

# Optimisation on Manifolds

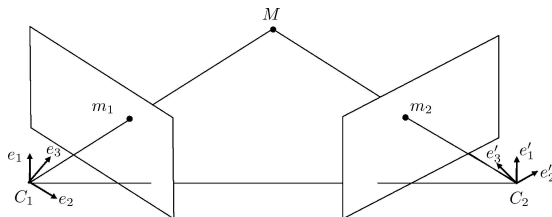
K. Hüper

MPI Tübingen & Univ. Würzburg

# Contents

- 2 Examples
  - Essential matrix estimation
  - Stereo matching without correspondence
- Which manifolds are involved?
- Parameterizations
- Newton-type methods
- Algorithm as differentiable map
- Convergence properties

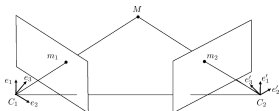
# Essential Matrix Estimation



**Assumption:** Identical pin hole cameras.

**Task:** Recover Euclidean transformation between two cameras.

# Epipolar Constraint



- Camera centers  $C_1, C_2 \in \mathbb{R}^3$ .
- Attached frames  $F_1 = \{e_1, e_2, e_3\}$  and  $F_2 = \{e'_1, e'_2, e'_3\}$ .
- Euclidean transformation  $(R, t) : F_1 \mapsto F_2$ .
- Let  $M_i = [X_i, Y_i, Z_i]^T$ ,  $i = 1, 2$  denote coords. of  $M$  w.r.t.  $F_i$ . Then

$$M_1 = RM_2 + t.$$

- Camera image points (pixel coords)  $m_i = [u_i, v_i, 1]^T$ . Then

$$m_1^T \hat{t} R m_2 = 0 \quad \text{(Epipolar Constraint)}$$

- **Essential matrix**  $E := \hat{t}R$ , with  $R$  orthonormal,  $\det R = 1$  and skew-symmetric  $\hat{t} = -\hat{t}^T$  with  $\|t\| = 1$ .

# Essential Matrix Estimation

**Facts:** The set of essential matrices  $\mathcal{E}$  is

- a compact connected 5-dimensional manifold diffeomorphic to

$$\mathbb{RP}^2 \times SO_3 \cong \left\{ E \in \mathbb{R}^{3 \times 3} \mid E = U \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} V^T; U, V \in SO_3 \right\},$$

(an orbit! “SVD action”),

- NOT a vector space,
- NOT easily described by equality constraints without redundancy,
- NOT diffeomorphic to the product of Stiefel manifolds  $S^2 \times SO_3$ .
- a reductive homogeneous space diffeomorphic to  $SO_3 \times SO_3 / O_2$

# Essential Matrix Estimation

**Task:** Given a set of  $N$  point correspondences  $(m_i, m'_i)_{i=1}^N$ , estimate  $E$  encapsulating the relative pose.

## Ideas:

- In the noise free case every pair has to fulfil the epipolar constraint  $m_i^\top \hat{t} R m'_i = 0$ .
- Choose suitable cost to be minimised over  $\mathcal{E}$ . Simplest choice (least squares) is

$$f : \mathcal{E} \rightarrow \mathbb{R}, \quad E \mapsto \sum_{i=1}^N (m_i^\top E m'_i)^2.$$

- Global minimisation of  $f$  over  $\mathcal{E}$ .
- Choose suitable family of parameterisations for  $\mathcal{E}$  ( $\dim \mathcal{E} = 5$ ).

# Tangent space of $\mathcal{E}$

**Fact:** The tangent space at the essential matrix  $E = UE_0V^T$  is

$$T_E\mathcal{E} = \left\{ U(\Omega E_0 - E_0\Psi)V^T \mid \Omega, \Psi \in \mathfrak{so}_3 \right\}$$
$$= \left\{ U \begin{bmatrix} 0 & \omega_{12} - \psi_{12} & -\psi_{13} \\ \psi_{12} - \omega_{12} & 0 & -\psi_{23} \\ -\omega_{13} & -\omega_{23} & 0 \end{bmatrix} V^T \mid \omega_{ij}, \psi_{ij} \in \mathbb{R}, i, j \in \{1, 2, 3\} \right\}$$

with  $\Omega = (\omega_{ij})$  and  $\Psi = (\psi_{ij})$ .

