SINGLE VIEW GEOMETRY AND SOME APPLICATIONS
Thank you for the slides. They come mostly from the following sources.

Marc Pollefeys  U. of North Carolina
Daniel DeMenthon  U. of Maryland
Alexei Efros  CMU
The most general transformation that can occur between a scene plane and an image plane under perspective imaging is a plane projective transformation (affine camera-affine transformation)
Action of projective camera on lines

forward projection

\[ X(\mu) = P(A + \mu B) = PA + \mu PB = a + \mu b \]

back-projection

\[ \Pi = P^T l \]

\[ \Pi^T X = l^T PX = 0 \]
The importance of the camera center

\[ P = KR[I \mid -\tilde{C}], \quad P' = K'R'[I \mid -\tilde{C}] \]

\[ P' = K'R'(KR)^{-1}P \]

\[ x' = P'X = K'R'(KR)^{-1}PX = K'R'(KR)^{-1}x \]

\[ x' = Hx \text{ with } H = K'R'(KR)^{-1} \]
Moving the image plane (zooming)

\[ x = K[I \mid 0]X \]
\[ x' = K'[I \mid 0]X = K'(K)^{-1}x \]

**first camera**

\[ H = K'(K)^{-1} = \begin{bmatrix} kI & (1-k)\tilde{x}_0 \\ 0^T & 1 \end{bmatrix} \]

**k = (x' - x_{\_0}) / (x_{\_} - x_{\_0})**

\[ x' = k x_{\_} + (1-k) x_{\_0} \]

**radiating from the principal point in the image -- \( x_{\_} \) and \( x_{\_0} \) on a line**

\[ k = f'/f \]

\[ K' = \begin{bmatrix} kI & (1-k)\tilde{x}_0 \\ 0^T & 1 \end{bmatrix} \]
\[ = \begin{bmatrix} kA & \tilde{x}_0 \\ 0^T & 1 \end{bmatrix} = K \begin{bmatrix} kI & 0 \\ 0^T & 1 \end{bmatrix} \]

**increase of focal length = magnification**

a compound lens will perturb also the camera center and the principal point

**inhomogeneous principal point**

**effect of zooming with factor \( k \) is to multiply \( K \) on the right with diag \((k, k, 1)\)**
Camera rotation

\[ x = K[I \mid 0]X \]
\[ x' = K[R \mid 0]X = KRK^{-1}x \]

\[ H = KRK^{-1} \]

same eigenvalues as \( R \)
but is a projective homography

\[ \{\mu, \mu e^{i\theta}, \mu e^{-i\theta}\} \]

\( \text{the angle theta from the phase} \)
\( \text{mu eigenvalue of} \ H \quad v = K \mu \text{ is the} \)
\( \text{vanishing point of the rotation axis in the image} \)

\( \text{mu unknown scale} \)
\( \det H = 1 \quad \mu = 1 \)

only rotation almost parallel with y-axis

translation too
Moving the camera center

motion parallax

the vector between $x'_{1}$ and $x'_{2}$

epipolar line

the center $C$ is moved
Synthetic view

(i) Compute the homography that warps some a rectangle to the correct aspect ratio
(ii) warp the image
Planar homography mosaicing

- Compute homography from each image to the reference.
- One image is the reference.
- Augment the reference image with the non-overlapping part of the warped image.
- Repeat...
Projective (reduced) notation

Let 

\[ X_1 = (1,0,0,0)^T, \quad X_2 = (0,1,0,0)^T, \quad X_3 = (0,0,1,0)^T, \quad X_4 = (0,0,0,1)^T \]

\[ x_1 = (1,0,0)^T, \quad x_2 = (0,1,0)^T, \quad x_3 = (0,0,1)^T, \quad x_4 = (1,1,1)^T \]

\[ P = \begin{bmatrix} a & 0 & 0 & -d \\ 0 & b & 0 & -d \\ 0 & 0 & c & -d \end{bmatrix} \]

give the camera matrix \[ x_i = P \cdot X_i \] \[ i = 1, \ldots, 4 \]

\[ a, b, c, d \] arbitrary constants

three degrees of freedom in \( P \) here but you can multiply it with a homography and \( P \) will be projective equivalent. e.g. \( 3 + 8 = 11 \) dof
Object Pose with 1D Image Plane

• What happens if we don’t know object’s angle?
More Points

• Limited number of object poses (2 or 1)
  ▪ Head lights and one taillight

Transparent car
Correspondence Problem

- When we know correspondences (i.e. matchings), pose is easier to find
- When we know the pose, correspondences are easier to find.
- But we need to find both at the same time
- Below, we first assume we know correspondences and describe how to solve the pose given \( n \) corresponding points in image and object
  - Perspective \( n \)-Point Problem
- Then we explore what to do when we don’t know correspondences
Figure 1: Perspective projection ($m_i$) and scaled orthographic projection ($p_i$) for an object point $M_i$ and a reference point $M_0$.

