

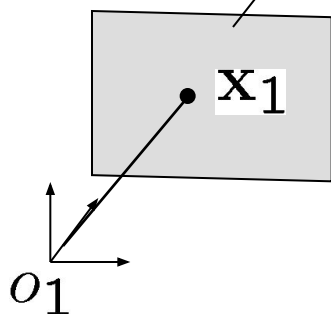


Lecture 4

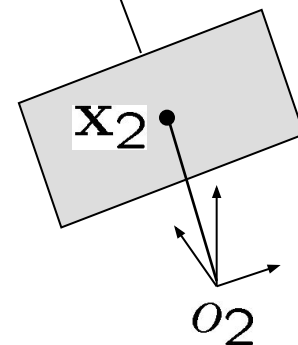
Two-View Geometry



General Formulation



Given two views of the scene
recover the unknown camera
displacement and 3D scene
structure





Pinhole Camera Model

- 3D points $\mathbf{X} = [X, Y, Z, W]^T \in \mathbb{R}^4, \quad (W = 1)$
- Image points $\mathbf{x} = [x, y, z]^T \in \mathbb{R}^3, \quad (z = 1)$
- Perspective Projection $\lambda \mathbf{x} = \mathbf{X}$

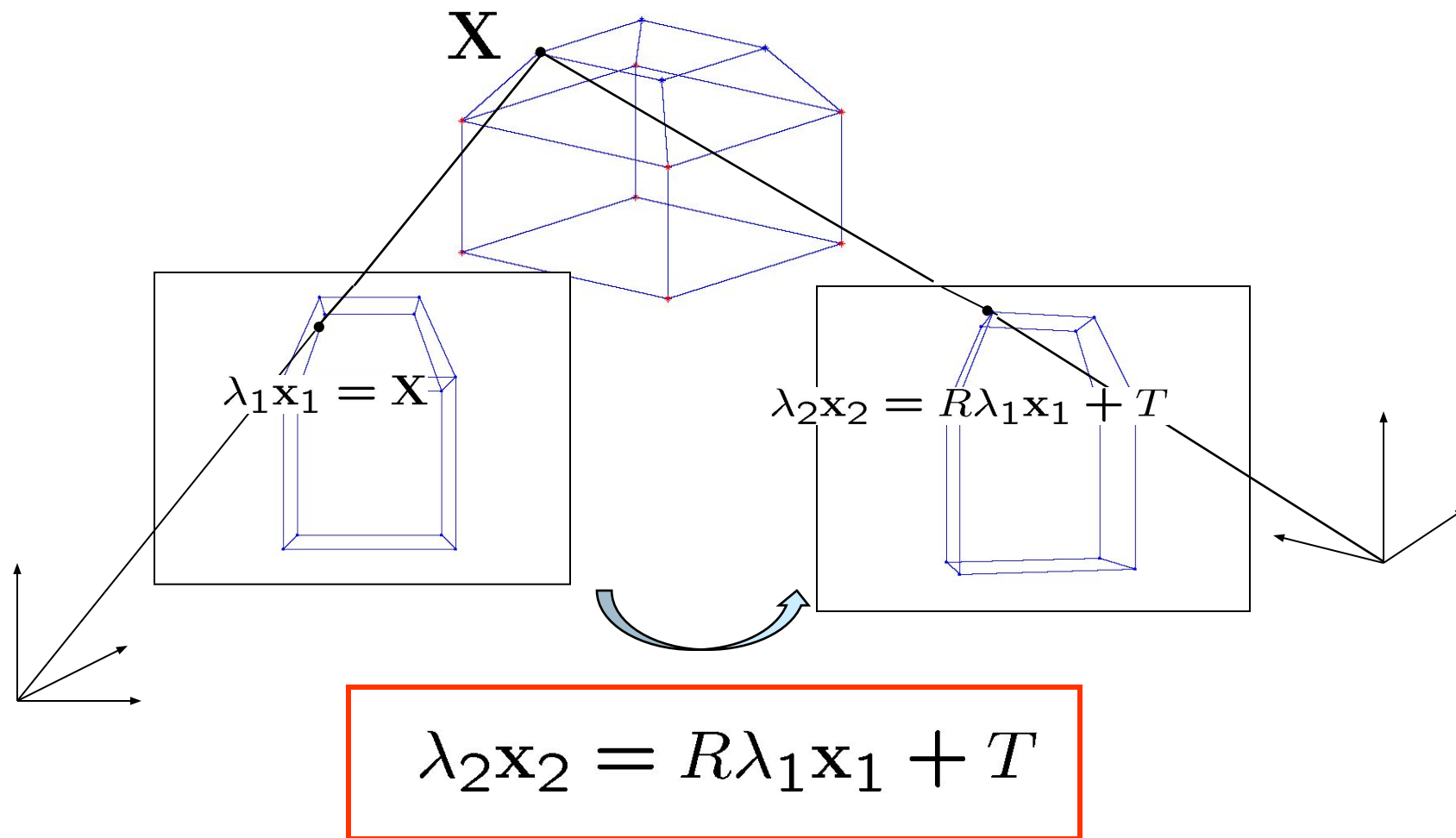
$$\lambda = Z \quad x = \frac{X}{Z} \quad y = \frac{Y}{Z}$$

