STRUCTURE COMPUTATION FROM TWO IMAGES
Scene planes and homographies

Fig. 13.1. The homography induced by a plane. The ray corresponding to a point $x$ is extended to meet the plane $\pi$ in a point $x_\pi$; this point is projected to a point $x'$ in the other image. The map from $x$ to $x'$ is the homography induced by the plane $\pi$. There is a perspectivity, $x = H_{1\pi}x_\pi$, between the world plane $\pi$ and the first image plane; and a perspectivity, $x' = H_{2\pi}x_\pi$, between the world plane and second image plane. The composition of the two perspectivities is a homography, $x' = H_{2\pi}H_{1\pi}^{-1}x = Hx$, between the image planes.

A plane induces homography between two views.
Homography given a plane

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

\[ \pi^T X = 0 \quad \pi = \begin{pmatrix} v^T, 1 \end{pmatrix}^T \]

Point on first image plane

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

Any point of the ray
\[ \pi^T X = 0 \]

Project in second view
\[ x' = (A - av^T)x \]

Compare with 4x4 matrices for ambiguity of \( F \)
\[ x' = P'H^{-1}HX \]

The sign of \( v \) changed for the \( H \). (not important)
Example: Calibrated stereo rig

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

\[ \pi = (n^T d)^T \quad \text{The 3D plane is} \quad n^T \tilde{X} + d = 0 \]

\[ v = n/d \]

Homography this case

\[ H_0 = R - t n^T / d \]

calibrated cameras

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

since \( x' = Hx = K' H_0 \tilde{X} = K' H_0 K^{\{-1\}} x \)

\( \tilde{X} \) is the inhomogeneous 3D point
Homographies and epipolar geometry

points on a plane also have to satisfy epipolar geometry!

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

\[H^T F \text{ has to be skew-symmetric; } H \text{ and } F \text{ be compatible}\]

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

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

I' epipolar line

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

\[F = e' \times H\]

6 - 1 = 5 DOF taken,
3 DOF family of planes for the homography
epipolar must obey

will show
A camera pair $P = [I | 0]$ $P' = [A | a]$ induces a homography $H = A - a v^T$ on a 3D plane having the normal $v$.

If choose $F = e' \times A$ then $P' = [A | e']$

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

the matrix $F^T H$ is skew-symmetric.
Equivalent with $P' = [A + e' v^T | ke']$ up the scale and sign. Then $F = e' \times H$.

If the normal $v = (0, 0, 0, 1)^T \implies P = [I | 0]$ $P' = [H | e']$. in world coordinate frame the plane at infinity $H = H_{\text{inf}}$
Homography maps epipole...

Compatibility constraints

intersects the baseline

\[ e' = He \]
...homography maps epipolar lines...

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

epipoles mapping also true
maps any point on a homography line to its corresponding epipolar line.

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

Method 1: compute plane through three 3D points ==> v
from F derive e' and find A  F=e'xA
(explicit reconstruction)

Method 2: use epipoles as 4th correspondence
to compute homography,  x' = Hx  i = 1,...,4
(implicit reconstruction)

should not be used if degeneracies are not present in method 1
Degenerate geometry for the second method.

Both $X_1$ and $X_2$ are in the same epipolar plane. The $H$ is not uniquely determined, collinear $E$.

**Solution**

\[
F = [e']_\times A \\
H = A - e'v^T \\
x_i' \times Hx_i = 0 \quad i = 1, 2, 3
\]

choose $A = [e']_\times F$

\[
x_i' \times Ax_i = (x_i' \times e')v^T x_i
\]

\[
x_i^T v = \frac{(x_i' \times Ax_i)^T (x_i' \times e')}{||x_i' \times e'||^2} = b_i \\
Mv = b \quad M = [x_1, x_2, x_3]^T
\]

The matrix $M$ has to be nonsingular $\det M \neq 0$. Here it is.

Consistency constraint: $e' \times x_i'$ is on an epipolar line.

Estimate first the 3D points $X_i$ and $x_i' = Hx_i$ to satisfy it.

triangulation because noise points
Plane homography: given a point and a line correspondence

First find $H$ function of one parameter family.