The points $m_i$ known camera coordinates

Internal camera calibrated $K$ is known

$M_0$ gives the center plane parallel with the image plane

four points is enough for pose since a $K$ is known already

We want to find $r_1, r_2, r_3, T_x, T_y, T_z$ in the camera coordinate system by iterations. $x_j = u_j/w_j, y_j = v_j/w_j$ has 6 d.o.f - an affine transformation
Iterative Pose Calculation

\[
\begin{bmatrix}
    u \\
v \\
w
\end{bmatrix}
= \begin{bmatrix}
    r_1^T & T_x \\
r_2^T & T_y \\
r_3^T & T_z
\end{bmatrix}
\begin{bmatrix}
    X_s \\
Y_s \\
Z_s \\
1
\end{bmatrix}
\Rightarrow
\begin{bmatrix}
    u \\
v \\
w
\end{bmatrix}
= \begin{bmatrix}
    r_1^T / T_z & T_x / T_z \\
r_2^T / T_z & T_y / T_z \\
r_3^T / T_z & 1
\end{bmatrix}
\begin{bmatrix}
    X \\
Y \\
Z \\
1
\end{bmatrix}
\]

\[
\begin{bmatrix}
    u_1 & v_1 \\
u_2 & v_2 \\
u_3 & v_3 \\
u_4 & v_4
\end{bmatrix}
= \begin{bmatrix}
    X_1 & Y_1 & Z_1 & 1 \\
X_2 & Y_2 & Z_2 & 1 \\
X_3 & Y_3 & Z_3 & 1 \\
X_4 & Y_4 & Z_4 & 1
\end{bmatrix}
\begin{bmatrix}
    r_1 / T_z & r_2 / T_z \\
T_x / T_z & T_y / T_z
\end{bmatrix}
\Rightarrow
\begin{bmatrix}
    u_1 & v_1 \\
u_2 & v_2 \\
u_3 & v_3 \\
u_4 & v_4
\end{bmatrix}
= M^{-1}
\begin{bmatrix}
    u_1 & v_1 \\
u_2 & v_2 \\
u_3 & v_3 \\
u_4 & v_4
\end{bmatrix}
\]

Non coplanar points needed (otherwise matrix \( M \) is singular). At least 4 points.

\[w_i = 1 + r_3 \cdot (X_i, Y_i, Z_i) / T_z\]
Iterative Pose Calculation

• Compute model matrix $M$ and its inverse

• Assume $\mathbf{r}_3 \cdot (X_i, Y_i, Z_i) / T_z = 0 \Rightarrow w_i = 1$

• Compute $u_i = w_i x_i$, $v_i = w_i y_i$

• Compute

\[
\begin{bmatrix}
\mathbf{r}_1 / T_z & \mathbf{r}_2 / T_z \\
T_x / T_z & T_y / T_z
\end{bmatrix}
= M^{-1}
\begin{bmatrix}
u_1 & v_1 \\
u_2 & v_2 \\
u_3 & v_3 \\
u_4 & v_4
\end{bmatrix}
\]

• Compute $T_z$, $T_x$, $T_y$, $\mathbf{r}_1$, $\mathbf{r}_2$, then $\mathbf{r}_3 = \mathbf{r}_1 \times \mathbf{r}_2$

• Compute $w_i = 1 + r_3 \cdot (X_i, Y_i, Z_i) / T_z$

• Go back to step 2 and iterate until convergence
Iterative Pose Calculation

1. Find object pose under scaled orthographic projection
2. Project object points on lines of sight
3. Find scaled orthographic projection images of those points
4. Loop using those images in step 1
Left: Actual perspective image for cube with known model
Top: Evolution of perspective image during iteration
Bottom: Evolution of scaled orthographic projection
4-3-2-?

