Notes on Multi-view Geometry in Computer Vision

Yufei Ye

October 3, 2019
Chapter 1

intro

uncalibrate problems has achieved great progress in recent decade:

- Given 2 images, compute matches between the images, and the 3D position of the points that generate these matches and the cameras that generate the images.

- Given three images, and no other information, similarly compute the matches between images of points and lines, and the position in 3D of these points and lines and the cameras.

- Compute the epipolar geometry of a stereo rig, and trifocal geometry of a trinocular rig, without requiring a calibration object.

- Compute the internal calibration of a camera from a sequence of images of natural scenes (i.e. calibration “on the fly”).

Why these achievement?

- the error that should be minimized in over-determined system

- robust estimation

Solved problem:

- Estimation of the multifocal tensors from image point correspondences, particularly the fundamental matrix and trifocal tensors (the quadrifocal tensor having not received so much attention).

- Extraction of the camera matrices from these tensors, and subsequent projective reconstruction from two, three and four views.
More to learn:

- bundle adjustment to solve more *general* reconstruction problems.
- Metric (Euclidean) reconstruction given minimal assumptions on the camera matrices.
- Automatic detection of correspondences in image sequences, and elimination of outliers and false matches using the multifocal tensor relationships.
Chapter 2

Intro - a Tour of Multiple View Geometry.

2.1 Projective geo

Projective Space: just expansion of Euclidean space. Euclidean space + points at infinity → projective space. Euclidean space is troublesome in one major respect: keeps making exception, e.g. parallel lines.

Coordinates: (homogeneous vectors) represent points by equivalence class of coordinate triples. \((kx, ky, k), k \neq 0!!\). Points at Infinity is represented as \((x, y, 0)\).

In Euclidean / projective geometry, all points are the same. It is just accident that particular point are selected to be original / points at infinity have final coordinate 0.

Transformation of Euclidean / projective space is represented by matrix multiplication. However, in projective space, points at infinity are not preserved, it could land on anywhere.

For practical reason, we sometimes treat line at infinity special / equal...

2.1.1 from Projective geo to euclidean space

Affine Geometry: map "line at infinity" to "line at infinity". The geometry of the projective plane and a distinguished line is known as affine geometry and any projective transformation that maps the distinguished line in one space to the distinguished line of the other space is known as affine transformation.

Euclidean Geometry: Specify line at infinity and two circular points. length ratio / angle can be defined in terms of circular points
CHAPTER 2. INTRO - A TOUR OF MULTIPLE VIEW GEOMETRY.

Absolute Conic: related to camera calibration.

In 2D, circular points: \((1, \pm i, 0)\) lie on every circle (in regular Euclidean coordinate). Circle with homogeneous coordinate \((x, y, w)\) is

\[(x - aw)^2 + (y - bw)^2 = r^2 w^2\]

5 points define an ellipse, 3 points + 2 circular points define a circle.

In 3D, all spheres intersect at a second-degree curve (conic) on the plane at infinity. It is called absolute conic. Angle can be defined in terms of the absolute conic in ANY ARBITRARY coordinate system.

2.2 Camera Projections

Mapping from \(P^3\) to \(P^2\). Generally,

\[
\begin{bmatrix}
  x \\
  y \\
  w
\end{bmatrix} = P_{3 \times 4} \begin{bmatrix}
  X \\
  Y \\
  Z \\
  T
\end{bmatrix}
\]

If all points lie on a plane, say choose this plane as \(z = 0\), then the linear mapping reduces to

\[
\begin{bmatrix}
  x \\
  y \\
  w
\end{bmatrix} = H_{3 \times 3} \begin{bmatrix}
  X \\
  Y \\
  T
\end{bmatrix}
\]

Cameras as points: All points in a ray passing through the center of projection projects to the same point in an image. Thus, all points along such a ray are equal. Thus, the set of all image points = the set of rays through the camera center. Thus, the rays themselves are represented by homogeneous coordinates, specifically, 2-d space of rays in \(P^2\).

camera center as essence. All that is important is the camera center, for it alone determines the set of rays forming the image. Any two images taken from the same point in space are projectively equivalent. (can be transformed only by projective transformation without any information about 3D points.)

Calibrated cameras. For a camera not located on the plane at infinity, the plane at infinity in the world maps one-to-one onto the image plane. It is because any point in the image defines a ray in space that meets the plane at infinity in a single point. Thus\(\text{WTF????}??\), the plane at infinity in the world does not tell us anything new about the image.

However, the absolute conic does. It projects to IAC (Image of the Absolute Conic). Location of IAC is known \(\equiv\) calibrated camera.

There is dual: The angle between two rays / lines in the world, is determined by where they meet the plane at inf, relative to the absolute conic. The
2.3. RECONSTRUCTION FROM MORE THAN ONE VIEW

The relationship between the two image points and \( \omega \) is exactly equal to the relationship between the intersections of the back-projected rays with the plane at infinity, and \( \omega_\infty \).

**Comment:** Analogy in DDG, It’s like you define a metric \( g <,> \), in the real space (3D world) but you need to work on another manifold (image). can we further do better?

If we want to reconstruct the scene from single view, we normally determine the line at infinity for some observed planes in the image. And upgrade the affine reconstruction to Euclidean by assuming angles observed in the scene (most particularly orthogonal lines / planes.)

2.3 Reconstruction from more than one view

**Comment:** normally the usual input is a set of point correspondences!! but how do we get the correspondence? how robust are they? how accurate / noisy should the correspondence? in what scale??