Fig. 13.4. (a) Image lines $l$ and $l'$ determine planes $\pi$ and $\pi'$ respectively. The intersection of these planes defines the line $\bar{L}$ in 3-space. (b) The line $L$ in 3-space is contained in a one parameter family of planes $\pi(\mu)$. This family of planes induces a one parameter family of homographies between the images.
The lines \( l \) and \( l' \) back-project to the planes \( P^\top l \) and \( P'^\top l' \). Having \( P = [I \ | \ 0] \) and \( P = [A \ | \ e'] \) and pencil of plane in 3D is given by \((\text{normal vector})^\top \cdot (3\text{D points}) = 0\)

\[
\pi(\mu) = \mu P^\top l + P'^\top l' = \mu \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} A'^\top l' \\ e'^\top l' \end{pmatrix} = (v^\top, 1)^\top.
\]

Using \( A = [e']_x F \) and \( H = A - e'v(\mu)^\top \)

\[
H = [e']_x F - e' \left( \frac{(e'^\top l' I - e'l'^\top) [e']_x F - \mu e' l'^\top}{e'^\top l'} \right)^\top
\]

Can verify that \( e'^\top l' I_{3 \times 3} - e'l'^\top = -[l']_x [e']_x \) and following is true

\[
[e']_x [e']_x F = [e']_x [e']_x [e']_x A \sim -[e']_x A = F
\]

\[
H = -\frac{[l']_x F + \mu e' l'^\top}{e'^\top l'}
\]

where the minus and the scale can be discarded. The projective variable \( \mu \) is found from \( x' \times H x = 0 \)

\[
x' \times ([l']_x F + \mu e' l'^\top) x = 0
\]

\[
\mu (l^\top x) (x' \times e') = -x' \times [l']_x F x
\]

since \([l']_x (F x) = -(F x) \times l' \)

\[
\mu = \frac{(x' \times e')^\top (x' \times (F x) \times l')}{{|x' \times e'|}^2 (l^\top x)}
\]

If the point pair are on the lines \( l, l' \), then \( F x = l'_c = l' \) and \( x' = [l']_x F x \). There is point-to-point map for this point pairs.
Degenerate homographies

\[ l' \text{ is not an epipolar line} \]

\[ x' = l' \times Fx \]

\[ \text{H rank-2} \]

All \( \mathbf{x} \) maps to \( l' \)

\[ \pi' \]

\[ \mathbf{C} \rightarrow x \rightarrow \pi' \rightarrow x' \rightarrow \mathbf{C'} \]

rank-2

\[ \mathbf{C} \rightarrow x \rightarrow \pi' \rightarrow x' \rightarrow \mathbf{C'} \]

\[ \mathbf{e} \rightarrow x' \rightarrow \mathbf{C'} \]

\[ \mathbf{p_i} = \mathbf{p_i'} \text{ no } \mathbf{L} \]

rank-1
Computing $F$: given homography induced by a plane

$x \mapsto \tilde{x}'$ on epipolar line

two point pairs $\Rightarrow e'$

$F = e' \times H$

Fig. 13.7. Plane induced parallax. The ray through $X$ intersects the plane $\pi$ at the point $X_\pi$. The images of $X$ and $X_\pi$ are coincident points at $x$ in the first view. In the second view the images are the points $x'$ and $\tilde{x}' = Hx$ respectively. These points are not coincident (unless $X$ is on $\pi$), but both are on the epipolar line $I_x$ of $x$. The vector between the points $x'$ and $\tilde{x}'$ is the parallax relative to the homography induced by the plane $\pi$. Note that if $X$ is on the other side of the plane, then $\tilde{x}'$ will be on the other side of $x'$.

left  right  left on right
chinese text is the homography
Example: 6-point algorithm F. Virtual parallax.

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

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

\[
e' = (x'_5 \times H x_5) \times (x'_6 \times H x_6)
\]

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

F: 6 instead of 7DOF
Projective depth

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

\[ \pi \]

\[ P = [I \mid 0] \]
\[ P' = [H \mid e'] \]

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

\( \rho = 0 \) on plane

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

\( x', e', Hx = \tilde{x}' \) are collinear
Fig. 13.10. **Binary space partition.** (a) (b) Left and right images. (c) Points whose correspondence is known. (d) A triplet of points selected from (c). This triplet defines a plane. The points in (c) can then be classified according to their side of the plane. (e) Points on one side. (f) Points on the other side.

