

2d Transform

- Method: finding correspondences
- transformations: translation, rotation, aspect, affine (change shape), perspective
 - p' = T(p) = Mp
 - when p is in R^2 -> 2x2 M
- uniform scaling: same scalar for all components
- non uniform -> new aspect ratio

shear

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x + sh_x \cdot y \\ sh_y \cdot x + y \end{bmatrix}$$

Translate

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Scale

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Rotate

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Shear

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & sh_x & 0 \\ 0 & 1 & sh_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

parallel lines remain parallel

feature based alignment

eigenvectors are orthonormal

$$\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ w \end{bmatrix}$$

symmetric A eigenvector w largest λ max $\frac{x^T A x}{x^T x}$ unit vectors max num

non symmetric -> sVD: $U \Sigma V^T$

u rotation eigenvectors of AA^T

σ sqrt of $A^T A$ λ 's

solving least squares invert A find v (m,b) point -> minimizing error

by argmin:

Satisfying:

Solution satisfies $(A^T A)v^* = A^T y$

or

$$v^* = (A^T A)^{-1} A^T y$$

- when A is square, rank < rows or rank < cols
- homo ls -> orthogonal to vectors $\text{argmin}(\|Av\|^2)$ smallest ev from ATA, smallest singular v of A
- for 2 variables need at least 2 eq

ransac - reduce noise

- not true matches -> outlier (minimising lq creates error)
- look for inliers

For N times

- select random seed group s points
 - more points = more robust
- compute transformation for seed group
- find inliers to this transformation
 - point whose d is < t
- if large then recomputer estimate M from all of inliers
 - d > inliers accept and refit using all inliers

photometric: type, direction, intensity of light, surfaces reflectance property
optical: focal length, fov, aperture, shutter speed

- image formation
 - every point on a tree blends its colors all across the film
 - > barrier known as aperture (only one makes it through) flipped
 - f = focal length, c = center of camera
 - f distance from pinhole and img
 - lines all to vanishing point
- facts about projection
 - line in 3d -> line in 3d
 - parallel 3d -> loses, but converge to centr. point
 - distant objects are smaller
- increase aperture (when too dark)

$$K[R]tX$$

Camera Properties

World Properties

Image Coordinates: (u,v)

Intrinsic Matrix (3x3)

Extrinsics: (R,t)

- extrinsic ass: no rotation, camera at 000 4th col
- intrinsic ass: opticalcenter000,unit aspe,!skew
- rotations:

Rotation around the coordinate axes, counter-clockwise:

$$R = R_x(\alpha)R_y(\beta)R_z(\gamma)$$

$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix}$$

$$R_y(\beta) = \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix}$$

$$R_z(\gamma) = \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \end{bmatrix}$$

5

6

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & s & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Stereo: given 2 cameras/correspondences
find: world coordinate for known point

Epipolar Geometry

- baseline connects both origins
- epioles where baseline intersects w img planes
- epipolar lines btw point and e -> p' (must be from same light ray)
 - rays given by ep line
- only translation -> ep lines horizontal, e ∞
- forward motion -> ep lines out from optical cent
- Fe = 0 F^te' = 0, F r=2, F 7 deg freedom, !unique
- > need at least 7 correspondences

eight point algo

Each point gives an equation:

$$[u_i', u_i'', u_i, v_i', v_i'', v_i, w_i', w_i'', w_i, 1] \cdot \begin{bmatrix} f_x & 0 & 0 & 0 & 0 & 0 \\ 0 & f_y & 0 & 0 & 0 & 0 \\ 0 & 0 & f_x & 0 & 0 & 0 \\ 0 & 0 & 0 & f_y & 0 & 0 \\ 0 & 0 & 0 & 0 & f_x & 0 \\ 0 & 0 & 0 & 0 & 0 & f_y \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ t_x \\ t_y \\ t_z \end{bmatrix} = 0$$

Stack equations to yield U:

$$U = \begin{bmatrix} u_1' u_1'' & u_1 u_1' & u_1 v_1' & u_1 w_1' & u_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

How to solve for t (if unrolled):

$$\arg \min_{\|U\|_F} \|U\|_F^2 \rightarrow \text{Eigenvector of } U^T U \text{ with smallest eigenvalue}$$

- f output is not perfect (bc not perfectly aligned)
- closest matrix F with lower rank = take highs σ
 - multiply SVD out for new F

- normalize img coordinates: center img data @ origin 0> scale msd btw origina & data points = 2
- apply 8-point algo: compute F
- enforce rank-2 constraint
- unnormailize coordinates: F back to og units
- $T^T F T =$ og coords