**Ambiguity.** The ambiguity in the reconstruction is expressed by **projective transformations** WTF?!

Because:

\[
P_j X_i = (P_j H^{-1})(H X_i)
\]

This projective ambiguity is unavoidable for 2-view camera. Up to that, the scene can be reconstructed for \( \leq 7 \) points which do not lie in critical configurations.

**Correspondence:** \( x_i \prec - \succ x'_i \)

**Fundamental matrix & fundamental-matrix method.** Basic tool for recon of points sets from 2-views!!!

\[
x_i^T F x_i = 0, \text{rank}(F) = 2, F \in \mathbb{R}^{3 \times 3}
\]

**Comment:** almost forget!!! Fundamental matrix is the basic algebraic entity!!

2.4 Three-view Geometry

**Trifocal tensor.** \( 3 \times 3 \times 3 \) tensor, which relate the coordinates of corresponding in 3 views. It is determined by 3 camera matrices, and determines them, up to projective transformations.

\[
\sum_{ijk} x^i j^i j^i k^i T_{jk} = 0 \text{ WTF}???
\]

The constraints: internal constraints.
correspondence:
\[ x \leftrightarrow l' \leftrightarrow l'' \]

**pros of 3-view over 2 view**
1) allow mixture of line and point correspondences, instead of just point-point correspondence. 2) stability.

**comment:** if we have more views, say view from category, can we further relax the correspondence requirement?? how messy can it be?

### 2.5 Four view geometry and n-view reconstruction.

In general, quadrifocal is the most. the tensor method does not extend to more than four views.

**relax:** affine camera. + a set of points are visible for \( n \) views, then, the factorization algorithm can be used.

**relax:** projective camera. but also requires all points to be visible in all images.

**relax:** various assumptions.

**dominant general method:** bundle adjustment. relation to maximum likelihood solution!!

### 2.6 Euclidean reconstruction

Till now, the camera are all un-calibrated. If we have complete calibration of each of the camera, then some ambiguity can be removed.

the distortion / equality in projective space will arise with same probability. To human, it is not correct in Euclidean sense. Knowledge of the camera calibration is equivalent to being able to determine the Euclidean structure of the scene.
Chapter 3

Projective Geometry and Transformation of 3D

Geometric distortion arises when a plane is imaged by a perspective camera. The imaging process can be modeled by projective geometry.

**comment:** 1. how to rectify planes? how to remove perspective distortion from an image? what is perspective camera?

### 3.1 Planar Geometry

<table>
<thead>
<tr>
<th>geometry</th>
<th>algebra</th>
</tr>
</thead>
<tbody>
<tr>
<td>point</td>
<td>vector</td>
</tr>
<tr>
<td>line</td>
<td>vector</td>
</tr>
</tbody>
</table>
| conic section / conic | symmetric matrix | WTF????

### 3.2 2D projective plane

**homogeneous representation of lines.** $ax + by + c = 0$ represents a line, $(a, b, c)^T \equiv k(a, b, c)^T, k \neq 0$. The set of equivalence classes of vectors in $\mathbb{R}^3 + (0, 0, 0)^T$ gives us projective space $\mathbb{P}^2$.

**homogeneous representation of points.** A point $x = (x, y)^T$ lines on the line $(a, b, c)$ is the same as $(x, y, 1) \cdot (a, b, c)^T = 0$. Homogeneous vector representative of a point

Both line and point in 2-space has 2-dof. **Result** point $x$ lines on the line $l$ iff $x^T l = 0$
### CHAPTER 3. PROJECTIVE GEOMETRY AND TRANSFORMATION OF 3D

<table>
<thead>
<tr>
<th>3-space</th>
<th>( P^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>rays</td>
<td>point</td>
</tr>
<tr>
<td>plane</td>
<td>line</td>
</tr>
<tr>
<td>two non-identcal rays lie on exactly one plane</td>
<td>two points define a line</td>
</tr>
<tr>
<td>two non-identical planes intersect in one ray</td>
<td>two lines intersect at a point</td>
</tr>
<tr>
<td>image plane at ( x_3 = 1 )</td>
<td>homogeneous coordinate to represent point</td>
</tr>
</tbody>
</table>

**Result** intersection of two lines \( l, l' \) is the point \( x = l \times l' \). They are so convinient thanks to homogeneous expression. proof:

\[
\begin{align*}
l(l \times l') = l'(l \times l')
\end{align*}
\]

. Let \( x = l \times l' \), then \( l.x = l'.x = 0 \), \( x \) is at two lines.

**Result** a line joining two points \( x, x' \) is \( l = x \times x' \)

### Ideal points and the line at infinity.

idea points / points at infinity: \( x_3 = 0 \). The whole set lies on a single line, which is \( l_\infty = (0,0,1) \). For any line \( l = (a,b,c)^T \), its ideal point meets \( l_\infty \) at \( (b,-a,0)^T \), which is the direction of lines. Thus, the line at infinity can be thought of as the set of directions of lines in the plane.

### Relation to projective plane / imaging.

**Result** Duality Principle: To any theorem of 2-D projective geometry, there corresponds a dual theorem, which may be derived by interchanging the roles of points and lines.

### Conics and dual conics.

Conic: 2-degree equation curve, and can be expressed homogeneously as symmetric matrix. **Proof:**

a conic in inhomogeneous coordinate is

\[
ax^2 + bxy + cy^2 + dx + ey + f = 0
\]

. Now we homogenize it by \( x \rightarrow x_1/x_3, y \rightarrow x_2/x_3 \), rewrite it:

\[
ax_1^2 + bx_1x_2 + cx_2^2 + dx_1x_3 + ex_2x_3 + fx_3^2 = 0
\]

\[
x^TCx = 0, C = \begin{bmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{bmatrix}
\]

It has 5-dof, and full rank if it is not degenerated case.

