Module VII

Stereovision

mostly covered by

Šára, R. How To Teach Stereoscopic Vision. Proc. ELMAR 2010 referenced as [SP]

additional references

C. Geyer and K. Daniilidis. Conformal rectification of omnidirectional stereo pairs. In *Proc Computer Vision and Pattern Recognition Workshop*, p. 73, 2003.

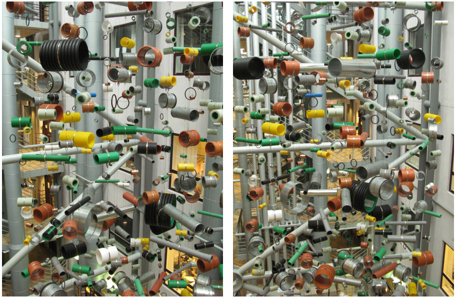
J. Gluckman and S. K. Nayar. Rectifying transformations that minimize resampling effects. In *Proc IEEE CS Conf on Computer Vision and Pattern Recognition*, vol. 1:111–117. 2001.

M. Pollefeys, R. Koch, and L. V. Gool. A simple and efficient rectification method for general motion. In *Proc Int Conf on Computer Vision*, vol. 1:496–501, 1999.

What Are The Relative Distances?

- monocular vision already gives a rough 3D sketch because we understand the scene

What Are The Relative Distances?

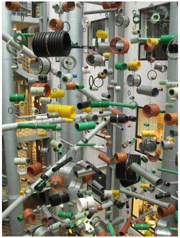


Centrum för teknikstudier at Malmö Högskola, Sweden

The Vyšehrad Fortress, Prague

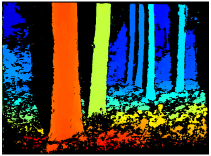
- left: we have no help from image interpretation
- right: ambiguous interpretation due to a combination of missing texture and occlusion

▶How Difficult Is Stereo?

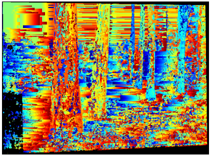


- when we do not recognize the scene and cannot use high-level constraints the problem seems difficult (right, less so in the center)
- most stereo matching algorithms do not require scene understanding prior to matching
- the success of a model-free stereo matching algorithm is unlikely:

left image



a good disparity map



disparity map from WTA

WTA Matching:

for every left-image pixel
find the most similar right-image pixel along the corresponding epipolar line [Marroquin 83]

A Summary of Our Observations and an Outlook

1. simple matching algorithms do not work
2. stereopsis requires image interpretation in sufficiently complex scenes

   or another-modality measurement

we have a tradeoff: model strength $\leftrightarrow$ universality

Outlook:

1. represent the occlusion constraint: correspondences are not independent due to occlusions
   - epipolar rectification
   - disparity
   - uniqueness as an occlusion constraint
2. represent piecewise continuity the weakest of interpretations; piecewise: object boundaries
   - ordering as a weak continuity model
3. use a consistent framework
   - looking for the most probable solution (MAP)

▶Linear Epipolar Rectification for Easier Correspondence Search

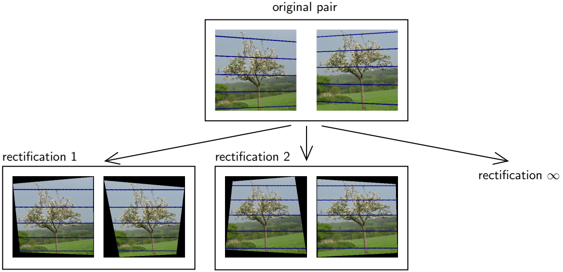
Obs:

- if we map epipoles to infinity, epipolar lines become parallel
- we then rotate them to become horizontal
- we then scale to make correspoding epipolar lines colinear
- this can be achieved by a pair of homographies applied to the images

Problem: Given fundamental matrix $\mathbf{F}$ or camera matrices $\mathbf{P}_1$, $\mathbf{P}_2$, compute a pair of homographies that maps epipolar lines to horizontal with the same row coordinate.