- 4 - point perspective solution
  - Unique solution for 6 pose parameters
  - Computational complexity of $n^4m^4$
- 3 - point perspective solution
  - Generally two solutions per triangle pair, but sometimes four.
  - Reduced complexity of $n^3m^3$
3 Points

- Each correspondence between scene point and image point determines 2 equations
- Since there are 6 degrees of freedom in the pose problems, the correspondences between 3 scene points in a known configuration and 3 image points should provide enough equations for computing the pose of the 3 scene points
- The pose of a triangle of known dimension is defined from a single image of the triangle
  - But nonlinear method, 2 to 4 solutions
Triangle Pose Problem

• There are two basic approaches
  ▪ Analytically solving for unknown pose parameters
    • Solving a 4th degree equation in one pose parameter, and then using the 4 solutions to the equation to solve for remaining pose parameters
    • problem: errors in estimating location of image features can lead to either large pose errors or failure to solve the 4th degree equation
  ▪ Approximate numerical algorithms
    • find solutions when exact methods fail due to image measurement error
    • more computation
Numerical Method for Triangle Pose

- If distance $R_c$ to C is known, then possible locations of A (and B) can be computed
  - they lie on the intersections of the line of sight through A' and the sphere of radius AC centered at C
  - Once A and B are located, their distance can be computed and compared against the actual distance $AB$
Numerical Method for Triangle Pose

• Not practical to search on $R_c$ since it is unbounded
  Instead, search on one angular pose parameter, $\alpha$.
  - $R_c = \frac{AC \cos \alpha}{\sin \delta}$
  - $R_a = R_c \cos \delta \pm AC \sin \alpha$
  - $R_b = R_c \cos \gamma \pm \text{BC sin beta}$
• This results in four possible lengths for side AB
• Keep poses with the right AB length
Choosing Points on Objects

• Given a 3-D object, how do we decide which points from its surface to choose for its model?
  ▪ Choose points that will give rise to detectable features in images
  ▪ For polyhedra, the images of its vertices will be points in the images where two or more long lines meet
    • These can be detected by edge detection methods
  ▪ Points on the interiors of regions, or along straight lines are not easily identified in images.
Choosing the Points

• Example: why not choose the midpoints of the edges of a polyhedra as features
  ▪ midpoints of projections of line segments are not the projections of the midpoints of line segments
  ▪ if the entire line segment in the image is not identified, then we introduce error in locating midpoint
Reducing the Combinatorics of Pose Estimation

- Reducing the number of matches
  - Consider only quadruples of image features that
    - Are connected by edges
    - Are “close” to one another
      - But not too close or the inevitable errors in estimating the position of an image vertex will lead to large errors in pose estimation
  - Generally, try to group the image features into sets that are probably from a single object, and then only construct quadruples from within a single group

(easier to distinguish separate object)
RANSAC

• **RANdom SAmple Consensus**

• Randomly select a set of 3 points in the image and a select a set of 3 points in the model

• Compute triangle pose and pose of model

• Project model at computed pose onto image

• Determine the set of projected model points that are within a distance threshold $t$ of image points, called the *consensus set*

• After $N$ trials, select pose with largest consensus set
A7.2 Planar homologies

A planar homology is a plane projective transformation which has a line \( a \) of fixed points, called the axis, and a distinct fixed point \( v \), not on the line, called the centre or vertex of the homology. There is a pencil of fixed lines through the vertex. Algebraically, two of the eigenvalues of the transformation matrix are equal, and the fixed line corresponds to the 2D invariant space of the matrix (here the repeated eigenvalues are \( \lambda_2 \) and \( \lambda_3 \)).

Consider two images obtained by a camera rotating about its centre (as in figure 2.3-(p36)b); then as shown in section 8.4.2(p204), the images are related by a conjugate rotation. In this case the complex eigenvalues determine the angle \( \theta \) through which the camera rotates, and the eigenvector corresponding to the real eigenvalue is the vanishing point of the rotation axis. Note that \( \theta \) — a metric invariant — can be measured directly from the projective transformation.

Properties of a planar homology include:

- Lines joining corresponding points intersect at the vertex, and corresponding lines (i.e. lines through two pairs of corresponding points) intersect on the axis. This is an example of Desargues' Theorem. See figure A7.2a.
- The cross ratios defined by the vertex, a pair of corresponding points, and the intersection of the line joining these points with the line of fixed points, are the same for all points related by the homology. See figure A7.2b.
- For curves related by a planar homology, corresponding tangents (the limit of neighbouring points defining corresponding lines) intersect on the axis.
- The vertex (2 dof), axis (2 dof) and invariant (1 dof) are sufficient to define the homology completely. A planar homology thus has 5 degrees of freedom.
Appendix 7 Some Special Plane Projective Transformations