**Comment:** The method of fitting a geometric entity (or relation) by determining a null space will be used frequently in the computation chapters throughout this book.

**Result** \( l \) tangent to \( C \) at point \( x \) is given by \( l = Cx \) **Proof:**
Dual conics / conic envelopes. $C^*$ The is that is the dual of $x$s forms a conic $C^* = C^{-1}$ (if full rank) in dual space. $C^*$ is the adjoint matrix of $C$.

Degenerate conics. $C$ is not full rank. rank = 2 two lines, rank = 1 repeated line.

(comment: from mathematic view: adjoint matrix of $C$ and dual of $C$???)

3.3 Projective transformations

Geometry is the study of properties invariant under groups of transformations – Felix Klein.

(Geometry def): Projectivity / collineation / projective transformation / homography. a group of transformations. an invertible mapping $h : \mathbb{P}^2 \rightarrow \mathbb{P}^2$ such that three points $x_1, x_2, x_3$ lie on the same line iff $h(x_i), i = 1, 2, 3$ do. It preserves line!

(Algebraic def) Projectivity. $h$ is Projectivity iff there exists a non-singular $3 \times 3$ matrix $H$, for any point in $\mathbb{P}^2$ represented by a vector $x$ it is true that $h(x) = Hx$. $h$ can be represented by $H$. There are two equivalent ways to define, 1) Any invertible linear transformations of homogeneous coordinates is projectivity. 2) Any projectivity arises as such a linear transformations.

Projective transformation. planar projective transformation is a linear transformation on homogeneous 3-vector (coordinates). It can be represented by a non-singular $3 \times 3$ matrix. $x' = Hx$. Scale of $H$ does not matter, so $H$ is homogeneous matrix, with 8-dof.

After projective transformation, the projective properties remain invariance.

Perspectivity. If two coordinate system defined in two planes are both Euclidean coordinate, then the mapping is called perspectivity, with 6 – dof.

(comment: Figure 2.5 in page 36. If the world lie in a planar, even locally, or camera rotating, or shadow, there is perspective images, where lots of good properties arise.

Fundamentally different ways to transform line / points / conics. if $x' = Hx$, $l' = H^{-T}l$. $C' = H^{-T}CH^{-1}$, $C' = HC^*H^T$. WTF???? think more!!

(comment: Conic and dual conic undergo two different transformation!)
CHAPTER 3. PROJECTIVE GEOMETRY AND TRANSFORMATION OF 3D

<table>
<thead>
<tr>
<th>class</th>
<th>expression</th>
<th>invariants</th>
<th>dof</th>
<th>#point</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isometries</td>
<td>$R \ t$</td>
<td>length, angle, area</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Similarity</td>
<td>$sR \ t$</td>
<td>ratio of length, angle, parallel</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Affinity</td>
<td>$A \ t$</td>
<td>ratio of length of $\parallel, \parallel$, area ratio</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Projective</td>
<td>$A \ t$</td>
<td>cross ratio of 4 collinear points</td>
<td>8</td>
<td>4</td>
</tr>
</tbody>
</table>

3.4 A hierarchy of transformations / projective linear group

**Isometries.** preserves Euclidean distance (iso=same, metric). Orientation-preserving if upper left $\text{det}(R) = 1$. It form a group.

**Metric Structure.** structure is defined up to a similarity.

**Essence of Affinity.** 1) scaling in orthogonal direction. $A = R(\theta)R(-\phi)DR(\phi)$ always decomposable! 2) affinity is the most general linear transformation that fix $l_\infty$

**Affinity is the middle ground for similarity and projective transformation.** in affinity, $\text{det}(A)$ fully define scaling anywhere on the plane, and orientation. Also, ideal point remains ideal.

**Decomposition of projective transformation.** transformation higher in the hierarchy than the previous one. $H_P(2\text{dof})$ moves the line at infinity, $H_A(2\text{dof})$ affects the affine properties. $H_S(A\text{dof})$ is similarity transformation.

$$H = H_{sim}H_AH_P = \begin{bmatrix} sR & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} K & 0 \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} I & 0 \\ v^T & v \end{bmatrix} = \begin{bmatrix} A & t \\ v^T & v \end{bmatrix}$$

WTF???? think! Where $A = sRK + tv^T$. $K$ is a normalized upper-tri matrix $\text{det}K = 1$

**comment:** Well structured composition is easier to learn / compute??

3.5 The projective geometry of 1D

**Cross Ratio.**

$$\text{cross}(x_1, x_2, x_3, x_4) = \frac{|x_1x_2||x_3x_4|}{|x_1x_3||x_2x_4|}, \text{det} \begin{bmatrix} x_{i1} & x_j1 \\ x_{i2} & x_j2 \end{bmatrix}$$
The definition of $|x_i x_j|$ is the signed distance from $i$ to $j$, if $x_{last} = 1$. Cross ratio is valid for infinity points, invariant with choice of coordinate.

**Comment:** Once you fix the camera position / points in the world lie in one plane, the cross ratio of those points projected on the plane, is the same!!!, irrelavant of where my imaging plane puts.

### 3.6 Topology of hte projective plane

$\mathbb{P}^2$ is equivalent of the set of all homogeneous 3-vector!! $x = (x_1, x_2, x_3)$ can be normalized as $x_1^2 + x_2^2 + x_3^2 = 1$. It lies in a sphere $S^2$ in $\mathbb{R}^3$. But $x, -x$ is the same point. Thus, 2-1 correspondence from points in $\mathbb{P}^2$, $S^2$, line corresponds to great circle.