# Useful Parameterisations for $\mathcal{E}$

Let  $E_0 := \begin{bmatrix} 1 & & \\ & 1 & \\ & & 0 \end{bmatrix}$  and

$$\Omega_1(x) = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}, \quad \Omega_2(x) = \begin{bmatrix} 0 & x_3 & x_5 \\ -x_3 & 0 & -x_4 \\ -x_5 & x_4 & 0 \end{bmatrix}.$$

$$\mu_{(U,V)}^{\text{exp}} : \mathbb{R}^5 \rightarrow \mathcal{E}, \quad x \mapsto U e^{\Omega_1(x)} E_0 e^{-\Omega_2(x)} V^\top,$$

$$\mu_{(U,V)}^{\text{GS}} : \mathbb{R}^5 \rightarrow \mathcal{E}, \quad x \mapsto U \left( I + \Omega_1(x) \right)_{\text{GS}} E_0 \left( I - \Omega_2(x) \right)_{\text{GS}} V^\top,$$

$$\mu_{(U,V)}^{\text{SVD}} : \mathbb{R}^5 \rightarrow \mathcal{E}, \quad x \mapsto U \left( E_0 + \Omega_1(x) E_0 - E_0 \Omega_2(x) \right)_{\text{SVD}} V^\top.$$

## Gram-Schmidt:

Let  $X = QR$  unique QR-dec of invertible  $X$ . Then  $X_{\text{GS}} := Q$ .

## SVD:

Let  $X = \hat{U} \Sigma \hat{V}^\top$  ordered SVD, then  $X_{\text{SVD}} := \hat{U} E_0 \hat{V}^\top$ .

# Algorithm

Let

$$\mu : \mathcal{E} \times \mathbb{R}^5 \rightarrow \mathcal{E}, \quad \mu(E, 0) = \mu_{(U,V)}(0) = E.$$

With the gradient of  $f \circ \mu_{(U,V)}$  at 0

$$\nabla_{f \circ \mu(E, \cdot)}(0),$$

and the Hessian of  $f \circ \mu_{(U,V)}$  at 0

$$H_{f \circ \mu(E, \cdot)}(0).$$

The algorithmic map to be iterated is then:

$$s : \mathcal{E} \rightarrow \mathcal{E},$$
$$s(E) = \mu \left( E, - (H_{f \circ \mu(E, \cdot)}(0))^{-1} \cdot \nabla_{f \circ \mu(E, \cdot)}(0) \right)$$

# Discussion

- For GS or SVD based parameterizations  $D f(E) = 0 \iff E$  is fixed point of algorithm.
- For EXP this is not true due to finite injectivity radius.
- Locally all three algorithms are smooth maps  $\mathcal{E} \rightarrow \mathcal{E}$  around minima with nondegenerate Hessian.
- For all three algorithms:  
 $D s(E_*) = 0 \implies \|s(E_k - E_*)\| \leq \sup \|D^2 s(\bar{E})\| \cdot \|E_k - E_*\|^2$ .  
That is, we have local quadratic convergence.

# Discussion

- For GS or SVD based parameterizations  $D f(E) = 0 \iff E$  is fixed point of algorithm.
- For EXP this is not true due to finite injectivity radius.
- Locally all three algorithms are smooth maps  $\mathcal{E} \rightarrow \mathcal{E}$  around minima with nondegenerate Hessian.
- For all three algorithms:

$$D s(E_*) = 0 \implies \|s(E_k - E_*)\| \leq \sup \|D^2 s(\bar{E})\| \cdot \|E_k - E_*\|^2.$$

That is, we have local quadratic convergence.

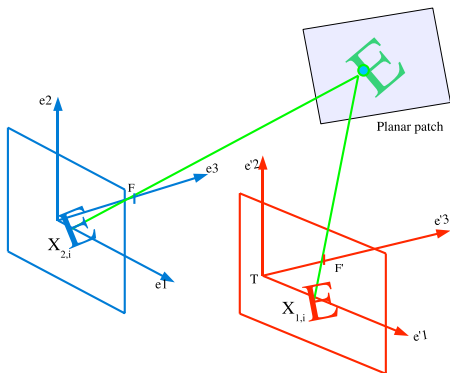
## Some details (we might discuss off-line)

- Why are our **algorithms well defined**? As there exist orthogonal  $U_1 \neq U_2$  and  $V_1 \neq V_2$  with  $\mu(U_1, V_1) = \mu(U_2, V_2)$ .
- Why is  $\mathcal{E}$  not **globally** diffeomorphic to a product of Stiefel manifolds? Any consequences?
- Are our algorithms **intrinsic Newton** methods, i.e. can we find a corresponding Riemannian metric?

# References

- **Multiple View Geometry in Computer Vision** by R. Hartley, A. Zisserman; Cambridge University Press 2004.
- **An Invitation to 3-D Vision** by Y. Ma *et al.*; Springer 2005.
- **Essential Matrix Estimation Using Gauss-Newton Iterations on a Manifold** by U. Helmke *et al.*; Int. J. of Computer Vision 74(2), 2007.
- **Optimization Criteria and Geometric Algorithms for Motion and Structure Estimation** by Y. Ma *et al.*; Int. J. of Computer Vision 44(3), 2001.
- **Pose Estimation of a Moving Humanoid Using Gauss-Newton Optimization on a Manifold** by M. Sarkis *et al.*; Humanoids 2007.

