▶Geometric Interpretation of Linear Rectification

What pair of physical cameras is compatible with $\mathbf{F}^*$?

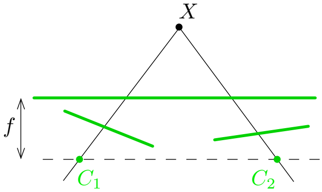
- we know that $\mathbf{F} = (\mathbf{Q}_1 \mathbf{Q}_2^{-1})^\top [\underline{\mathbf{e}}_1]_\times$ Slide 77
- we choose $\mathbf{Q}_1^* = \mathbf{K}_1^*$, $\mathbf{Q}_2^* = \mathbf{K}_2^* \mathbf{R}^*$; then
$$(\mathbf{Q}_1^* \mathbf{Q}_2^{*-1})^\top [\underline{\mathbf{e}}_1^*]_\times = (\mathbf{K}_1^* \mathbf{R}^{*\top} \mathbf{K}_2^{*-1})^\top \mathbf{F}^*$$
- we look for $\mathbf{R}^*$, $\mathbf{K}_1^*$, $\mathbf{K}_2^*$ compatible with
$$(\mathbf{K}_1^* \mathbf{R}^{*\top} \mathbf{K}_2^{*-1})^\top \mathbf{F}^* = \lambda \mathbf{F}^*, \qquad \mathbf{R}^* \mathbf{R}^{*\top} = \mathbf{I}, \qquad \mathbf{K}_1^*, \mathbf{K}_2^* \text{ upper triangular}$$
- we also want $\mathbf{b}^*$ from $\underline{\mathbf{e}}_1^* \simeq \mathbf{P}_1^* \underline{\mathbf{C}}_2^* = \mathbf{K}_1^* \mathbf{b}^*$ — $\mathbf{b}^*$ in cam. 1 frame
- result:
$$\mathbf{R}^* = \mathbf{I}, \quad \mathbf{b}^* = \begin{bmatrix} b \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{K}_1^* = \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{K}_2^* = \begin{bmatrix} k_{21} & k_{22} & k_{23} \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{29}$$

- rectified cameras are in canonical position with respect to each other — not rotated, canonical baseline
- rectified calibration matrices can differ in the first row only
- when $\mathbf{K}_1^* = \mathbf{K}_2^*$ then the rectified pair is called the standard stereo pair and the homographies standard rectification homographies

▶cont'd

- rectification is a homography (per image)
  - ⇒ rectified camera centers are equal to the original ones
- standard rectified cameras are in canonical orientation
  - ⇒ rectified image projection planes are coplanar
- standard rectification guarantees equal rectified calibration matrices
  - ⇒ rectified image projection planes are equal

standard rectification homographies reproject onto a common image plane parallel to the baseline



Corollary

- the standard rectified stereo pair has vanishing disparity for 3D points at infinity
  - but known $\mathbf{F}$ alone does not give any constraints to obtain standard rectification homographies
  - for that we need either of these:
    1. projection matrices, or
    2. calibrated cameras, or
    3. a few points at infinity calibrating k_{1i}, k_{2i}, $i = 1, 2, 3$ in (29)

▶Primitive Rectification

Goal: Given fundamental matrix $\mathbf{F}$, derive some simple rectification homographies $\mathbf{H}_1$, $\mathbf{H}_2$

1. Let the SVD of $\mathbf{F}$ be $\mathbf{U}\mathbf{D}\mathbf{V}^\top = \mathbf{F}$, where $\mathbf{D} = \mathrm{diag}(1, d^2, 0)$, $\quad 1 \geq d^2 > 0$
2. decompose $\mathbf{D} = \mathbf{A}^\top \mathbf{F}^* \mathbf{B}$, where ($\mathbf{F}^*$ is given → Slide 151)

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & d & 0 \\ 1 & 0 & 0 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & -d & 0 \end{bmatrix}$$

3. then

$$\mathbf{F} = \mathbf{U}\mathbf{D}\mathbf{V}^\top = \underbrace{\mathbf{U}\mathbf{A}^\top}_{\hat{\mathbf{H}}_2^\top} \mathbf{F}^* \underbrace{\mathbf{B}\mathbf{V}^\top}_{\hat{\mathbf{H}}_1}$$

   and the primitive rectification homographies are

$$\hat{\mathbf{H}}_2 = \mathbf{A}\mathbf{U}^\top, \qquad \hat{\mathbf{H}}_1 = \mathbf{B}\mathbf{V}^\top$$

⊛ P1; 1pt: derive some $\mathbf{A}$, $\mathbf{B}$ from the admissible class

- rectification homographies do exist
- there are other primitive rectification homographies, these suggested are just simple to obtain