**WTF???** In the language of topology, the sphere $S^2$ is a 2-sheeted covering space of $\mathbb{P}^2$. This implies that $\mathbb{P}^2$ is not simply-connected, which means that there are loops in $\mathbb{P}^2$ which cannot be contracted to a point inside $\mathbb{P}^2$. To be technical, the fundamental group of $\mathbb{P}^2$ is the cyclic group of order 2.

$P^2 \equiv$ a disk with opposite points on its boundary identified. $\equiv$ glued together. $P^2$ is not orientable.

**Comment:** will topology constraints helps learn a canonical uv map?? **Comment:** the projection is camera dependent? can we map images from a categories to the same mapping?? **Comment:** what does images from a category really capture??

### 3.7 Recovery of affine and metric properties from images

**DoF** We actually only want to restore to a similarity transformation. so, only 4 dof is needed. Line at infinity (2dof) + 2 circular points (2dof). $\equiv$ restore in the decmoposition chain .

**Vanishing line.** affine properties can be recovered from identifying imaged vanishing line.

**Length ratio.** affine properties can also be recovered by length / cross ratio. To reconstruct similarity / metric.

**Result** circular points $I$, $J$ are fixed points under $H$ iff $H$ is a similarity.

**Proof:**

To easily calculate.

$I, J$ determines the orthogonal direction of Euclidean geometry. Some intuition: $(0, 1, 0)^T, (1, 0, 0)^T$ are packed into $I = (1, 0, 0) + i(0, 1, 0)$. 


The conic dual to the circular points.

\[ C^\ast_\infty = IJ^T + JI^T \leftrightarrow I^T C^\ast_\infty I = 0, \ J^T C^\ast_\infty J = 0 \]

. It is a degenerate rank2 conic.

**Degenerate rank2 conic.** passing two lines \( l, m \), \( C = lm^T + ml^T \). Result dual conic \( C^\ast_\infty \) of circular points \( I, J \) are fixed points under \( H \) iff \( H \) is a similarity. Result under euclidean coordinate, \( C^\ast_\infty = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \)

1. \( C^\ast_\infty \) is generally a symmetric \( 3 \times 3 \) matrix, with 5-dof. But \( \det(C^\ast_\infty) = 0 \), thus, it has 4-dof.

2. \( l_\infty \) is in the null space of \( C^\ast_\infty \), \( C^\ast_\infty l_\infty = 0 \)

Angles on the projective plane

\[ \cos \theta = \frac{l_1 m_1 + l_2 m_2}{\sqrt{(l_1^2 + l_2^2)(m_1^2 + m_2^2)}} = \frac{i^T C^\ast_\infty m}{\sqrt{(i^T C^\ast_\infty i)(m^T C^\ast_\infty m)}} \]

The later one to express angle is invariant of projective transformation. bc,

\[ i^T C^\ast_\infty m \rightarrow i^T H^{-1} H C^\ast_\infty H^T H^{-T} m = i^T C^\ast_\infty m \]

**Result** Once the conic \( C^\ast_\infty \) is defined, 1) the Euclidean angles can be measured. 2) the ratio of length can also be measured. (by \( \sin A, \sin B \)).

Recovery of metric from image

\[ C^\ast_\infty' = HC^\ast_\infty H^T = (H_P H_A H_s) C^\ast_\infty H^T = \begin{bmatrix} K K^T & K K^T v \\ v^T K K^T & v^T K K^T v \end{bmatrix} \]

So, if we can observe \( C^\ast_\infty' \), the \( H = U \) can be obtained by SVD.

\[ C^\ast_\infty' = U \Sigma U^T, \Sigma = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \]

**How to observe** \( C^\ast_\infty' \)? It can either from two orthogonal lines / known angle, or from length ratio.

**Orthogonal lines.** two lines \( l, m \), \( C^\ast_\infty \) can be expressed as \( \begin{bmatrix} K K^T & 0 \\ 0^T & 0 \end{bmatrix} \) (because we have affinity, set \( v = 0 \) in the previous result. \( S = K K^T \) to guarantee symmetric.) Therefore, one pair of orthogonal lines forms one constraints:

\[ (l'_1 m'_1, l'_2 m'_2 + l'_2 m'_2) (s_{11}, s_{12}, s_{22})^T = 0 \]

WTF????this always has \( \sigma = 1, 1, 0 \)??
3.8. More Properties of Conics

length ratio. The image conic (an ellipse) intersect with $l_\infty$ at circular points.

Orthogonal lines. (one step) Suppose we start from perspective image (instead of affine transformation as above) one pair of orthogonal lines provides one constraints on $C^\infty$

$$
(l_1m_1,(l_1m_2+l_2m_1)/2,l_2m_2,(l_1m_3+l_3m_1)/2,(l_2m_3+l_3m_2)/2,l_3m_3)c = 0
$$

c = (a, b, c, d, e, f)^T

comment: but the two-step approach (termed stratified / stratification) also apply in 3D???

3.8 More properties of conics

Polarity. for any $x, C$ defines a line $l = Cx$. This line has a property: for the two points $x_1, x_2$ it intersects with the conic, their tangent will meet back at $x$. $x$ is the pole of $l$ w.r.t $C$, $l$ is the polar of $x$ w.r.t. $C$.

A map: Correlation between pts and lines. Hence! the conic $C$ induces a map between points and lines of $\mathbb{P}^2$. And interestingly, the mapping is invariant under projective transformation, since it is all based on incidence.