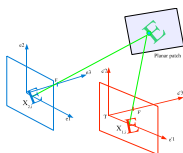
# Stereo Matching



Two coplanar cameras observe a planar patch.

**Task:** Recover Euclidean transformation between two cameras.

# Stereo Matching



**Assumption:** Two sets of image points  $\{X_{1,i}\}$ , and  $\{X_{2,i}\}$ ,  $X_{1,i}, X_{2,i} \in \mathbb{R}^3$  are unordered, i.e. the **pointwise correspondence** between both sets is **unknown**. The only available information then is the Euclidian displacement  $(R, \tau)$ ,  $R \in SO_3$  and  $\tau \in \mathbb{R}^3$  between the cameras.

## Typical tasks:

- (i) recover the geometry of the observed patch from the two images,
- (ii) establish a pointwise correspondence of both sets of image points,
- (iii) find a homographic transformation from the points of one image to the points of the other.

# Stereo Matching without Correspondence

**Idea (Zhou/Ghosh CDC 1996):** From normalized image points form the Gramians  $N, Q \in \mathbb{R}^{n \times n}$

$$N = \frac{1}{k} \sum_{i=1}^k X_{1,i} X_{1,i}^T, \quad Q = \frac{1}{k} \sum_{i=1}^k X_{2,i} X_{2,i}^T,$$

and find a transformation  $A$  via following a gradient flow which minimises the cost  $\|Q - ANA^T\|^2$ .

## Questions:

- What kind of transformations are the  $A$  matrices?
- What is the structure of the set of critical points?
- Can one do better than following a gradient flow?
- Are there closed form solutions available in the noise free case?

# Stereo Matching without Correspondence

## Fact:

The set of  $A$ -matrices forms a noncompact 3-dim. Lie group

$$G = \left\{ I_3 + \mathbf{e}_1 \mathbf{a}^\top \in \mathbb{R}^{3 \times 3} \mid 1 + \mathbf{e}_1^\top \mathbf{a} > 0, \mathbf{a} \in \mathbb{R}^3 \right\}.$$

with Lie algebra

$$\mathfrak{g} := \left\{ \mathbf{e}_1 \mathbf{b}^\top \mid \mathbf{b} \in \mathbb{R}^3 \right\}$$

and Lie bracket the matrix commutator.

# Stereo Matching without Correspondence

## Parameterization:

By exponentiating Lie algebra elements we obtain for any  $A \in G$  the parameterization map

$$\nu: \mathbb{R}^3 \rightarrow G, \quad \nu(b) := \exp(\mathbf{e}_1 b^\top) = I_3 + h(\mathbf{e}_1^\top b) \mathbf{e}_1 b^\top \quad (1)$$

with

$$h(b_1) = \begin{cases} \frac{e^{b_1} - 1}{b_1} & b_1 \neq 0 \\ 1 & b_1 = 0 \end{cases} . \quad (2)$$

Note that  $\nu$  satisfies  $\nu(0) = I_3$  and  $\nu$  defines a **global diffeomorphism onto the group  $G$** .

# Stereo Matching without Correspondence

## For convenience:

### Lemma

Given an  $(n \times n)$ -matrix  $N = N^\top > 0$  and let  $M = \{ANA^\top \mid A \in G\}$ . Then  $M$  is a smooth and connected 3-dimensional manifold. The map

$$\phi : G \rightarrow M, \quad \phi(A) := ANA^\top$$

is a global diffeomorphism. The tangent space of  $M$  at  $X \in M$  is  $T_X M = \{BX + XB^\top \mid B \in \mathfrak{g}\}$

Correspondingly, we obtain a family of global parameterizations of the manifold  $M$  as

$$\mu_X : \mathbb{R}^3 \rightarrow M, \quad \mu_X(\mathbf{b}) := \mathbf{e}^{\mathbf{e}_1 \mathbf{b}^\top} X (\mathbf{e}^{\mathbf{e}_1 \mathbf{b}^\top})^\top.$$

Thus  $\mu_X$  satisfies  $\mu_X(\mathbf{0}) = X$  and  $\mu_X$  defines a global diffeomorphism onto the manifold  $M$ .

# Stereo Matching without Correspondence

## Lemma

*Let  $N = N^\top$  be positive definite. The function  $f : M \rightarrow \mathbb{R}$ ,  $f(X) = \|Q - X\|^2$  with  $f(X) = \|Q - X\|^2$  has a unique critical point  $X_c \in M$ . The critical point  $X_c$  is characterized by the property that the first column coincides with that of  $Q$ .*

## Idea:

Minimise  $f$  over  $M$  via Newton-on-manifold approach to find the unique global minimum.