- Rigid Body Motion $\Pi = [R, T] \in \mathbb{R}^{3 \times 4}$
- Rigid Body Motion + Projective projection

$$\lambda \mathbf{x} = \Pi \mathbf{X} = [R, T] \mathbf{X}$$



Rigid Body Motion – Two views





3D Structure and Motion Recovery

Euclidean transformation

$$\lambda_2 \mathbf{x}_2 = R \lambda_1 \mathbf{x}_1 + T$$

measurements

unknowns

$$\sum_{j=1}^n \|\mathbf{x}_1^j - \pi(R_1, T_1, \mathbf{X})\|^2 + \|\mathbf{x}_2^j - \pi(R_2, T_2, \mathbf{X})\|^2$$

Find such **Rotation** and **Translation** and **Depth** that
the reprojection error is minimized

Two views \sim 200 points

6 unknowns – **Motion** 3 Rotation, 3 Translation

- **Structure** 200x3 coordinates

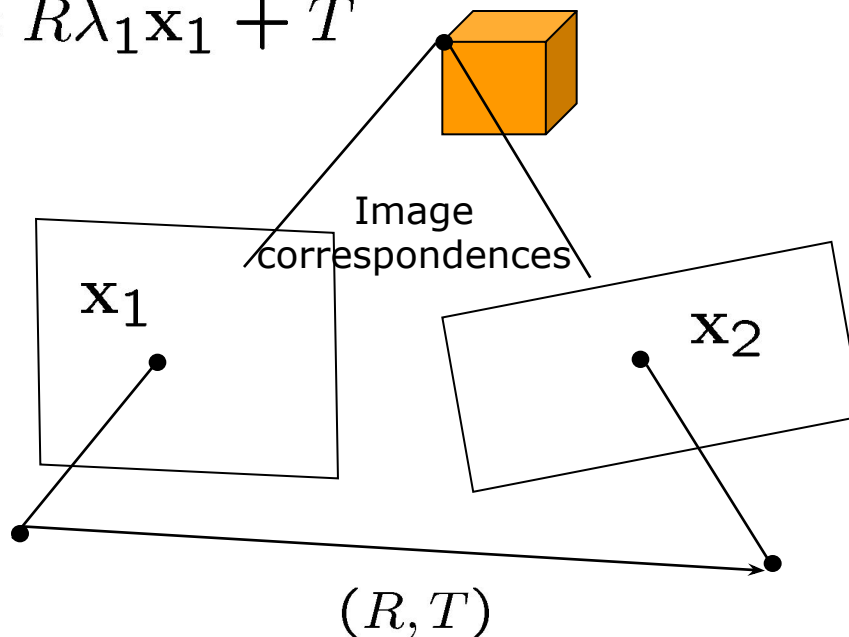
- (-) universal scale

Difficult optimization problem



Epipolar Geometry

$$\lambda_2 \mathbf{x}_2 = R\lambda_1 \mathbf{x}_1 + T$$



- Algebraic Elimination of Depth [Longuet-Higgins '81]:

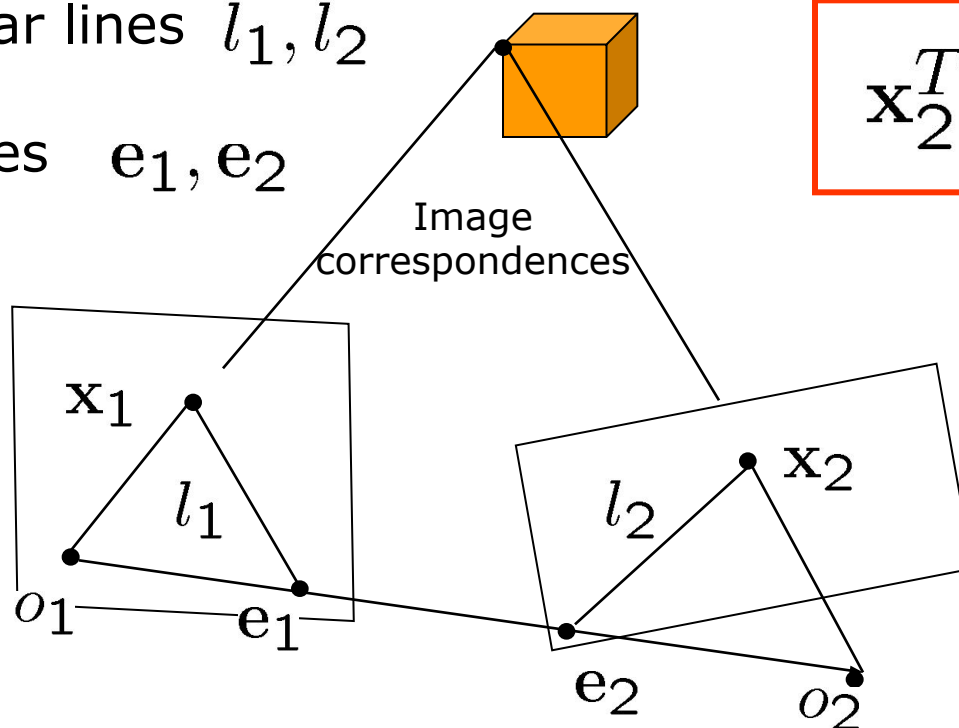
$$\mathbf{x}_2^T \underbrace{\hat{T}R}_E \mathbf{x}_1 = 0$$

- Essential matrix $E = \hat{T}R$



Epipolar Geometry

- Epipolar lines l_1, l_2
- Epipoles e_1, e_2



$$\mathbf{x}_2^T E \mathbf{x}_1 = 0$$

$$E = \hat{T}R$$

$$l_1 \sim E^T \mathbf{x}_2$$

$$l_i^T \mathbf{x}_i = 0$$

$$l_2 \sim E \mathbf{x}_1$$

$$E \mathbf{e}_1 = 0$$

$$l_i^T \mathbf{e}_i = 0$$

$$\mathbf{e}_2 E^T = 0$$



Characterization of the Essential Matrix

$$\mathbf{x}_2^T \hat{T} R \mathbf{x}_1 = 0$$

- Essential matrix $E = \hat{T} R$ Special 3x3 matrix

$$\mathbf{x}_2^T \begin{bmatrix} e_1 & e_2 & e_3 \\ e_4 & e_5 & e_6 \\ e_7 & e_8 & e_9 \end{bmatrix} \mathbf{x}_1 = 0$$

Theorem 1a (Essential Matrix Characterization)

A non-zero matrix E is an essential matrix iff its SVD: $E = U \Sigma V^T$ satisfies: $\Sigma = \text{diag}([\sigma_1, \sigma_2, \sigma_3])$ with $\sigma_1 = \sigma_2 \neq 0$ and $\sigma_3 = 0$ and $U, V \in SO(3)$