▶Primitive Rectification Suffices for Calibrated Cameras

Obs: calibrated cameras: $d = 1 \Rightarrow \hat{\mathbf{H}}_1$, $\hat{\mathbf{H}}_2$ are orthogonal

1. determine primitive rectification homographies $(\hat{\mathbf{H}}_1, \hat{\mathbf{H}}_2)$ from the essential matrix
2. choose a suitable common calibration matrix $\mathbf{K}$, e.g.

$$\mathbf{K} = \begin{bmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad f = \frac{1}{2}(f^1 + f^2), \quad u_0 = \frac{1}{2}(u_0^1 + u_0^2), \quad \text{etc.}$$

3. the final rectification homographies are

$$\mathbf{H}_1 = \mathbf{K}\hat{\mathbf{H}}_1, \quad \mathbf{H}_2 = \mathbf{K}\hat{\mathbf{H}}_2$$

- we got a standard camera pair and non-negative disparity

$$\mathbf{P}_i^+ \stackrel{\text{def}}{=} \mathbf{K}_i^{-1}\mathbf{P}_i = \mathbf{R}_i \begin{bmatrix} \mathbf{I} & -\mathbf{C}_i \end{bmatrix}, \quad i = 1, 2$$

note we started from $\mathbf{E}$, not $\mathbf{F}$

$$\mathbf{H}_1\mathbf{P}_1^+ = \mathbf{K}\hat{\mathbf{H}}_1\mathbf{P}_1^+ = \mathbf{K}\underbrace{\mathbf{B}\mathbf{V}^\top\mathbf{R}_1}_{\mathbf{R}^*} \begin{bmatrix} \mathbf{I} & -\mathbf{C}_1 \end{bmatrix} = \mathbf{K}\mathbf{R}^* \begin{bmatrix} \mathbf{I} & -\mathbf{C}_1 \end{bmatrix}$$

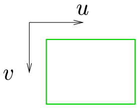
$$\mathbf{H}_2\mathbf{P}_2^+ = \mathbf{K}\hat{\mathbf{H}}_2\mathbf{P}_2^+ = \mathbf{K}\underbrace{\mathbf{A}\mathbf{U}^\top\mathbf{R}_2}_{\mathbf{R}^*} \begin{bmatrix} \mathbf{I} & -\mathbf{C}_2 \end{bmatrix} = \mathbf{K}\mathbf{R}^* \begin{bmatrix} \mathbf{I} & -\mathbf{C}_2 \end{bmatrix}$$

- one can prove that $\mathbf{B}\mathbf{V}^\top\mathbf{R}_1 = \mathbf{A}\mathbf{U}^\top\mathbf{R}_2$ with the help of (11)
- points at infinity project to $\mathbf{K}\mathbf{R}^*$ in both images $\Rightarrow$ they have zero disparity

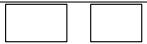
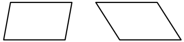
▶The Degrees of Freedom in Epipolar Rectification

Proposition 1 Homographies $\mathbf{A}_1$ and $\mathbf{A}_2$ are rectification-preserving if the images stay rectified, i.e. if $\mathbf{A}_2^{-\top}\,\mathbf{F}^*\,\mathbf{A}_1^{-1} \simeq \mathbf{F}^*$, which gives

$$\mathbf{A}_1 = \begin{bmatrix} l_1 & l_2 & l_3 \\ 0 & s_v & t_v \\ 0 & q & 1 \end{bmatrix}, \qquad \mathbf{A}_2 = \begin{bmatrix} r_1 & r_2 & r_3 \\ 0 & s_v & t_v \\ 0 & q & 1 \end{bmatrix},$$



where $s \neq 0,\ u_0,\ l_1,\ l_2 \neq 0,\ l_3,\ r_1,\ r_2 \neq 0,\ r_3,\ q$ are 9 free parameters.

general	transformation		canonical	type
$l_1,\ r_1$	horizontal scales		$l_1 = r_1$	algebraic
$l_2,\ r_2$	horizontal skews		$l_2 = r_2$	algebraic
$l_3,\ r_3$	horizontal shifts		$l_3 = r_3$	algebraic
q	common special projective			geometric
s_v	common vertical scale			geometric
t_v	common vertical shift			algebraic
9 DoF			$9 - 3 = 6$ DoF	

- q is rotation about the baseline
- s_v changes the focal length

proof: find a rotation $\mathbf{G}$ that brings $\mathbf{K}$ to upper triangular form via RQ decomposition: $\mathbf{A}_1\mathbf{K}_1^* = \hat{\mathbf{K}}_1\mathbf{G}$ and $\mathbf{A}_2\mathbf{K}_2^* = \hat{\mathbf{K}}_2\mathbf{G}$

The Rectification Group

Corollary for Proposition 1 Let $\bar{\mathbf{H}}_1$ and $\bar{\mathbf{H}}_2$ be (primitive or other) rectification homographies. Then $\mathbf{H}_1 = \mathbf{A}_1\bar{\mathbf{H}}_1, \quad \mathbf{H}_2 = \mathbf{A}_2\bar{\mathbf{H}}_2$ are also rectification homographies.