Fig. A7.2. Homology transformation. (a) Under the transformation points on the axis are mapped to themselves; each point off the axis lies on a fixed line through \( v \) intersecting \( a \) and is mapped to another point on the line. Consequently, corresponding point pairs \( x \leftrightarrow x' \) and the vertex of the homology are collinear. Corresponding lines – i.e. lines through pairs of corresponding points – intersect on the axis: for example, the lines \( \langle x_1, x_2 \rangle \) and \( \langle x'_1, x'_2 \rangle \). (b) The cross ratio defined by the vertex \( v \), the corresponding points \( x, x' \) and the intersection of their join with the axis \( i \), is the characteristic invariant of the homology, and is the same for all corresponding points. For example, the cross ratios of the four points \( \{v, x_1, x_1, i_1\} \) and \( \{v, x_2, x_2, i_2\} \) are equal since they are perspective related by lines concurrent at \( p_{12} \). It follows that the cross ratio is the same for all points related by the homology.

- 3 matched points are sufficient to compute a planar homology. The 6 degrees of freedom of the point matches over-constrain the 5 degrees of freedom of the homology.

A planar homology arises naturally in an image of two planes related by a perspectivity of 3-space (i.e. lines joining corresponding points on the two planes are concurrent). An example is the transformation between the image of a planar object and the image of its shadow on a plane. In this case the axis is the imaged intersection of the two planes, and the vertex is the image of the light source, see figure 2.5(p36)c.

**Parametrization.** The projective transformation representing the homology can be parametrized directly in terms of the 3-vectors representing the axis \( a \) and vertex \( v \), and the characteristic ratio \( \mu \), as

\[
H = I + \left( \mu - 1 \right) \frac{va^T}{v^Ta}
\]

where \( I \) is the identity. It is straightforward to verify that the inverse transformation is given by

\[
H^{-1} = I + \left( \frac{1}{\mu} - 1 \right) \frac{va^T}{v^Ta}.
\]

The eigenvectors are

\[
\{e_1 = v, e_2 = a_1^\perp, e_3 = a_2^\perp\}
\]
with corresponding eigenvalues
\[
\{\lambda_1 = \mu, \lambda_2 = 1, \lambda_3 = 1\}
\]
where \(a_i^\perp\) are two vectors that span the space orthogonal to the 3-vector \(a\), i.e. \(a^\top a_i^\perp = 0\) and \(a = a_1^\perp \times a_2^\perp\).

If the axis or the vertex is at infinity then the homology is an affine transformation. Algebraically, if \(a = (0,0,1)^\top\), then the axis is at infinity; or if \(v = (v_1,v_2,0)^\top\), then the vertex is at infinity; and in both cases the transformation matrix \(H\) has last row \((0,0,1)\).

Planar harmonic homology. A specialization of a planar homology is the case that the cross ratio is harmonic \((\mu = -1)\). This planar homology is called a planar harmonic homology and has 4 degrees of freedom since the invariant is known. The transformation matrix \(H\) obeys \(H^2 = I\), i.e. the transformation is a square root of the identity, which is called an involution (also a collineation of period 2). The eigenvalues are, up to a common scale factor, \\{\(-1, 1, 1\)\}. Two pairs of point correspondences determine \(H\).

In a perspective image of a plane object with a bilateral symmetry, corresponding points in the image are related by a planar harmonic homology. The axis of the homology is the image of the symmetry axis. Algebraically, \(H\) is a conjugate reflection where the conjugating element is a plane projective transformation. In an affine image (generated by an affine camera) the resulting transformation is a skewed symmetry, and the conjugating element is a plane affine transformation. For a skewed symmetry the vertex is at infinity, and the lines joining corresponding points are parallel.

The harmonic homology can be parametrized as
\[
H = H^{-1} = I - 2\frac{va^\top}{v^\top a}.
\]
Again, if the axis or vertex is at infinity then the transformation is affine.

A7.3 Elations

An elation has a line of fixed points (the axis), and a pencil of fixed lines intersecting in a point (the vertex) on the axis. It may be thought of as the limit of a homology where the vertex is on the line of fixed points. Algebraically, the matrix has three equal eigenvalues, but the eigenspace is 2-dimensional. It may be parametrized as
\[
H = I + \mu va^\top \text{ with } a^\top v = 0 \quad (A7.1)
\]
where \(a\) is the axis, and \(v\) the vertex. The eigenvalues are all unity. The invariant space of \(H\) is spanned by \(a_1^\perp, a_2^\perp\). This is a line (pencil) of fixed points (which includes \(v\) since \(a^\top v = 0\)). The invariant space of \(H^\top\) is spanned by vectors \(v_1^\perp, v_2^\perp\) orthogonal to \(v\). This is a pencil of fixed lines, \(l = \alpha v_1^\perp + \beta v_2^\perp\), for which \(l^\top v = 0\), i.e. all lines of the pencil are concurrent at the point \(v\).