Estimating the Essential Matrix

- Estimate Essential matrix $E = \hat{T}R$
- Decompose Essential matrix into R, T

$$\mathbf{x}_2^T \hat{T} R \mathbf{x}_1 = 0$$

- Given n pairs of image correspondences:
 - Find such **Rotation** and **Translation** that the epipolar error is minimized

$$\min_E \sum_{j=1}^n \mathbf{x}_2^{jT} E \mathbf{x}_1^j$$

- Space of all **Essential Matrices** is 5 dimensional
- 3 Degrees of Freedom – Rotation
- 2 Degrees of Freedom – Translation (**up to scale !**)



Pose Recovery from the Essential Matrix

Essential matrix $E = \hat{T}R$

Theorem 1a (Pose Recovery)

There are two relative poses (R, T) with $T \in \mathcal{R}^3$ and $R \in SO(3)$ corresponding to a non-zero matrix essential matrix.

$$E = U\Sigma V^T$$

$$\begin{aligned}(\hat{T}_1, R_1) &= (UR_Z(+\frac{\pi}{2})\Sigma U^T, UR_Z^T(+\frac{\pi}{2})V^T) \\ (\hat{T}_2, R_2) &= (UR_Z(-\frac{\pi}{2})\Sigma U^T, UR_Z^T(-\frac{\pi}{2})V^T)\end{aligned}$$

$$\Sigma = \text{diag}([1, 1, 0]) \quad R_z(+\frac{\pi}{2}) = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

• Twisted pair ambiguity $(R_2, T_2) = (e^{\hat{u}\pi}R_1, -T_1)$



Estimating Essential Matrix

$$\mathbf{x}_2^T \hat{T} R \mathbf{x}_1 = 0$$

- Denote $\mathbf{a} = \mathbf{x}_1 \otimes \mathbf{x}_2$

$$\mathbf{a} = [x_1 x_2, x_1 y_2, x_1 z_2, y_1 x_2, y_1 y_2, y_1 z_2, z_1 x_2, z_1 y_2, z_1 z_2]^T$$

$$E^s = [e_1, e_4, e_7, e_2, e_5, e_8, e_3, e_6, e_9]^T$$

- Rewrite $\mathbf{a}^T E^s = 0$

- Collect constraints from all points

$$\chi E^s = 0$$

$$\min_E \sum_{j=1}^n \mathbf{x}_2^{jT} E \mathbf{x}_1^j \quad \longrightarrow \quad \min_{E^s} \|\chi E^s\|^2$$



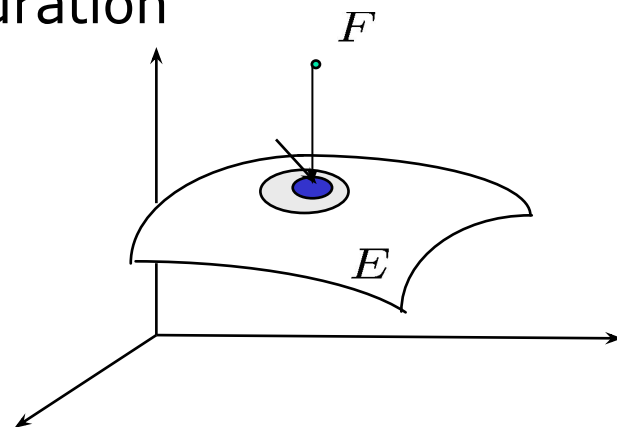
Estimating Essential Matrix

$$\min_E \sum_{j=1}^n (\mathbf{x}_2^{jT} E \mathbf{x}_1^j)^2 \longrightarrow \min_{E^s} \|\chi E^s\|^2$$

Solution

- Eigenvector associated with the smallest eigenvalue of $\chi^T \chi$
- if $\text{rank}(\chi^T \chi) < 8$ degenerate configuration

Projection on to Essential Space



Theorem 2a (Project to Essential Manifold)

If the SVD of a matrix $F \in \mathcal{R}^{3 \times 3}$ is given by $F = U \text{diag}(\sigma_1, \sigma_2, \sigma_3) V^T$ then the essential matrix E which minimizes the Frobenius distance $\|E - F\|_f^2$ is given by $E = U \text{diag}(\sigma, \sigma, 0) V^T$ with $\sigma = \frac{\sigma_1 + \sigma_2}{2}$