Minimizing via $U^T U$ minimizes sum of squared algebraic distances between points p_i and epipolar lines Fp_i' (or points p_i' and epipolar lines $F'p_i$):

$$\sum_i (p_i^T F p_i')^2$$

$$\sum_i d(p_i, F p_i')^2 + d(p_i', F^T p_i)^2$$

Calibrated case

- know intrinsic & extrinsic -> set coord to cam 1
 - $M_1 = K[I, 0], M_2 = K'[R, t]$, M produce p
- $K^{-1}p$ producing normalized coordinate p^{\wedge} with K as identity matrix (canonical view)
- coplanar: $Rp^{\wedge}, t, p^{\wedge} \rightarrow p^T (x \times R p) = 0$

$$\tilde{p}^T \begin{bmatrix} 0 & t_x & t_y \\ t_x & 0 & -t_z \\ -t_x & t_z & 0 \end{bmatrix} R \tilde{p} = 0$$

$x^T E y = 0$ E is t_x : essential matrix

bc of p^{\wedge} being normalized

$\tilde{p}^T E \tilde{p} = 0 \quad E = [t_x] R$

$E p^{\wedge}$ gives eq for ep line for o'

$E^T t p^{\wedge}$ gives ep in o

epipoles in nullspace of E: $E^T \tilde{e}' = 0$ and $E \tilde{e} = 0$
- Set: $F = K^{-1} E K^{-1}$ Then: $p^T F p = 0$ $F p$ $F^T p'$ are ep lines w p, p'

- F is estimate known as weak calibration
- $E = K'^T F K$
- from E calc relative rotation/translation
- scene point z direction
- disparity: $x - x'$, inversely proportional to depth

For each pixel

1. find corresponding ep line in right

2. search along and find best match (SSD)

3. tri matches to get depth

y points are the same

Stereo image rectification

- create virtual planes that only differ by translation
- \rightarrow pixel motion is horiz \rightarrow 2 homo for each img
- distance min, similarity max
 - sum(l - r)^2 vs

Triangulation

$$z = B^* f / (z - x')$$

- Method: finding correspondences
- transformations: translation, rotation, aspect, affine (change shape), perspective
 - $p' = T(p) = Mp$
 - when p is in $R^2 \rightarrow 2 \times 2 M$
- uniform scaling: same scalar for all components
- non uniform \rightarrow new aspect ratio

The diagram illustrates four types of 2D linear transformations, each with its corresponding transformation matrix and equations for the transformed coordinates x' and y' .

- Shear:**
 - Equations: $x' = x + sh_x \cdot y$, $y' = sh_y \cdot x + y$
 - Matrix: $\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & sh_x \\ 0 & 1 & sh_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$
- Rotate:**
 - Equations: $x' = -y$, $y' = x$
 - Matrix: $\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$
- Scale:**
 - Equations: $x' = s_x \cdot x$, $y' = s_y \cdot y$
 - Matrix: $\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$
- Shear (Skew):**
 - Equations: $x' = 1 \cdot x + sh_x \cdot y$, $y' = 1 \cdot y$
 - Matrix: $\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & sh_x & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$

- * parallel lines remain parallel
 - * feature based alignment
 - * eigenvectors are orthonormal
- $$\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ w \end{bmatrix}$$
- * symmetric A eigenvector w largest λ max $\frac{x^T A x}{x^T x}$ unit vectors max num
 - * non symmetric -> sVD: $U \Sigma V^T$
 - * u rotate eigenvectors of AA^t
 - * σ sqrt of λ 's

$$\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ w \end{bmatrix}$$

- solving least squares invert A find v (m,b) point -
 > minimizing error

- by argmin:
-
- Start with two points (x, y)
- $$y = A v$$
- $$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix}$$
- $$\|y - A v\|^2 = \left\| \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} m x_1 + b \\ m x_2 + b \end{bmatrix} \right\|^2$$
- $$= (y_1 - (m x_1 + b))^2 + (y_2 - (m x_2 + b))^2$$

- * satisfying:
Solution satisfies $(A^T A)v^* = A^T y$
or
$$v^* = (A^T A)^{-1} A^T y$$

- when A is square, rank < rows or rank < cols
- homo ls -> orthogonal to vectors
 $\text{argmin}(\|Av\|^2)$ smallest ev from AtA ,
 smallest singular v of A
- for 2 variables need at least 2 eq

ransac - reduce noise

- not true matches -> outlier (minimising lq creates error)
 - look for inliers

For N times

1. select random seed group s points
 1. more points = more robust
2. compute transformation for seed group
3. find inliers to this transformation
 1. point whose d is $< t$
4. if large then recompute estimate M on all of inliers
 1. $d > \text{inliers}$ accept and refit using all inliers

photometric: type, direction, intensity of light,
surfaces reflectance property


optical: focal length, fov, aperture, shutter speed

- image formation
 - every point on a tree blends its colors all across the film
 - -> barrier known as aperture (only one makes it through) flipped
 - f = focal length, c = center of camera
 - f distance from pinhole and img
 - lines all to vanishing point
- facts about projection
 - line in 3d -> line in 3d
 - parallel 3d -> loses, but converge to center point
 - distant objects are smaller
- increase aperture (when too dark)

Diagram illustrating the geometry of a camera model. It shows a 3D coordinate system with axes x , y , and z . A point $P(x, y, z)$ is projected onto the image plane ($z=0$) at point $p(x', y', 0)$. The image plane is at a distance f (focal length) from the optical center ($0, 0, 0$). The projection is shown as a ray from the optical center through P to the image plane. The diagram also shows the image plane as a circle with radius r . The image plane is labeled "Image plane" and the optical center is labeled "Optical center".