Proposition 2 Pairs of rectification-preserving homographies $(\mathbf{A}_1, \mathbf{A}_2)$ form a group with group operation $(\mathbf{A}_1', \mathbf{A}_2') \circ (\mathbf{A}_1, \mathbf{A}_2) = (\mathbf{A}_1'\,\mathbf{A}_1, \mathbf{A}_2'\,\mathbf{A}_2)$.

Proof:

- closure by Proposition 1
- associativity by matrix multiplication
- identity belongs to the set
- inverse element belongs to the set by $\mathbf{A}_2^\top\mathbf{F}^*\mathbf{A}_1 \simeq \mathbf{F}^* \Leftrightarrow \mathbf{F}^* \simeq \mathbf{A}_2^{-\top}\mathbf{F}^*\mathbf{A}_1^{-1}$

Optimal and Non-linear Rectification

Optimal choice for the free parameters

- by minimization of residual image distortion, eg.
  [Gluckman & Nayar 2001]

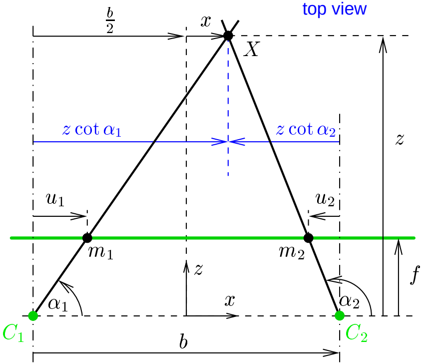
$$\mathbf{A}_1^* = \arg\min_{\mathbf{A}_1} \iint_{\Omega} \left(\det J(\mathbf{A}_1 \hat{\mathbf{H}}_1 \underline{\mathbf{x}}) - 1\right)^2 d\mathbf{x}$$

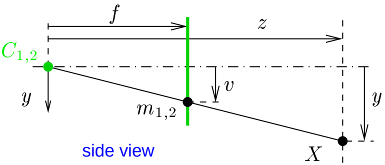
- by minimization of image information loss
  [Matoušek, ICIG 2004]
- non-linear rectification   suitable for forward motion
  [Pollefeys et al. 1999], [Geyer & Daniilidis 2003]

forward egomotion

rectified images, Pollefeys' method

▶Binocular Disparity in Standard Stereo Pair





- Assumptions: single image line, standard camera pair

$$b = z \cot \alpha_1 - z \cot \alpha_2$$

$$u_1 = f \cot \alpha_1 \qquad u_2 = f \cot \alpha_2$$

$$b = \frac{b}{2} + x - z \cot \alpha_2$$

$X = (x, z)$ from **disparity** $d = u_1 - u_2$:

$$\boxed{z = \frac{b\,f}{d}}, \quad x = \frac{b}{d}\,\frac{u_1 + u_2}{2}, \quad y = \frac{b\,v}{d}$$

f, d, u, v in pixels, b, x, y, z in meters

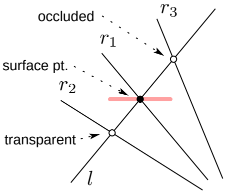
Observations

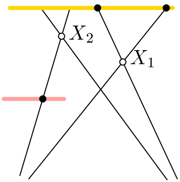
- constant disparity surface is a frontoparallel plane
- distant points have small disparity
- relative error in z is large for small disparity

$$\frac{1}{z}\,\frac{dz}{dd} = -\frac{1}{d}$$

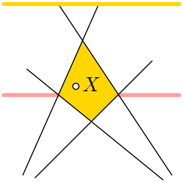
- increasing baseline increases disparity and reduces the error

▶Understanding Basic Occlusion Types





half occlusion

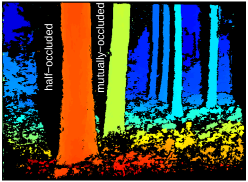


mutual occlusion

- surface point at the intersection of rays l and r_1 occludes a world point at the intersection (l, r_3) and implies the world point (l, r_2) is transparent, therefore