Two view linear algorithm

$$E = \{\hat{T}R | R \in SO(2), T \in S^2\}$$

- Solve the **LLSE** problem:

$$\min_E \sum_{j=1}^n \mathbf{x}_2^{jT} E \mathbf{x}_1^j$$

$\chi E^S = 0$ followed by projection

- **Project** onto the essential manifold:

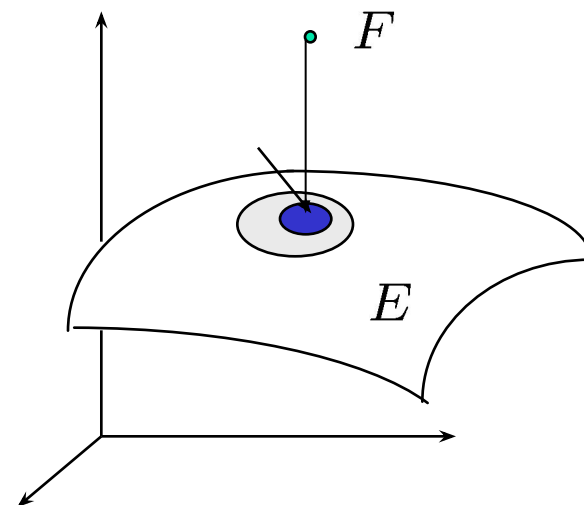
$$\text{SVD: } F = U \Sigma V^T$$

$$\Sigma' = \text{diag}(1, 1, 0)$$

$$E = U \Sigma' V^T$$

- **Recover** the unknown pose:

$$(\hat{T}, R) = (UR_Z(\pm \frac{\pi}{2}) \Sigma U^T, UR_Z^T(\pm \frac{\pi}{2}) V^T)$$



E is 5 diml. sub. mnfld. in

- 8-point linear algorithm



Pose Recovery

- There are exactly **two** pairs (R, T) corresponding to each essential matrix E .
- There are also **two** pairs (R, T) corresponding to each essential matrix $-E$.
- Positive depth constraint - used to disambiguate the physically impossible solutions
- Translation has to be non-zero
- Points have to be in general position
 - degenerate configurations – planar points
 - quadratic surface
- Linear 8-point algorithm
- Nonlinear 5-point algorithms yield up to 10 solutions



3D structure recovery

$$\lambda_2 \underline{\mathbf{x}_2} = \underline{R} \lambda_1 \underline{\mathbf{x}_1} + \underline{\gamma T}$$

- Eliminate one of the scale's

$$\lambda_1^j \widehat{\mathbf{x}_2^j} R \mathbf{x}_1^j + \gamma \widehat{\mathbf{x}_2^j} T = 0, \quad j = 1, 2, \dots, n$$

- Solve LLSE problem

$$M^j \bar{\lambda}^j \doteq \begin{bmatrix} \widehat{\mathbf{x}_2^j} R \mathbf{x}_1^j, & \widehat{\mathbf{x}_2^j} T \end{bmatrix} \begin{bmatrix} \lambda_1^j \\ \gamma \end{bmatrix} = 0$$

If the configuration is non-critical, the Euclidean structure of then points and motion of the camera can be reconstructed up to a universal scale.