Conjugate points. if point $y$ is on the line of $l = Cx$, then $y^TCx = 0$. $y, x$ are conjugate w.r.t $C$.

Classification of conics under perspective. $C$ is symmetric thus always has real eigenvalue. under projective transformation, $C' = U^{-T}CU^{-1} = D = \text{diag}(\sigma_1d_1, \sigma_2d_2, \sigma_3d_3)$, $\sigma_i = \pm 1, d_i > 0$, $D = \text{diag}(s_i)^T\text{diag}(\sigma_i)\text{diag}(s_i)$. Thus, the type of conics can be enumerated by enumerating $\sigma_i$. See table 2.2 in page 60.

Classification of conics under affinity. $l_\infty$ is preserved, thus, has 3 classes, by intersecting line and conic: 1) ellipse (does not intersect) 2) hyperbola (tangent) 3) parabola (2 points)

3.9 Fixed points and lines

fixed points corresponds to eigen vector. comment: does not need $\lambda = 1$ because $e, \lambda e$ is the same point. Fixed line are usually not mapped pointwise.

Euclidean matrix. two circular points $I, J$, which corresponds to eigenvalues $\{e, e^{-d}\}$. Third eigenvector is called pole, $\lambda_3 = 1$. The Euclidean transformation is equivalent of rotate $\theta$ around pole, without translation. comment: KDC!
Similarity. ideal fixed points: circular points. eigenvalue \( \{1, se^{i\theta}, se^{-i\theta}\} \)

Affine. two eigen vector: ideal points \( x_3 = 0 \) (not circular points?). tird eigen vector is finite in general.
Chapter 4

Projective Geometry and Transformation of 3D

similar properties: \( l_\infty \) in \( \mathbb{P}^2 \), \( \pi_\infty \) in \( \mathbb{P}^3 \).

different properties: lines always intersects.

4.1 Points and projective transformation
dual, points - planes, lines- lines.

4.1.1 Planes
Incidence and relations. 1) points on the plane: \( \pi^T X = 0, \pi = (n, d) \). 2) note that two points defines a line is not trivial as in \( \mathbb{P}^2 \), \( l = x \times y \).

Three points define a plane.

\[
\begin{bmatrix}
X_1^T \\
X_2^T \\
X_3^T
\end{bmatrix}_{3 \times 4}
\pi = 0
\]

rank(3), then 1-dim null-space. To have similar expression as in \( \mathbb{P}^2 \), \( l = x \times y \equiv [x, y]^T \cdot l = 0 \). we have \( \text{det}[X, X_1, X_2, X_3] = 0 \), for any points on \( \pi \).

\[\text{det}M = X_1D_{234} - X_2D_{134} + X_3D_{124} - X_4D_{123} = 0, \text{so} \pi = (D_{234}, -D_{134}, D_{124}, -D_{123})^T\]

Three points defines a point. Direct apply dual.

Projective transformation. Under point transformation \( X' = HX, \pi' = H^{-T}\pi \)
Parametrized points on a plane. Points \( X \) on the plane \( \pi \) can be written as \( X = M_{4 \times 3}x \), \( x \) in any point in \( \mathbb{P}^2 \), \( M \) generate null-space of \( \pi^T \) (\( \pi^TM = 0 \)).

4.1.2 Lines

DoF and awkward problem. Lines have 4-dof. the natural coordinate are 5. homogeneous 5 vetor cannot easily be used in math expressions togetehr with 4-vectors representing points and planes. 

WTF????comment: it is counter-intuitive. coz first defines a point, and defines a direction. it's 6-dof. what is misisng here???

Null-space and span representation. Two points A, B, the line can be represented as \( W = \begin{bmatrix} A^T \\ B^T \end{bmatrix} \). with the properties: (1) span of \( W^T \) is the pencil of points \( \lambda A + \mu B \) on the line (2) span of 2d right null-space of \( W \) is the pencile of planes with the line as axis. comment: use span to cancel the choice of points! this is not in dual space. WTF????comment: are these represent the same line???

Two plane P, Q. \( W^* = \begin{bmatrix} P^T \\ Q^T \end{bmatrix} \) (1) span of \( W^*^T \) is the pencil of planes with the line as axis. (2) span of 2d right null-space of \( W^* \) is the pencil of points. comment: use span to cancel the choice of points! this is not in dual space.

Result \( W^*W^T = 0_{2 \times 2} \).

Join and incidence. the join of a point \( X \) and a line \( W \) is

\[
\text{null}(M) = \text{null} \begin{bmatrix} W \\ XT \end{bmatrix}
\]

the point defined by intersection of line \( W \) and plane \( \pi \) is

\[
\text{null}(M) = \begin{bmatrix} W^* \\ \pi^T \end{bmatrix}
\]

Plucker matrices. where line is represented by a \( 4 \times 4 \) skew-symmetric homogeneous matrix. line joining 2 points \( A, B \) is represented by \( L \)

\[
L = AB^T - BA^T
\]

comment: formally similar to cross product?

It has properties: 1) \( L \) in rank 2. 2-dimensional null-space is spanned by the pencil of planes with the line as axis. \( LW^*^T = 0 \) 2) \( L \) has 4dof. because a) skew-symmetric matrix has 6dof. 1 in scale, 1 is constrained by \( \text{det}(L) = 0 \) 3) is gerazalized of \( l = x \times y \) 4) \( L \) is independent of choice of A, B, because if \( C = A + \mu B \), we can derive the same \( L \) 5) point transformation, matrix is transformed as \( L' = HLH^T \)

\[
L^* = PQ^T - QP^T, L^* = H^{-T} LH^{-1}
\]
Join and incidence.

\[ \pi = L^*X, X = L\pi \]

**Plucker line coordinates** : six non-zero elements of \( L_{4 \times 4} \).

\[ \mathcal{L} = \{l_{12}, l_{13}, l_{14}, l_{23}, l_{42}, l_{34}\}, \det L = l_{12}l_{34} + l_{13}l_{42} + l_{14}l_{23} = 0 \]

every line in \( P^3 \) defies a surface in \( P^5 \) – Klein quadric.