Several planes identifies 3-space region.
Two planes: two homographies
F is overdetermined

\[ H = H_2^{-1} H_1 \quad \text{He} = e \quad F = [e'_i] \times H_i \]

H has fixed point e. The fixed points on intersection line, map to themselves.
Till now projective transformation, now also an affine element.

The infinite homography, $H_\infty$, is the homography induced by the plane at infinity, $\pi_\infty$.

$$x' = K'RK^{-1}x + K't/Z = H_\infty x + K't/Z$$

depth measured from first camera $Z = \infty$.

$H_{\text{inf}}$ depends only on rotation and internal parameters.

$t = 0 \implies$ only rotation around the center.

$e' = K't$ so $1/Z$ is the parallax $\rho$ from the plane at infinity.

All points are vanishing points on plane at infinity $v' = H_{\text{inf}} v$.

Three (orthogonal) vanishing points gives $H_{\text{inf}}$ and $F$.

[Result 13.6] or vanishing line+vanishing point [Sec.13.2.2]
If the cameras are $P = [I \mid 0]$ and $P' = [H_{\text{inf}} \mid e']$ the reconstruction is \textit{affine} already.

If $p_{\text{inf}} = (0 \ 0 \ 0 \ 1)^T$ $H_{\text{inf}}$ obtained from the camera matrices
For a points $X = (x_{\text{inf}}^T \ 0)^T$ $x = M \ x_{\text{inf}}$ $x' = M' \ x_{\text{inf}}$
where $M$ and $M'$ a 3x3 submatrix of the cameras.

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

The image of the absolute conic specified in the first image can be mapped to the second image.

\[ \omega = (KK^T)^{-1} \]
\[ \omega' = H_{\infty}^{-T} \omega H_{\infty}^{-1} \]
\[ \omega' = (K'K'^T)^{-1} \]
given the conic transfer obtained with the infinite homography find \textbf{omega} and \textit{metric} reconstruction.
For stereo correspondence...

if $H_{\text{inf}}$ given [but needs oriented projective geometry].

---

Fig. 13.13. **Reducing the search region using $H_{\infty}$.** Points in 3-space are no ‘further’ away than $\pi_{\infty}$. $H_{\infty}$ captures this constraint and limits the search on the epipolar line in one direction. The baseline between the cameras partitions each epipolar plane into two. A point on one “side” of the epipolar line in the left image will be imaged on the corresponding “side” of the epipolar line in the right image (indicated by the solid line in the figure). The epipole thus bounds the search region in the other direction.
not covered: Affine epipolar geometry

The solution is a total least squares problem.

Fig. 14.1. **Affine epipolar geometry.** (a) Correspondence geometry: Projection rays are parallel and intersect at infinity. A point $x$ back-projects to a ray in 3-space defined by the first camera centre (at infinity) and $\hat{x}$. This ray is imaged as a line $\mathcal{Y}$ in the second view. The 3-space point $X$ which projects to $x$ lies on this ray, so the image of $X$ in the second view lies on $\mathcal{Y}$. (b) Epipolar lines and planes are parallel.
not covered: Trifocal tensor

Three epipolar constraints are not independent. Linear equations in tensor notation.
3D reconstruction of cameras and structure

reconstruction problem

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

$$x_i = PX_i, \quad x'_i = PX'_i$$

for all $i$

without additional information is only possible
up to projective ambiguity

*Only inliers in reconstruction.*
Reconstruction ambiguity: projective

Fundamental matrix. The angle between points changes (needs \textit{omega}!) but the 2D points remain the same, including the epipoles.
Reconstruction ambiguity: similarity

Essential matrix.

Angles remain the same.