Procedure:

1. find a pair of rectification homographies $\mathbf{H}_1$ and $\mathbf{H}_2$.
2. warp images using $\mathbf{H}_1$ and $\mathbf{H}_2$ and transform the fundamental matrix $\mathbf{F} \mapsto \mathbf{H}_2^{-\top}\mathbf{F}\mathbf{H}_1^{-1}$ or the cameras $\mathbf{P}_1 \mapsto \mathbf{H}_1\mathbf{P}_1$, $\mathbf{P}_2 \mapsto \mathbf{H}_2\mathbf{P}_2$.

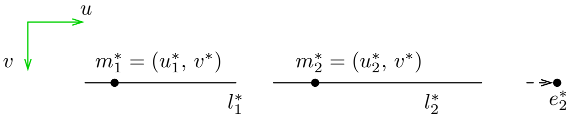


▶Rectification Homographies

Assumption: Cameras $(\mathbf{P}_1, \mathbf{P}_2)$ are rectified by a homography pair $(\mathbf{H}_1, \mathbf{H}_2)$:

$$\mathbf{P}_i^* = \mathbf{H}_i \mathbf{P}_i = \mathbf{H}_i \mathbf{K}_i \mathbf{R}_i \begin{bmatrix} \mathbf{I} & -\mathbf{C}_i \end{bmatrix}, \quad i = 1, 2$$

rectified entities: $\mathbf{F}^*$, $\underline{\mathbf{l}}_2^*$, $\underline{\mathbf{l}}_1^*$, etc:



- the rectified location difference $d = u_1^* - u_2^*$ is called disparity

corresponding epipolar lines must be:

1. parallel to image rows $\Rightarrow$ epipoles become $e_1^* = e_2^* = (1, 0, 0)$
2. equivalent $l_2^* = l_1^*$ $\Rightarrow$ (a) $\underline{\mathbf{l}}_2^* \simeq \underline{\mathbf{l}}_1^* \simeq \underline{\mathbf{e}}_1^* \times \underline{\mathbf{m}}_1 = [\underline{\mathbf{e}}_1^*]_\times \underline{\mathbf{m}}_1$, (b) $\underline{\mathbf{l}}_2^* \simeq \mathbf{F}^* \underline{\mathbf{m}}_1$

- therefore the canonical fundamental matrix is

$$\mathbf{F}^* \simeq \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

A two-step rectification procedure

1. find some pair of primitive rectification homographies $\hat{\mathbf{H}}_1$, $\hat{\mathbf{H}}_2$
2. upgrade to a pair of optimal rectification homographies while preserving $\mathbf{F}^*$

▶Geometric Interpretation of Linear Rectification

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

- we know that $\mathbf{F} = (\mathbf{Q}_1 \mathbf{Q}_2^{-1})^\top [\underline{\mathbf{e}}_1]_\times$ →78
- 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:

$$\mathbf{R}^* = \mathbf{I}, \quad \mathbf{b}^* = \begin{bmatrix} b \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{K}_1^* = \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{K}_2^* = \begin{bmatrix} k_{21} & k_{22} & k_{23} \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{33}$$

- rectified cameras are in canonical relative pose — 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 <u>standard stereo pair</u> and the homographies <u>standard rectification homographies</u>
- standard rectification homographies: points at infinity have zero disparity

$$\mathbf{P}_i^* \underline{\mathbf{X}}_\infty = \mathbf{K} \begin{bmatrix} \mathbf{I} & -\mathbf{C}_i \end{bmatrix} \underline{\mathbf{X}}_\infty = \mathbf{K} \mathbf{X}_\infty \qquad i = 1, 2$$

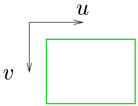
$K^* = K_1^* = K_2^*$

- this does not mean that the images are not distorted after rectification

▶The Degrees of Freedom in Epipolar Rectification

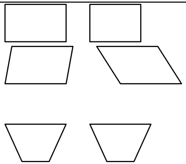
Proposition 1 Homographies $\mathbf{A}_1$ and $\mathbf{A}_2$ are rectification-preserving if the images stay rectified, i.e. if $\mathbf{A}_2^{-\top}\,\mathbf{F}^*\,\mathbf{A}_1^{-1} \simeq \mathbf{F}^*$, which gives