**comment**: It must satisfy the \( \det L = 0 \) constraints. It is not homogeneous, but a subspace

**Result** Two lines \( \hat{L} \) are coplannar (and thus intersect) iff \( (L|\hat{L}) = 0 \) \( (L|\hat{L}) \) is bilinear product!

some properties: 1) \( \det[A, B, \hat{A}, \hat{B}] = (L|\hat{L}) \) it does not depends on the choice of \( A, B, \hat{A}, \hat{B} \) 2) it can also defy by dual plane \( \det[P, Q, \hat{P}, \hat{Q}] = (L|\hat{L}) \) 3) \( L \) only represent a line in \( P^3 \) if \( (L|L) = 0 \) 4) if \( L \) is defied by \( P, Q, \hat{L} \) defind by \( A, B \). \( (L|L) = (P^T A)(Q^T B) - (Q^T A)(P^T B) \) WTF????

### 4.1.3 Quadrics and dual quadrics. (conic in \( P^3 \))

\[ X^T QX = 0 \]

1. 9-dof

2. 9 points defines a quadratic. **comment**: why???

3. a quadirc defines as polarity between a point and a plane \( \pi = QX \)

4. quadric is transformed by \( Q' = H^{-T}QH^{-1} \)

5. dual quadrics are equations on planes \( \pi^T Q^*\pi = 0, Q^* = \text{adjoint} Q \text{or} Q^{-1} \)

**comment**: what is adjoint mean??

6. Dual quadric is transformed by \( Q'^* = HQ^*HT \)

### 4.1.4 classification of quadrics

**signature** quadric is classified by diagnose matrix \( D \) where the element are only 1, 0, \(-1\). Number of +1- number of \(-1\).

**classification** the projective type of a quadric is uniquely determined by rank and signature. Table 3.1 p74.

### 4.2 Twisted Cubics \( c \)

it is basically a parameterized curve!
2-d projective plane as parameterized curve.

\[
\begin{bmatrix}
  x_1 \\
  x_2 \\
  x_3
\end{bmatrix} = A \begin{bmatrix}
  1 \\
  \theta \\
  \theta^2
\end{bmatrix}
\]

A twisted cubic is defined to be a curve in \(P^3\). Similarly we have \(A\) in \(4 \times 4\) for \(P^3\).

The standard form / canonical form is \(c(\theta) = (1, \theta, \theta^2, \theta^3)^T\)

Application. 1) horopter for 2-view geometry. 2) defining the degenerate set for camera resectioning.

4.3 Hierarchy of Transformation

Hierarchy Table 3.2. 7-dof for similarity – (3 rotation, 3 translation, 1 isotropic scaling. ; 5 for affinity \((3 \times 3 - 3 - 1 = 5)\) WTF???? other intuition? s); – 3 projective part.

Screw decomposition a decomposition of a euclidean transformation.

comment: will be useful to discuss special motions.

Result Any particular translation and rotation is equivalent to a rotation about a screw axis together with a translation along the screw axis. The screw axis is parallel to the rotation axis. It can be determined from the fixed points of the \(4 \times 4\) matrix representing the Euclidean transformation.

4.4 The plane at infinity

\(\pi_\infty/l_\infty\) allow affine properties to be measured. \(\Omega_\infty\) / circular points allows metric properties to be measured.

By defining infinity plane, we have \(\mathbb{P}^3\) that any two pairs of planes intersect in a line.

comment: \(\pi_\infty\) has 3 dof. why??? scale???

Result the plane at infinity is a fixed plane under \(H\) iff \(H\) is an affinity. The plane is fixed as a set, not pointwise.

4.5 The absolute conic

\(X_1^2 + X_2^2 + X_3^2 = 0, X_4 = 0\)

Firstly, it means those points lie on the plane at infinity, \(X_4 = 0\); Secondly, it means those points are in a conic \(\sum_i^3 X_i^2 = 0\). \(\Omega_\infty\) corresponds to a conic \(C\) with \(C = I_{3 \times 3}\). 5dof. The points are all imaginary points / no real points.

Result Absolute conic, \(\Omega_\infty\) is a fixed conic under the projective transformation \(H\) iff \(H\) is a similarity transformation.
4.6. THE ABSOLUTE DUAL QUADRIC

Orthogonality and polarity. \( d_1, d_2 \) are orthogonal if \( d_1^T \Omega_\infty d_2 = 0 \). It is the definition of conjugacy w.r.t \( \Omega_\infty \).

If image points are conjugate w.r.t IAC, then corresponding rays are orthogonal.

**comment:** The 3 \( \times \) 3 matrix indicates that the conic lies in the plane at infinity.

4.6 The absolute dual quadric

**comment:** it is actually more important than primal absolute quadric. Because it can be expressed in one algebraic form.