- extrinsic ass: no rotation, camera at 000 4th col
- intrinsic ass: opticalcenter000,unit aspe,!skew
- rotations:

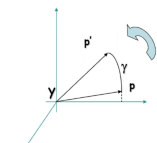
Rotation around the coordinate axes, **counter-clockwise**:

$$\mathbf{R} = \mathbf{R}_x(\alpha) \mathbf{R}_y(\beta) \mathbf{R}_z(\gamma)$$


$$\mathbf{R}_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix}$$

$$\mathbf{R}_y(\beta) = \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix}$$

$$\mathbf{R}_z(\gamma) = \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



$$w \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & s & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Stereo: given 2 cameras/correspondence

- find: world coordinate for known point

Epipolar Geometry

baseline connects both origins

- epipoles where baseline intersects w img planes
- epipolar lines btw point and $e \rightarrow p'$ (must be from same light ray)
 - rays given by ep line
- only translation \rightarrow ep lines horizontal, $e \infty$
- forward motion \rightarrow ep lines out from optical cent
- $F_e = 0$ $F^{\wedge} te' = 0$, $F = r, 2, 7$ deg freedom, !unique
- \rightarrow need at least 7 correspondences

eight point algo

Each point gives an equation:

$$[u u', u v', u, v u', v v', v, u', v', 1] \cdot [f_{11}, f_{12}, f_{13}, f_{21}, f_{22}, f_{23}, f_{31}, f_{32}, f_{33}] = 0$$

Stack equations to yield U:

[illegible]

How to solve for f (F unrolled)?:

$$\arg \min_{\|f\|=1} \|Uf\|_2^2 \rightarrow \text{Eigenvector of } U^T U \text{ with smallest eigenvalue}$$

- f output is not perfect (bc not perfectly aligned)
- closest matrix F with lower rank = take highs σ
 - multiply SVD out for new F

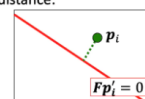
1. normalize img coordinates: center img data @ origin 0 > scale msd btw origina & data points = 2
2. apply 8-point algo: compute F
3. enforce rank-2 constraint
4. unnormalize coordinates: F back to og units
5. $T^T F T =$ og coords

Minimizing via UTU minimizes sum of squared *algebraic* distances between points \mathbf{p}_i and epipolar lines $\mathbf{F}\mathbf{p}'_i$ (or points \mathbf{p}'_i and epipolar lines $\mathbf{F}'\mathbf{p}_i$):

$$\sum_i (p_i^T F p'_i)^2$$

May want to minimize *geometric* distance:

$$\sum_i d(p_i, Fp'_i)^2 + d(p'_i, F^T p_i)^2$$



Process:

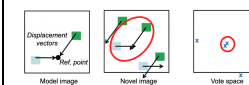
- $y = mx + b \rightarrow$ point based on parameters m and b
- point in $x, y \rightarrow b = -xm + y$ (line in hough space)
- intersection of lines in hough space = line that passes through two coordinates in image space
 > quantize and bin each grid spot for votes
- polar representation:

Diagram illustrating the geometric interpretation of the Hough transform. A point (x, y) in the image plane is shown. A line is drawn through this point, perpendicular to the x-axis at an angle θ . The perpendicular distance from the origin to this line is labeled d . The diagram shows that the coordinates x and y can be expressed as $x = d \cos(\theta)$ and $y = d \sin(\theta)$.

- point in img -> sinusoid segment in hough

Algorithm

1. $H[d, \theta] = 0\$$
 2. for each edge point $I[x, y]$
 1. for $\theta = [\theta_{min} to \theta_{max}]$
 2. $\theta = \text{gradient at } x, y$
 3. $x \cos(\theta) + y \sin(\theta) = d$
 4. $H[d, \theta] += 1$
 3. find (d, θ) where $H[\] = \text{max}$
 4. detect line given by d
- now circles instead of lines intersecting -> intersecting circles centered at points of hough space -> on edges of circle in img
 - unknown radius = cone in hough space
 - can use gradient angle of img space -> hough space line ray



- offline procedure: at each boundary point find $r = a - p_{-}(i)$ is a potential center, r vote
 - displacement vectors in table indexed by gradient orientation
 - look for overlaps in r

Tuning

- * same d and theta can't vote for multiple lines
- * still see peaks in noise voting space -> gradient descent
- * orientation
- * or give more votes to stronger edges
- * or snapping of d, theta to give more or less votes