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



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

general	transformation	standard
$l_1,\ r_1$	horizontal scales	$l_1 = r_1$
$l_2,\ r_2$	horizontal shears	$l_2 = r_2$
$l_3,\ r_3$	horizontal shifts	$l_3 = r_3$
q	common special projective	
s_v	common vertical scale	
t_v	common vertical shift	
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 $\mathbf{A}_2^\top\,\mathbf{F}^*\mathbf{A}_1 \simeq \mathbf{F}^* \Leftrightarrow \mathbf{F}^* \simeq \mathbf{A}_2^{-\top}\mathbf{F}^*\mathbf{A}_1^{-1}$

▶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. Write $\mathbf{D}$ as $\mathbf{D} = \mathbf{A}^\top \mathbf{F}^* \mathbf{B}$ for some regular $\mathbf{A}$, $\mathbf{B}$. For instance ($\mathbf{F}^*$ is given →152)

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

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 other admissible $\mathbf{A}$, $\mathbf{B}$

- rectification homographies do exist →152
- 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 applied as $\mathbf{P}_i \mapsto \mathbf{H}_i \, \mathbf{P}_i$ are

$$\mathbf{H}_1 = \mathbf{K}\hat{\mathbf{H}}_1\mathbf{K}_1^{-1}, \quad \mathbf{H}_2 = \mathbf{K}\hat{\mathbf{H}}_2\mathbf{K}_2^{-1}$$

- we got a standard stereo pair (→153) and non-negative disparity

let $\mathbf{K}_i^{-1}\mathbf{P}_i = \mathbf{R}_i \begin{bmatrix} \mathbf{I} & -\mathbf{C}_i \end{bmatrix}, \quad i = 1, 2$ note we started from $\mathbf{E}$, not $\mathbf{F}$

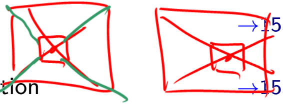
$$\mathbf{H}_1\mathbf{P}_1 = \mathbf{K}\hat{\mathbf{H}}_1\mathbf{K}_1^{-1}\mathbf{P}_1 = \mathbf{K}\underbrace{\mathbf{B}\mathbf{V}^\top\mathbf{R}_1}_{\mathbf{R}^*}\begin{bmatrix} \mathbf{I} & -\mathbf{C}_1 \end{bmatrix} = \mathbf{K}\mathbf{R}^*\begin{bmatrix} \mathbf{I} & -\mathbf{C}_1 \end{bmatrix}$$

$$\mathbf{H}_2\mathbf{P}_2 = \mathbf{K}\hat{\mathbf{H}}_2\mathbf{K}_2^{-1}\mathbf{P}_2 = \mathbf{K}\underbrace{\mathbf{A}\mathbf{U}^\top\mathbf{R}_2}_{\mathbf{R}^*}\begin{bmatrix} \mathbf{I} & -\mathbf{C}_2 \end{bmatrix} = \mathbf{K}\mathbf{R}^*\begin{bmatrix} \mathbf{I} & -\mathbf{C}_2 \end{bmatrix}$$

- one can prove that $\mathbf{B}\mathbf{V}^\top\mathbf{R}_1 = \mathbf{A}\mathbf{U}^\top\mathbf{R}_2$ with the help of essential matrix decomposition (13)
- points at infinity project to $\mathbf{K}\mathbf{R}^*$ in both images ⇒ they have zero disparity →160

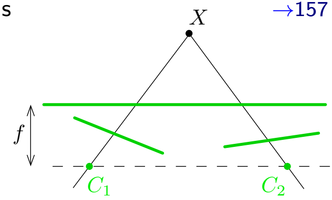
▶Summary & Remarks: Linear Rectification

