

► Gauge Freedom

1. The external frame is not fixed:

See Projective Reconstruction Theorem →130

$$\underline{\mathbf{m}}_i \simeq \mathbf{P}_j \underline{\mathbf{X}}_i = \mathbf{P}_j \mathbf{H}^{-1} \mathbf{H} \underline{\mathbf{X}}_i = \mathbf{P}'_j \underline{\mathbf{X}}'_i$$

2. Some representations are not minimal, e.g.

- \mathbf{P} is 12 numbers for 11 parameters
- we may represent \mathbf{P} in decomposed form $\mathbf{K}, \mathbf{R}, \mathbf{t}$
- but \mathbf{R} is 9 numbers representing the 3 parameters of rotation

As a result

- there is no unique solution
- matrix $\sum_r \mathbf{L}_r^\top \mathbf{L}_r$ is singular

Solutions

1. fixing the external frame (e.g. a selected camera frame) explicitly or by constraints
- 2a. either imposing constraints on projective entities

- cameras, e.g. $\mathbf{P}_{3,4} = 1$
- points, e.g. $\|\underline{\mathbf{X}}_i\|^2 = 1$

$$\mathbf{R}\mathbf{R}^\top = \mathbf{I}$$

this excludes affine cameras
this way we can represent points at infinity

- 2b. or using minimal representations $\det \mathbf{R} = +1$

- points in their Euclidean representation $\underline{\mathbf{X}}_i$ but finite points may be an unrealistic model
- rotation matrix can be represented by axis-angle or the Cayley transform see next

Implementing Simple Constraints

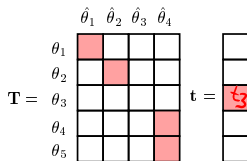
What for?

1. fixing external frame as in $\theta_i = \mathbf{t}_i$ 'trivial gauge'
2. representing additional knowledge as in $\theta_i = \theta_j$ e.g. cameras share calibration matrix \mathbf{K}

Introduce reduced parameters $\hat{\theta}$ and replication matrix \mathbf{T} :

$$\theta = \mathbf{T} \hat{\theta} + \mathbf{t}, \quad \mathbf{T} \in \mathbb{R}^{p, \hat{p}}, \quad \hat{p} \leq p$$

then \mathbf{L}_r in LM changes to $\mathbf{L}_r \mathbf{T}$ and everything else stays the same $\rightarrow 107$



these \mathbf{T} , \mathbf{t} represent

$\theta_1 = \hat{\theta}_1$	no change
$\theta_2 = \hat{\theta}_2$	no change
$\theta_3 = t_3$	constancy
$\theta_4 = \theta_5 = \hat{\theta}_4$	equality

- \mathbf{T} deletes columns of \mathbf{L}_r that correspond to fixed parameters it reduces the problem size
- consistent initialisation: $\theta^0 = \mathbf{T} \hat{\theta}^0 + \mathbf{t}$ or filter the init by pseudoinverse $\theta^0 \mapsto \mathbf{T}^\dagger \theta^0$
- no need for computing derivatives for θ_j corresponding to all-zero rows of \mathbf{T} fixed θ
- constraining projective entities $\rightarrow 145-146$
- more complex constraints tend to make normal equations dense
- implementing constraints is safer than explicit renaming of the parameters, gives a flexibility to experiment
- other methods are much more involved, see [Triggs et al. 1999]
- BA resource: <http://www.ics.forth.gr/~lourakis/sba/> [Lourakis 2009]

Matrix Exponential

- for any square matrix we define

$$\text{expm } \mathbf{A} = \sum_{k=0}^{\infty} \frac{1}{k!} \mathbf{A}^k$$

$e^{\mathbf{A}}$

note: $\mathbf{A}^0 = \mathbf{I}$

- some properties

$$\text{expm } \mathbf{0} = \mathbf{I}, \quad \text{expm}(-\mathbf{A}) = (\text{expm } \mathbf{A})^{-1}, \quad \text{expm}(\mathbf{A} + \mathbf{B}) \neq \text{expm}(\mathbf{A}) \text{expm}(\mathbf{B})$$

$$\text{expm}(\mathbf{A}^{\top}) = (\text{expm } \mathbf{A})^{\top} \text{ hence if } \mathbf{A} \text{ is skew symmetric then } \text{expm } \mathbf{A} \text{ is orthogonal:}$$

$$e^{[\mathbf{a}]_{\times}} = \mathbf{R} \quad (\text{expm}(\mathbf{A}))^{\top} = \text{expm}(\mathbf{A}^{\top}) = \text{expm}(-\mathbf{A}) = (\text{expm}(\mathbf{A}))^{-1}$$

$$\text{expm}(a \mathbf{A}) \text{expm}(b \mathbf{A}) = \text{expm}((a + b)\mathbf{A}), \quad \underline{\det \text{expm } \mathbf{A} = \text{expm}(\text{tr } \mathbf{A})}$$

Ex:

- homography can be represented via exponential map with 8 numbers e.g. as

$$\mathbf{H} = \text{expm } \mathbf{Z} \quad \text{such that} \quad \text{tr } \mathbf{Z} = 0, \quad \text{eg. } \mathbf{Z} = \begin{bmatrix} z_{11} & z_{12} & z_{13} \\ z_{21} & z_{22} & z_{23} \\ z_{31} & z_{32} & -(z_{11} + z_{22}) \end{bmatrix}$$

► Minimal Representations for Rotation

- \mathbf{o} – rotation axis, $\|\mathbf{o}\| = 1$, φ – rotation angle
- **wanted**: simple mapping to/from rotation matrices