2D points are the same.

\[ x_i = PX_i = (PH_S^{-1})(H_SX_i) \]

\[ H_S = \begin{bmatrix} R & t \\ 0^T & \lambda \end{bmatrix} \]

\[ P = K[R_P \mid t_P] \]

\[ PH_S^{-1} = K[R_PR^{-1} \mid t'] \]
Terminology

$x_i \leftrightarrow x'_i$

In 3D the original scene has $X_i$

*Projective, affine, similarity reconstruction*

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

Literature: Metric and Euclidean reconstruction

= similarity reconstruction

**Stratified reconstruction**

(i) Projective reconstruction
(ii) Affine reconstruction
(iii) Metric reconstruction
projective reconstruction

(a) Compute F from correspondences
(b) Compute camera matrices from F
(c) Compute 3D point for each pair of corresponding points

**computation of F**
use $x'^T F x_i = 0$ equations, linear in coeff. F
8 points (linear), 7 points (nonlinear), LM (nonlinear)

**computation of camera matrices**
use $P = [I \mid 0]$  $P' = [[e']_x F] \lambda e'$

**triangulation**
compute intersection of two backprojected rays
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.

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

\[ P_2 = P_1 H^{-1} \quad P'_2 = P'_1 H^{-1} \quad X_2 = H X_1 \quad \text{(except: } F x_i = x_i^T F = 0) \]

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

⇒ along same ray of \( P_2 \) \quad \text{same for } P'_2

\[ H \text{ any } 4 \times 4 \quad \text{nonsingular} \]

two possibilities: \( X_{2i} = H X_{1i} \), \( \text{or distinct points along baseline coincide with the epipoles} \)

key result: allows reconstruction from pair of \textbf{uncalibrated images}
Fig. 10.3. **Projective reconstruction.** (a) Original image pair. (b) 2 views of a 3D projective reconstruction of the scene. The reconstruction requires no information about the camera matrices, or information about the scene geometry. The fundamental matrix $F$ is computed from point correspondences between the images, camera matrices are retrieved from $F$, and then 3D points are computed by triangulation from the correspondences. The lines of the wireframe link the computed 3D points.
Projective to affine reconstruction

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

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

\[ \pi = (A, B, C, D)^T \mapsto (0, 0, 0, 1)^T \]

plane at infinity

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

3D plane transformation

\[ H = \begin{bmatrix} I & 0 \\ \pi^T & \pi \end{bmatrix} \]

if \( P_{14} \neq 0 \)

otherwise, Householder matrix directly from the 3D transformation A4.1.2

a projective transformation which has fixed the plane at infinity, \( \pi \), becomes affine transformation

can be sufficient depending on some application e.g. mid-point, centroid, parallelism
Scene constraints

*Translational motion only*. Need two views. Points in plane at infinity does not move, \( x_i = x'_i \). From three such point pairs find \( X_i \) and recover the plane \( \pi_{\text{inf}} \). Matrix \( H \) determined.

*Parallel lines*. Intersect at infinity. Three sets of such parallel lines allow to uniquely determine \( \pi_{\text{inf}} \). ...or in 2D only
The \( v \) is a vanishing points in the first image ==> epipolar line \( l' \) in the second image. Gives two equations in 3D point \( X\):
first image \( v \times PX = 0 \)  
second image \( l'^{\top} P'X = 0 \)

Infinite homography. After affine reconstruction \( P = [M \mid m] \) \( P' = [M' \mid m'] \) and \( X^{\top} = (X, 0) \). Thus \( H_{\text{inf}} = M'M^{\{-1\}} \)
Fig. 10.4. **Affine reconstruction.** The projective reconstruction of figure 10.3 may be upgraded to affine using parallel scene lines. (a) There are 3 sets of parallel lines in the scene, each set with a different direction. These 3 sets enable the position of the plane at infinity, $\pi_\infty$, to be computed in the projective reconstruction. The wireframe projective reconstruction of figure 10.3 is then affinely rectified using the homography (10.2). (b) Shows two orthographic views of the wireframe affine reconstruction. Note that parallel scene lines are parallel in the reconstruction, but lines that are perpendicular in the scene are not perpendicular in the reconstruction.
Affine to metric reconstruction

Identify absolute conic on the plane of infinity and back-project to \( \omega \) the image of the absolute conic.

This image is independent of particular reconstruction.
To move from affine to similarity a 3D transformation

\[ P = [M \mid m] \text{ is a conic } \omega \]

after affine reconstruction

\[ \omega^* = \omega^{-1} = KK^T. \text{ Combining this with } M_M = MA = KR \]

How to obtain the image of the absolute conic? One camera!