Euclidean / canonical coordinate. The envelope of all plane tangent to \( \Omega_\infty \):

\[
Q_\infty^* = \begin{bmatrix}
I_{3 \times 3} & 0 \\
0 & 0
\end{bmatrix}
\]

**WTF????** **comment:** why \( \Omega_\infty \) can not be expressed as the same form???

**Proof:**
\( \pi = (n, d) \). \( \pi^T Q_\infty^* \pi = 0 \), thus, \( n^T n = 0 \). On the other hand, \( n \) represents the lines that \( \pi \) meets with \( \pi_\infty \). From polarity w.r.t \( \Omega_\infty \), \( n^T I n = 0 \) iff it is tangent to the conic.

Another **Proof:**
Consider all conic in form of \( Q = \text{diag}(1,1,1,k) \). when \( k \to \infty \), it only contains points \( \sum X_i^2 = 0, X_4 = 0 \) (expand to infinity). The dual of \( Q \) is \( Q^* = \text{diag}(1,1,1,k^{-1}) \to \text{diag}(1,1,1,0) \).

**Result** \( Q_\infty^* \) is fixed iff \( H \) is similarity.

**Result** \( \pi_\infty \) is the null-vector of \( Q_\infty^* \).

**Result** angle is given by \( \cos \theta = \frac{\pi_1^T Q_\infty^* \pi_2}{\sqrt{00}} \), which is invariant to \( H \). **comment:** this is a metric invariant to \( f(X) \)

**WTF????** **comment:** p84: should be \( \pi_\infty, Q_\infty^* \)?
Chapter 5

Estimation – 2D Projective Transformation

5.1 DLT

Setting. 2D 2D point correspondence.

Key equation.

\[ x_i' = Hx_i, \text{or, } x_i' \times (Hx_i) \]

**comment:** This is a common trick to convert inhomogeneous equation to homogeneous.

Solution. Inhomogeneous solution (not recommended), vs homogeneous solution.

**Solution from lines and other entities.** \( l' = H^{-T}l. \) Count dof. Care be taken for mixed type.

5.2 Different cost functions

**Algebraic distance.** [Bookstein-79] pros: linear, cheap to compute. Starting point for non-linear min of a geometric / statistical cost function. Cons: no geometrically / statistically meaningful, or not expected intuitively. – can be solved by choice of normalization. **comment:** You might want to consider the following when designing the loss function / supervision!!

**Geometric distance.** \( x: \) measured imaged coordinates, \( \hat{x}: \) estimated value of points, \( \bar{x}: \) true values of the points.
CHAPTER 5. ESTIMATION – 2D PROJECTIVE TRANSFORMATION

Error in one image. assume in the first image, \( \bar{x} = x \). True in calibration pattern where points are measured to a very high accuracy. **Transfer error:**

\[
\sum_i d(x_i', H\bar{x}_i)^2
\]

Symmetric transfer error. errors occur in both images. Forward, backward transformation.

\[
\sum_i d(x_i', Hx_i)^2 + \sum_i d(H^{-1}x_i', x_i)^2
\]

Estimated homography is the one for the above is minimized.

Reprojection error. Estimating a "correction" for each correspondence. Estimate both \( \hat{H} \) and \( \hat{x}_i \) perfectly matched correspondence \( \hat{x}_i \)

\[
\sum_i d(x_i, \hat{x}_i)^2 + \sum_i d(x_i', \hat{x}_i'). s.t. \hat{x}' = \hat{H}\hat{x}
\]

MLE of homography and correspondence. *WTF????* is there sth deep??

Comparison of geometric and algebraic. Fitting \( V_H \) on point \( X = (x, y, x', y') \).

Conic analogue. fitting conic to 2D points. *WTF????* I’m lost...

Sampson error. A middle ground in between algebraic and geometric cost function in terms of complexity. Close approximation to geometric error. 1st order method! The key is to consider constraints \( Ah = 0 \) as a cost depends on \( X, C_H(X) + C_H(\hat{X}) = 0 \). We wanna solve for \( \|\delta_X\|^2 \) subject to Taylor Expansion on \( X = \delta X + \hat{X} \).

**Result** Sampson Error is:

\[
\|\delta_X\|^2 = \epsilon^T (J J^T)^{-1} \epsilon
\]

, where \( J \) is \( \nabla C_X \)

*WTF????* need to run through it yourself.

To find \( H \) for all points,

\[
D = \sum_i \epsilon_i^T (J_i J_i^T)^{-1} \epsilon_i
\]

Both \( \epsilon, J \) depends on \( H \).

Linear cost function. \( C_H(X) = A(X)h \) is linear w.r.t \( X \) is important. 1) the Taylor expansion is exact – Sampson error is geometric error. 2) Finding \( H \) becomes a hyperplane fitting problem.
5.3. **STATISTICAL COST FUNCTIONS AND MAXIMUM LIKELIHOOD ESTIMATION**

A geometric interpretation. all measurements is represented by a single point in a measurement space \( \mathbb{R}^N \).

1. measurement space \( \mathbb{R}^N \)

2. model: a subset \( S \) of points in \( \mathbb{R}^N \). if \( X \) in the subset, it satisfy the model.

Given the \( X \), wanna find vector \( \hat{X} \) closes to \( X \) that satisfy the model

**Error in both image.** \( N = 4n \), if \( x \) and \( H \) are selected, the model defines \( x' \). thus, the feasible subset \( S \) has \( 2n + 8 \) dof. The geometric distance becomes: given a set of measured point pairs, which corresponds to a point \( X \) in \( \mathbb{R}^N \), and an estimated points \( \hat{X} \in \mathbb{R}^N \) lying on \( S \).

\[
\min \|X - \hat{X}\|^2 = \sum_i d(x_i, \hat{x}_i)^2 + d(x'_i, \hat{x}'_i)^2
\]

**Error in one image.** model becomes \( x'_i = H \bar{x}_i \)

**5.3 Statistical cost functions and Maximum Likelihood estimation**

Assume noise is Gaussian on each image coord without bias, with uniform standard deviation \( \sigma \). comment: this might not hold true for imaging reason.