- rectification is done with a pair of homographies (one per image) →151
  - ⇒ rectified camera centers are equal to the original ones
  - binocular rectification: a 9-parameter family of rectification homographies
  - trinocular rectification: has 9 or 6 free parameters (depending on additional constrains)
  - in general, linear rectification is not possible for more than three cameras
- rectified cameras are in canonical orientation →153
  - ⇒ rectified image projection planes are coplanar



- equal rectified calibration matrices give standard rectification →153
  - ⇒ rectified image projection planes are equal
- primitive rectification is standard in calibrated cameras →157

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



Corollary

- standard rectified pair: disparity vanishes when corresponding 3D points are at infinity
  - known $\mathbf{F}$ used alone gives no constraints on standard rectification homographies
  - for that we need either of these:
    1. projection matrices, or calibrated cameras, or
    2. a few points at infinity calibrating k_{1i}, k_{2i}, $i = 1, 2, 3$ in (33)

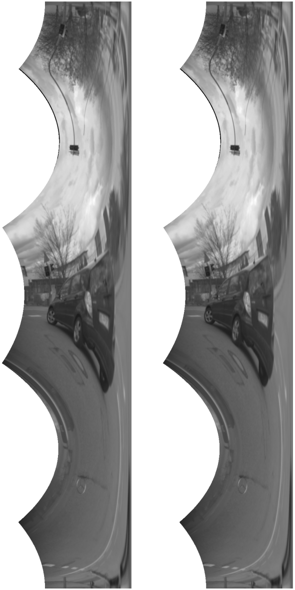
Optimal choice for the free parameters

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

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

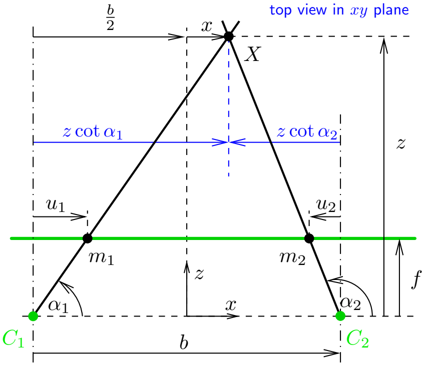
- by minimization of image information loss [Matoušek, ICIG 2004]
- non-linear rectification suitable for forward motion
  non-parametric: [Pollefeys et al. 1999]
  analytic: [Geyer & Daniilidis 2003]

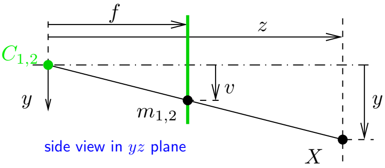
forward egomotion



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** $d = u_1 - u_2$:

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

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

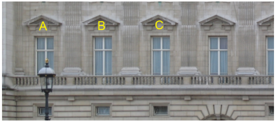
$$\frac{1}{z}\,\frac{dz}{dd} = -\frac{1}{d}$$

- increasing the baseline or the focal length increases disparity and reduces the error

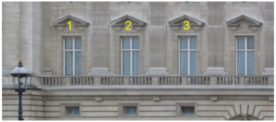
Structural Ambiguity in Stereovision

- we can recognize matches but have no scene model
- lack of an occlusion model
- lack of a continuity model

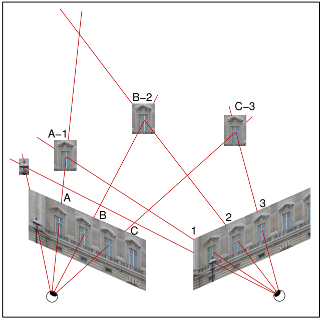
⇒ structural ambiguity in the presence of repetitions (or lack of texture)



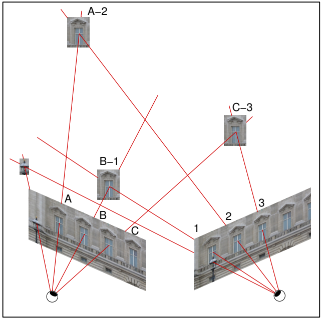
left image



right image



interpretation 1



interpretation 2

Thank You