# Gradient and Hessian

$$\begin{aligned}\nabla(f \circ \mu_X)(0) &= 4X(X - Q)e_1, \\ \mathbf{H}_{f \circ \mu_X}(0) &= 4(X^2 + Xe_1e_1^\top X + e_1^\top(X - Q)e_1X) \\ &\quad + \frac{1}{2}(X(X - Q)e_1e_1^\top + e_1e_1^\top(X - Q)X).\end{aligned}$$

Note that the Hessian at the unique critical point  $X_c$  simplifies to

$$\mathbf{H}_{f \circ \mu_{X_c}}(0) = 4 \left( X_c^2 + X_c e_1 e_1^\top X_c \right)$$

# Newton's Method

Iterate the map

$$s : M \rightarrow M.$$

Let  $x^{\text{opt}}(X)$  denote the solution of

$$\widehat{H}_{f \circ \mu_X}(0) x = -\nabla(f \circ \mu_X)(0),$$

where for any  $X \in M$

$$\widehat{H}_{f \circ \mu_X}(0) = 4 \left( X^2 + X e_1 e_1^\top X \right).$$

Thus

$$x^{\text{opt}}(X) = X^{-1} \left( I_n - \frac{1}{2} e_1 e_1^\top \right) (Q - X) e_1$$

is well-defined for any  $X \in M$ . The algorithmic map  $s$  is given as

$$s(X) = \mu_X \left( x^{\text{opt}}(X) \right).$$

# Convergence Properties

## Lemma

*The algorithm  $s(X) = \mu_X(x^{\text{opt}}(X))$  converges locally quadratically fast.*

## Proof:

The mapping is smooth. The first derivative of the mapping at the critical point is equal to zero.

# Alternative Approach: Cholesky Factorisation

Unique Cholesky factors

$$N = U_N U_N^\top, \quad Q = U_Q U_Q^\top$$

with upper triangular matrices  $U_N, U_Q$  with positive diagonal entries.  
Thus for group elements

$$A(x_1, x_2, x_3) = \begin{bmatrix} x_1 & x_2 & x_3 \\ & 1 & \\ & & 1 \end{bmatrix} \quad (3)$$

we introduce

$$\begin{aligned} \tilde{f}: \mathbb{R}^3 &\rightarrow \mathbb{R}, \\ \tilde{f}(x_1, x_2, x_3) &:= \|A(x_1, x_2, x_3)U_N - U_Q\|^2 \end{aligned}$$

to be minimized. The function  $\tilde{f}$  is convex and its gradient and Hessian can be easily computed.

# Alternative Approach: Cholesky Factorisation

$$U_N = \begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & g \end{bmatrix}, \quad U_Q = \begin{bmatrix} r & s & t \\ 0 & u & v \\ 0 & 0 & w \end{bmatrix}.$$

$$\nabla \tilde{f}(x, y, z) = 2U_N(A(x, y, z)U_N - U_Q)^\top \mathbf{e}_1$$

$$H_{\tilde{f}(x,y,z)} = 2U_N U_N^\top = 2N = 2 \begin{bmatrix} a^2+b^2+c^2 & bd+ce & cg \\ bd+ce & d^2+e^2 & eg \\ cg & eg & g^2 \end{bmatrix}.$$

Clearly,  $H_{\tilde{f}(x,y,z)} \succ 0$ . A Newton iteration step for this problem then moves right into the minimum

$$\begin{bmatrix} x_{t+1} \\ y_{t+1} \\ z_{t+1} \end{bmatrix} = \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} - H_{\tilde{f}(x_t, y_t, z_t)}^{-1} \nabla \tilde{f}(x_t, y_t, z_t) = \begin{bmatrix} \frac{r}{a} \\ \frac{as-rb}{ad} \\ \frac{adt-cdr-aes+ber}{adg} \end{bmatrix}. \quad (4)$$

# Alternative Approach: Cholesky Factorisation

Thus,  $A(x, y, z) \in G$  with

$$x = \frac{r}{a}, \quad y = \frac{as-rb}{ad}, \quad z = \frac{adt-cdr-aes+ber}{adg}$$

is the unique group element minimizing  $\tilde{f}$ .

In the noise free case at the minimum

$$d = u, \quad e = v, \quad g = w,$$

$AU_N = U_Q$  and therefore the minimal value is equal to zero.

# Numerical Examples

# Outlook

- Cholesky updating (learning approach).
- Cholesky based solution seems to be more sensitive to noise than Newton's method.
- Generalisation to non-planar patches?
- Non-coplanar cameras, possibly different ones?

# References

- **A Gradient Algorithm for Stereo Matching without Correspondence** by J. Zhou, B.K. Ghosh; IEEE TAC 41(11) 1996.
- **Stereo Matching for Calibrated Cameras without Correspondence** by U. Helmke *et al.*; IEEE CDC 2008.

End

Thanks!