**Error in one image.** MLE of the homography \( \hat{H} \) maximizes the log-likelihood, which is equivalent as \( \sum_i d(x'_i, H \bar{x}_i)^2 \). In short, MLE is equivalent to minimizing the geometric error function.

\[
\log P\{x'_i|H\} = -\frac{1}{2\sigma^2} \sum_i d(x'_i, H \bar{x}_i)^2 + \text{constant}.
\]

**Error in both images.** MLE is identical with minimizing reprojection error function. \( \sum_i d(x_i, \hat{x}_i)^2 + d(x'_i, \hat{x}'_i)^2 \)

Mahalanobis distance. now we assume covariance matrix \( \Sigma \). Then max log-likelihood is equivalent to minimizing the Mahalanobis distance

\[
\|X - \bar{X}\|_\Sigma^2 = (X - \bar{X})^T \Sigma^{-1} (X - \bar{X})
\]

**5.4 Transformation invariance and normalization**

Question: Invariance of the algorithm to different choices of coordinates.
5.4.1 Invariance to image coordinate transformations

To what extent is the result of an algorithm that minimizes a cost function to estimate a homography dependent on the choice of coordinate in the image?

coordinates $x$ in one image are replaced by $\tilde{x} = Tx$, and the in the other image $\tilde{x}' = T'x'$. Then, $H$ are transform correspondingly: $\tilde{H} = T'HT^{-1}$

The next question is, whether the outcome of the algorithm is independent of the $T,T'$. 

5.4.2 Non-invariance of the DLT algorithm

Setup: a set of correspondence $x_i \leftrightarrow x'_i$ and $H$ that is the result of the DLT. Consider further a related set $Tx_i \leftrightarrow T'x'_i$, let $\tilde{H} := T'HT^{-1}$. Question: does DLT applied to the $\tilde{x}_i, \tilde{x}'_i$ yield $\tilde{H}$?

Result $T'$ be a similarity transformation with scale factor $s$, $T$ is any arbitrary projective transformation. $H$ is any homography and $\tilde{H} := T'HT^{-1}$. Then, $\|\tilde{A}\| = s\|A\|$. In other word,

$$d_{\text{algebraic}}(\tilde{x}'_i, \tilde{H}\tilde{x}_i) = s d_{\text{alg}}(x'_i, Hx_i)$$

Remark: WTF????(miss the argument p106).

1. No one-to-one correspondence between $H$ and $\tilde{H}$.
2. $\|H\| = 1$ is not equivalent to $\|\tilde{H}\| = 1$

comment: only dependent on $T'$. WTF????

5.4.3 Invariance of geometric error

Minimizing geometric error is invariant to similarity transformation.

$$d(\tilde{x}', \tilde{H}\tilde{x}) = d(T'x', T'HT^{-1}Tx) = d(T'x', T'Hx) = d(x', Hx)$$

, where Euclidean distance is unchanged under Euclidean transformation $T'$.

5.4.4 Normalizing transformation

We have seen in the previous section, that there are some coordinate systems better than others for computing a 2D homography. Some normalization should be carried out before applying the DLT algorithm. Two pros: 1) result is more accurate 2) it undo the arbitrary choice of scale and origin. The algebraic min is carried out in a canonical frame.

comment: here the term canonical comes with reason for computational accuracy.
5.5. **ITERATIVE MINIMIZATION METHODS**

**Isotropic scaling.**

1. translate points to the origin.
2. scale points so that distance from the origin is equal to $\sqrt{2}$
3. apply those transformation to both image independently.

**Why is normalization essential? (pre-conditioning)** (must not consider as optional) Think in detail of DLT. we essentially solve a SVD for $A = U\Sigma V^T$, to solve $h$, $Ah = 0$. $A$ is $2n \times 9$ but should have rank 8. However, it is impossible due to noisy data. Thus we want to find $h$ to minimize $\|Ah\|$. It is equivalent to find a rank 8 matrix $\hat{A}$ that satisfy exactly $\|\hat{Ah}\| = 0$ and closes to $A$ in Frobenious form. $\hat{A} = U\hat{\Sigma}V^T$. $\min_{\hat{A}}\|A - \hat{A}\|_F = \|UDV^T - U\hat{D}V^T\|_F = \|D - \hat{D}\|_F$, s.t.$\text{rank}(\hat{A}) = 8$.

The element of $A$ is just $xx', xy', ww'...$. The $xx'$ will be in the order of $10^4$, $ww'$ will be one. The min above is just increase / detrese the value such that the sum of the change is minimum to reach rank 8 matrix. But changing small amount of $w$ (100) will have a huge effect in $H$ but $xx'$ will not.

**comment:** It is insightful to switch between linear equation, SVD, null space, rank, optimize matrix, find vector.

From the condition number aspect, the condition number of DLT is $d_1/d_n-1$. The exact arithmetic results is independent of normalization, but it will diverge from the correct result in the presence of noise. large condintion number will amplify the effect.

**Non-isotropic scaling and variants.** experiment suggests it does not lead to significantly better results. Another variant is based on the observation that some column of $A$ are not affected by noise ($w,w'$, thus those column should not be varied to find $A$).

**Scaling with points near infinity.** It makes no sense to normalize the coordinates of points in the infinite plane by setting the centroid at the origin, since centroid may have very large coordinates. 

5.5 **Iterative minimization methods**