Scene orthogonality: vanishing points/lines

\[ v_1^T \omega v_2 = 0 \quad l = \omega v \]

Known internal parameters

\[ \omega_{12} = \omega_{21} = 0 \]

zero skew and \( \alpha_x = \alpha_y \)

\[ \omega_{11} = \omega_{22} \]

If two images, not same internal para.
Fig. 10.5. **Metric reconstruction.** The affine reconstruction of figure 10.4 is upgraded to metric by computing the image of the absolute conic. The information used is the orthogonality of the directions of the parallel line sets shown in figure 10.4, together with the constraint that both images have square pixels. The square pixel constraint is transferred from one image to the other using $H_\infty$. (a) Two views of the metric reconstruction. Lines which are perpendicular in the scene are perpendicular in the reconstruction and also the aspect ratio of the sides of the house is veridical. (b) Two views of a texture mapped piecewise planar model built from the wireframes.
two views

Direct metric reconstruction using $\omega$

\forall

assume the same internal parameters

**approach 1**

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

calibrated reconstruction, essential matrix, 4 solutions

**approach 2**

compute projective reconstruction, $P$ and $P'$

back-project $\omega$ from *both* images

defines two $\Omega_\infty$ and plane at infinity $\pi_\infty$

two solutions of absolute conics.
Twisted pair ambiguity in reconstruction.
Objective
Given two uncalibrated images compute a metric reconstruction \((P_M, P'_M, \{X_{M_i}\})\) of the cameras and scene structure, i.e. a reconstruction that is within a similarity transformation of the true cameras and scene structure.

Algorithm

(i) **Compute a projective reconstruction** \((P, P', \{X_i\})\):

(a) **Compute the fundamental matrix** from point correspondences \(x_i \leftrightarrow x'_i\) between the images.

(b) **Camera retrieval**: compute the camera matrices \(P, P'\) from the fundamental matrix.

(c) **Triangulation**: for each point correspondence \(x_i \leftrightarrow x'_i\), compute the point \(X_i\) in space that projects to these two image points.

(ii) **Rectify the projective reconstruction to metric**:

- **either Direct method**: Compute the homography \(H\) such that \(X_{e;i} = HX_i\) from five or more ground control points \(X_{e;i}\) with known Euclidean positions. Then the metric reconstruction is

\[
P_M = PH^{-1}, \quad P'_M = P'H^{-1}, \quad X_{M_i} = X_i.
\]

- **or Stratified method**:

(a) **Affine reconstruction**: Compute the plane at infinity, \(\pi_\infty\), as described in section 10.4.1, and then upgrade the projective reconstruction to an affine reconstruction with the homography

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

(b) **Metric reconstruction**: Compute the image of the absolute conic, \(\omega\), as described in section 10.4.2, and then upgrade the affine reconstruction to a metric reconstruction with the homography

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

where \(A\) is obtained by Cholesky factorization from the equation \(AA^T = (M^T \omega M)^{-1}\), and \(M\) is the first \(3 \times 3\) submatrix of the camera in the affine reconstruction for which \(\omega\) is computed.

Algorithm 10.1. *Computation of a metric reconstruction from two uncalibrated images.*
<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>$\pi_\infty$</td>
<td>projective</td>
</tr>
<tr>
<td>point correspondences</td>
<td>$F, H$</td>
<td>$\pi_\infty$</td>
<td>affine</td>
</tr>
<tr>
<td>including vanishing points</td>
<td></td>
<td>$\pi_\infty$</td>
<td>metric</td>
</tr>
<tr>
<td>Points correspondences</td>
<td>$F, H$</td>
<td>$\omega, \omega'$</td>
<td>metric</td>
</tr>
<tr>
<td>and internal camera</td>
<td></td>
<td>$\Omega_\infty$</td>
<td></td>
</tr>
<tr>
<td>calibration</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Direct reconstruction using ground truth

use control points $X_{Ei}$ with known Euclidean coordinates to go from projective to metric, $X_i$ projective coor.

Find $4 \times 4 \, H$.

From $x \leftrightarrow x'$ obtain $X$.

3 equation $H$ per $X$

$n \geq 5$ points in 3D

$X_{Ei} = HX_i$

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

image plane 2/point

If $x_i$ and $x'_i$ are visible for $X_{Ei}$, coplanarity three eqn. only.