An elation has 4 degrees of freedom: one less than a homology due to the constraint \(a^\top v = 0\). It is defined by the axis \(a\) (2 dof), the vertex \(v\) on \(a\) (1 dof) and the parameter \(\mu\) (1 dof). It can be determined from 2 point correspondences.
How can we model this scene?

1. Find world coordinates \((X,Y,Z)\) for a few points
2. Connect the points with planes to model geometry
   - Texture map the planes

Finding world coordinates \((X,Y,Z)\)

1. Define the ground plane \((Z=0)\)
2. Compute points \((X,Y,0)\) on that plane
3. Compute the *heights* \(Z\) of all other points
Perspective cues

Comparing heights

Vanishing Point
Measuring height

Intersect p₁q₁ with p₂q₂
\[ v = (p₁ \times q₁) \times (p₂ \times q₂) \]

Least squares version
- Better to use more than two lines and compute the "closest" point of intersection
- See notes by Bob Collins for one good way of doing this:
Vanishing point

Vanishing line

Vertical vanishing point (at infinity)

Vanishing point

Measuring height without a ruler

Compute $Z$ from image measurements

- Need more than vanishing points to do this
Measuring height

\[ v \approx (b \times b_0) \times (v_x \times v_y) \]

\[ t \equiv (v \times t_0) \times (r \times b) \]

What if the point on the ground plane \( b_0 \) is not known?

- Here the guy is standing on the box
- Use one side of the box to help find \( b_0 \) as shown above
The cross ratio

A Projective Invariant
- Something that does not change under projective transformations (including perspective projection)

The cross-ratio of 4 collinear points

\[ \frac{\| \mathbf{p}_3 - \mathbf{p}_1 \| \, \| \mathbf{p}_4 - \mathbf{p}_2 \|}{\| \mathbf{p}_3 - \mathbf{p}_2 \| \, \| \mathbf{p}_4 - \mathbf{p}_1 \|} \]

Can permute the point ordering
- \( 4! = 24 \) different orders (but only 6 distinct values)
- This is the fundamental invariant of projective geometry

Measuring height

\[ \frac{\| \mathbf{T} - \mathbf{B} \|}{\| \mathbf{R} - \mathbf{B} \|} \times \frac{\| \mathbf{\infty} - \mathbf{R} \|}{\| \mathbf{\infty} - \mathbf{T} \|} = \frac{H}{R} \]

scene cross ratio

\[ \frac{\| \mathbf{t} - \mathbf{b} \| \, \| \mathbf{v}_z - \mathbf{r} \|}{\| \mathbf{r} - \mathbf{b} \| \, \| \mathbf{v}_z - \mathbf{t} \|} = \frac{H}{R} \]

image cross ratio

scene points represented as \( \mathbf{p} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \)

image points as \( \mathbf{p} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \)
What if $v_z$ is not infinity?  
\[ \frac{d_1}{d_2} = \frac{t_\sim_1}{t_2} \]

2HZFig.8.20

\[ t_\sim_1 (v - t_2) / t_2 (v - t_\sim_1) \]
Measuring heights of people

Here we go!
Forensic Science: measuring heights of suspects

Are the heights of the 2 groups of people consistent with each other?

Flagellation, Piero della Francesca

Estimated relative heights

Corrected of radial distortion -- two references reduces uncertainty
Assessing geometric accuracy

The Marriage of the Virgin, Raphael

Estimated relative heights

Criminisi et al., ICCV 99

Complete approach

- Load in an image
- Click on lines parallel to X axis
  - repeat for Y, Z axes
- Compute vanishing points
- Specify 3D and 2D positions of 4 points on reference plane
- Compute homography H
- Specify a reference height
- Compute 3D positions of several points
- Create a 3D model from these points
- Extract texture maps
  - Cut out objects
  - Fill in holes
- Output a VRML model
Interactive silhouette cut-out

Occlusion filling

Geometric filling by exploiting:
- symmetries
- repeated regular patterns

Texture synthesis
- repeated stochastic patterns