2D point $x_{pi}$ is in the plane $pi$.

Plane induces homography between two views.

$x' = H_{2pi} H^{-1}_{1pi} x = H x$
Homography given plane

nonhomogeneous coordinates

\[ P = [I \mid 0] \quad P' = [A \mid a] \]

\[ \pi^T X = 0 \quad \pi = (v^T, 1)^T \]

point on plane

\[ x = PX = [I \mid 0]X \]

\[ X = (x^T, \rho)^T = (x^T, -v^T x)^T \]

project in second view

\[ x' = (A - av^T)x \]

\[ x' = Hx \quad H = A - av^T \]

v (and H) found from 3 linear correspondences

does NOT pass through the first camera center \((0^T, 1)^T\)

three-parameter of \(v\) inhomogeneous coor. (see \(\pi\) above)
Calibrated stereo rig

\[ P_E = K[I | 0] \quad P'_E = K'[R | t] \]

\[ \pi = \left(n^T d\right)^T \]  
world coordinate of the plane

\[ v = n/d \]

\[ H = K'\left(R - tn^T/d\right)K^{-1} \]  
three-parameter parametrized n/d

\[ n^T X + d = 0 \]  
the plane

homography induces by this plane

\[ H = H_{2pi} \quad H^{-1}_{-1pi} = K' (R - t v^T) K^{-1}_{-1pi} \]

\[ H_{2pi} \quad H^{-1}_{-1pi} \]
homographies and epipolar geometry

points on plane also have to satisfy epipolar geometry!

\[(Hx)^T Fx = x^T H^T Fx = 0, \forall x\]

\[H^T F + F^T H = 0\]

\[x^T H^T e' \times x' = 0, \forall x \leftrightarrow x'\]

\[Fx = e' \times x', \forall x\]

but 4 arbitrary points in BOTH images may satisfy the epipolar constraint only is

\[P = [I | 0] \quad P' = [A | e']\]

\[F = e' \times A\]

\[H = A - e' v^T\]

\[F^T H = A^T e' e [x] A - A^T (e' [x] e') v^T\]

then \(F^T H\) is skew-symmetric since \((e' [x] e') = 0\) the same epipolar point

the transformation \(H\) is the homography between two images induces by some world plane if and only if \(F = e' [x] H\)
Homography also maps epipole

$e' = He$

[point transformation]
Homography also maps epipolar lines

\[ l_e = H^T l'_e \]

inverse line transformation
 Compatibility constraint

\[ l'_e = Fx = x' \times (Hx) \]
plane homography
given F and 3 points correspondences

Method 1: reconstruct explicitly,
compute plane through 3 points
derive homography

Method 2: use epipoles as 4th correspondence
to compute homography

\[ x'_i = H x_i \quad i=1,2,3 \]
\[ e' = H e' \implies \text{find } H \]
degenerate geometry for an implicit computation of the homography

H cannot be computed uniquely X_1, X_2, baseline are same line in pi but can be solved with explicit (1) method
Estimation from 3 noisy points (+F)

Consistency constraint:
points have to be in exact epipolar correspondence

Determine MLE points given F and $x \leftrightarrow x'$

$$H = A - e'(M^{-1}b)^T$$

$$A = \begin{bmatrix} e' \\ x_i \end{bmatrix} F$$

$$b_i = (x'_i A x_i)^T (x'_i x e') / \|x'_i x e'\|^2$$

M v = b
M has the rows the points $x_i$
$v = M^{-1} b$
i = 1, 2, 3

3 image correspondences not collinear all three!
plane homography
given $F$, a point and a line

$x \leftrightarrow x'$

$P = [I \mid 0] \quad P' = [A \mid e'] \quad p_i(mu) = mu P^T l + P'^T l'$

$(4 \times 3 \text{ matrices}) \quad v(mu) = [v_1 \quad v_2 \quad v_3]^T$

$= (mu l + A^T l') / (e'^T l')$

$A = [e'] \times F \quad H = A - e'v^T = ((e'^T l')_3 \times e' l'^T) [e'] \times F - e' l'^T mu / (e'^T l')$

$= -([l'] \times [e'] \times [e'] \times F + mu e' l'^T) / (e'^T l')$

$= -([l'] \times F + mu e' l'^T) / (e'^T l') \uparrow \text{ up to a scale}$

$l'^T e' \text{ not equal to zero (line point l)}$

$H = [l']_x F + \mu e'l'^T$

$\mu = \frac{(x' \times e')^T (x' \times F x \times l')}{{\|x' \times e'\|^2 (l^T x)}}$

the point correspondence uniquely determines the plane
Degenerate homographies

\( l' \) in NOT the epipolar line!

words plane of H through the second camera center

\( x' = l' \times Fx \)

\[ H = [l'] \times F \text{ rank 2 matrix} \]

\( l' \times Fx = 0 \)
for any \( l' \) !!

rank-1 homography

both centers

rank-2 homography
plane induced parallax

homography induced virtual parallax

left  right images

points off the plane do not coincide
LINES joining corresponding point OFF the plane in the superimposed image intersect at the epipole.

\[ \mathbf{I}' = \mathbf{x}' \times \mathbf{H}_x \text{ two points give a line} \]
6-point algorithm

\( x_1, x_2, x_3, x_4 \) in plane, \( x_5, x_6 \) out of plane

Compute \( H \) from \( x_1, x_2, x_3, x_4 \)

\( e' = (x_5' \times Hx_5) \times (x_6' \times Hx_6) \)

\( F = [e']_x H \)

![Diagram a](image1)

![Diagram b](image2)
Projective depth

\[ X = (x^T, \rho)^T \]

\[ x' = Hx + \rho e' \]

\( \rho = 0 \) on plane

sign of \( \rho \) determines on which side of plane

\( x', Hx, e' \) are collinear

see 2HZfigure 13.10
Two planes

the two homographies induce off-plane information about the other, is enough for estimating $F$

$F = \left[ e_\mathbf{x} \right]_x H_i$

$H$ has fixed point and fixed line

$H = H_2^{-1} H_1$

$He = e$

$e' = H_i e \ i = 1,2$

$F$ is over-determined, the two homographies must satisfy consistency constraints

first image into itself

planar homology 5 dof
two eigenvalues equal

plus 3 dof of $\nu$...
on this segment correspondence