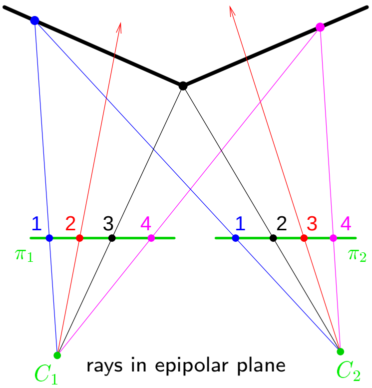
  (l, r_3) and (l, r_2) are <u>excluded</u> by (l, r_1)

- in half-occlusion, every world point such as X_1 or X_2 is excluded by a binocularly visible surface point $\Rightarrow$ decisions on correspondences <u>are not</u> independent
- in mutual occlusion this is no longer the case: any X in the yellow zone is <u>not excluded</u> $\Rightarrow$ decisions in the zone <u>are</u> independent on the rest

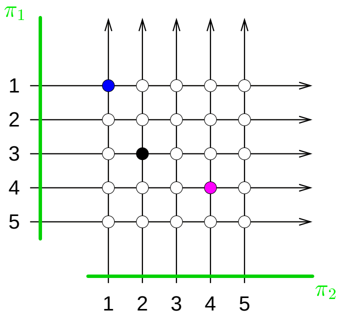


▸Matching Table

Based on the observation on mutual exclusion we expect each pixel to match at most once.



rays in epipolar plane



matching table T

matching table

- rows and columns represent optical rays
- nodes: possible correspondence pairs
- full nodes: correspondences
- numerical values associated with nodes: descriptor similarities

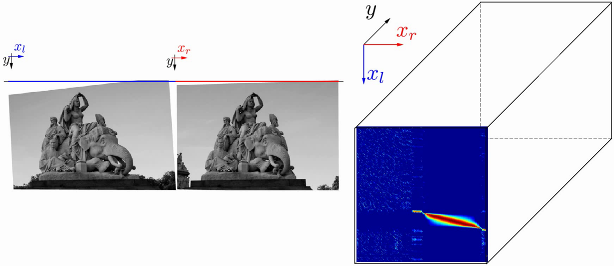
see next

Image Point Descriptors And Their Similarity

Descriptors: Tag image points by their (viewpoint-invariant) physical properties:

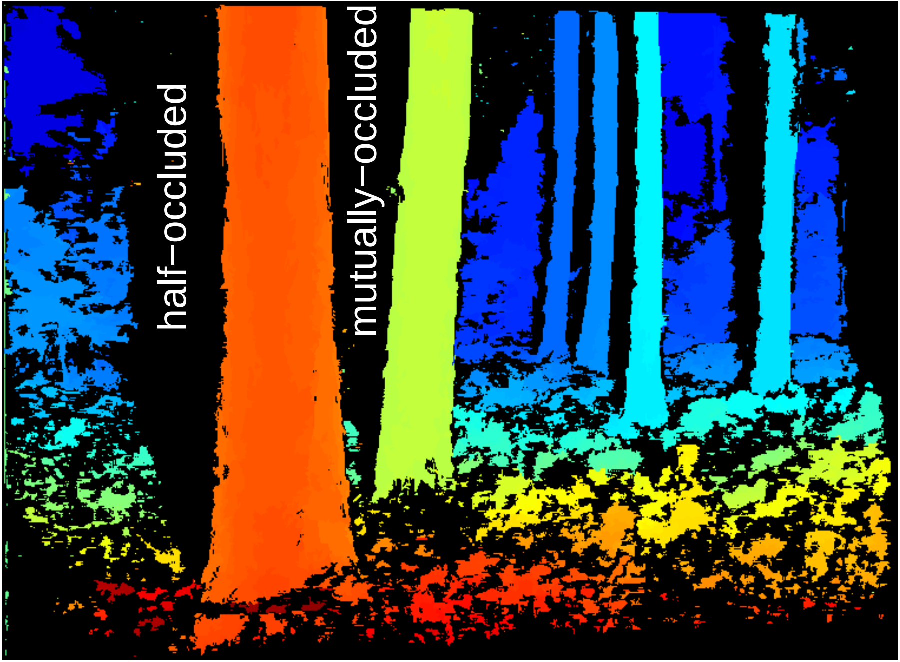
- texture window [Moravec 77]
- reflectance profile under a moving illuminant
- photometric ratios [Wolff & Angelopoulou 93-94]
- dual photometric stereo [Ikeuchi 87]
- polarization signature
- . . .

- similar points are more likely to match
- we will compute image similarity for all 'match candidates' and get the matching table



video

Thank You


half–occluded
mutually–occluded