1. Matrix exponential. Let $\boldsymbol{\omega} = \varphi \mathbf{o}$, $0 \leq \varphi < \pi$, then

$$\varphi = \|\boldsymbol{\omega}\|$$

$$\mathbf{R} = \expm[\boldsymbol{\omega}]_{\times} = \sum_{n=0}^{\infty} \frac{[\boldsymbol{\omega}]_{\times}^n}{n!} = \cdots \stackrel{\textcircled{*}}{=} \mathbf{I} + \frac{\sin \varphi}{\varphi} [\boldsymbol{\omega}]_{\times} + \frac{1 - \cos \varphi}{\varphi^2} [\boldsymbol{\omega}]_{\times}^2$$

- for $\varphi = 0$ we take the limit and $\mathbf{R} = \mathbf{I}$
- this is the Rodrigues' formula for rotation
- inverse (the principal logarithm of \mathbf{R}) from

$$0 \leq \varphi < \pi, \quad \cos \varphi = \frac{1}{2}(\text{tr } \mathbf{R} - 1), \quad [\boldsymbol{\omega}]_{\times} = \frac{\varphi}{2 \sin \varphi} (\mathbf{R} - \mathbf{R}^{\top}),$$

2. Cayley's representation; let $\mathbf{a} = \mathbf{o} \tan \frac{\varphi}{2}$, then

$$\mathbf{R} = (\mathbf{I} + [\mathbf{a}]_{\times})(\mathbf{I} - [\mathbf{a}]_{\times})^{-1}, \quad [\mathbf{a}]_{\times} = (\mathbf{R} + \mathbf{I})^{-1}(\mathbf{R} - \mathbf{I})$$

$$\mathbf{a}_1 \circ \mathbf{a}_2 = \frac{\mathbf{a}_1 + \mathbf{a}_2 - \mathbf{a}_1 \times \mathbf{a}_2}{1 - \mathbf{a}_1^{\top} \mathbf{a}_2} \quad \text{composition of rotations } \mathbf{R} = \mathbf{R}_1 \mathbf{R}_2$$

- again, cannot represent rotations for $\phi \geq \pi$
- no trigonometric functions
- explicit composition formula

► Minimal Representations for Other Entities

with the help of rotation we can minimally represent

1. fundamental matrix

$$\mathbf{F} = \mathbf{U}\mathbf{D}\mathbf{V}^\top, \quad \mathbf{D} = \text{diag}(1, d^2, 0), \quad \mathbf{U}, \mathbf{V} \text{ are rotations,} \quad 3 + 1 + 3 = 7 \text{ DOF}$$

2. essential matrix

$$\mathbf{E} = [-\mathbf{t}]_{\times} \mathbf{R}, \quad \mathbf{R} \text{ is rotation,} \quad \|\mathbf{t}\| = 1, \quad 3 + 2 = 5 \text{ DOF}$$

3. camera

$$\mathbf{P} = \mathbf{K} [\mathbf{R} \quad \mathbf{t}], \quad 5 + 3 + 3 = 11 \text{ DOF}$$

Interestingly, let

[Eade 2017]

$$\mathbf{B} = \begin{bmatrix} [\boldsymbol{\omega}]_{\times} & \mathbf{u} \\ \mathbf{0}^\top & 0 \end{bmatrix}, \quad \mathbf{B} \in \mathbb{R}^{4,4}$$

then, assuming $\|\boldsymbol{\omega}\| = \phi > 0$

for $\phi = 0$ we take the limits

$$\begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix} = \text{expm } \mathbf{B} = \mathbf{I}_4 + \mathbf{B} + h_2(\phi) \mathbf{B}^2 + h_3(\phi) \mathbf{B}^3 = \begin{bmatrix} \text{expm} [\boldsymbol{\omega}]_{\times} & \mathbf{V} \mathbf{u} \\ \mathbf{0}^\top & 1 \end{bmatrix}$$

$$\mathbf{V} = \mathbf{I}_3 + h_2(\phi) [\boldsymbol{\omega}]_{\times} + h_3(\phi) [\boldsymbol{\omega}]_{\times}^2, \quad \mathbf{V}^{-1} = \mathbf{I}_3 - \frac{1}{2} [\boldsymbol{\omega}]_{\times} + h_4(\phi) [\boldsymbol{\omega}]_{\times}^2$$

$$h_1(\phi) = \frac{\sin \phi}{\phi}, \quad h_2(\phi) = \frac{1 - \cos \phi}{\phi^2}, \quad h_3(\phi) = \frac{\phi - \sin \phi}{\phi^3}, \quad h_4(\phi) = \frac{1}{\phi^2} \left(1 - \frac{1}{2} \phi \cot \frac{\phi}{2} \right)$$

Stereovision

- 7.1 Introduction
- 7.2 Epipolar Rectification
- 7.3 Binocular Disparity and Matching Table
- 7.4 Image Similarity
- 7.5 Marroquin's Winner Take All Algorithm
- 7.6 Maximum Likelihood Matching
- 7.7 Uniqueness and Ordering as Occlusion Models

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

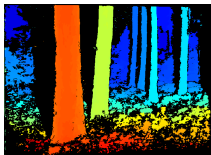
► How Difficult Is Stereo?



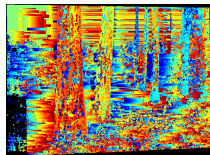
- when we do not recognize the scene and cannot use high-level constraints the problem seems difficult (right, less so in the center)
- most stereo matching algorithms do not require scene understanding prior to matching
- the success of a model-free stereo matching algorithm is unlikely:



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)



Thank You