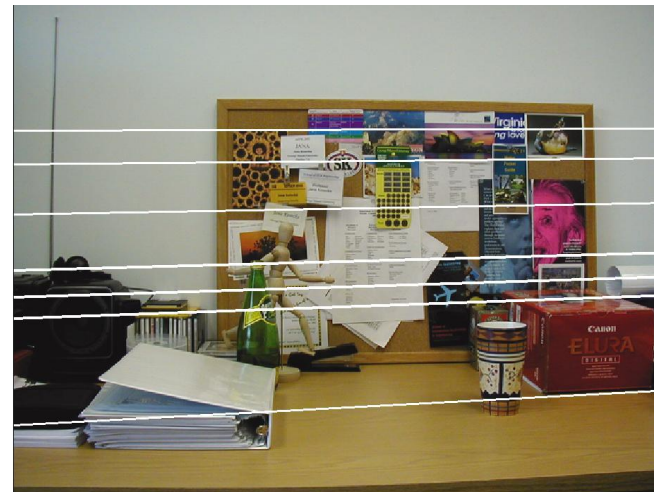
Example- Two views



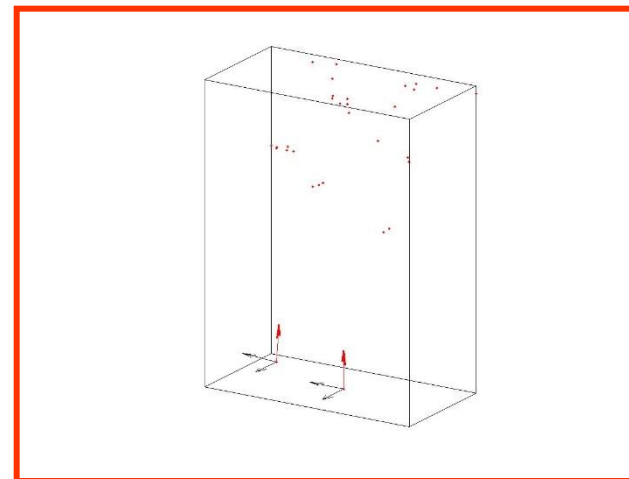
Point Feature Matching



Example – Epipolar Geometry



Camera Pose
and
Sparse Structure Recovery





Epipolar Geometry – Planar Case

- Plane in first camera coordinate frame

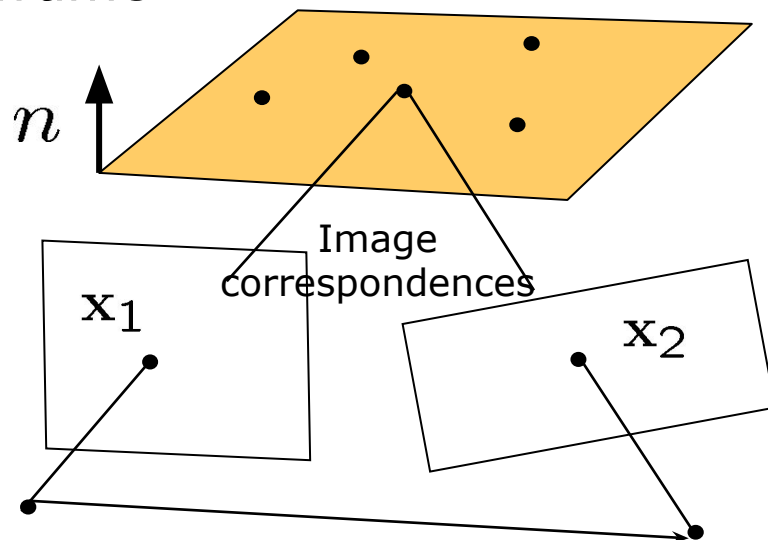
$$aX + bY + cZ + d = 0$$

$$\frac{1}{d}N^T\mathbf{X} = 1$$

$$\lambda_2\mathbf{x}_2 = R\lambda_1\mathbf{x}_1 + T$$

$$\lambda_2\mathbf{x}_2 = (R + \frac{1}{d}TN^T)\lambda_1\mathbf{x}_1$$

$$\mathbf{x}_2 \sim H\mathbf{x}_1$$



Planar homography

Linear mapping relating two corresponding planar points in two views

$$H = (R + \frac{1}{d}TN^T)$$



Decomposition of H

- Algebraic elimination of depth $\widehat{\mathbf{x}}_2^T H \mathbf{x}_1 = 0$
- H_L can be estimated linearly $H_L = \lambda H$
- Normalization of $H = H_L / \sigma_3$
- Decomposition of H into 4 solutions $H = (R + \frac{1}{d} T N^T)$

$R_1 = W_1 U_1^T$ $N_1 = \widehat{v}_2 u_1$ $\frac{1}{d} T_1 = (H - R_1) N_1$	$R_3 = R_1$ $N_3 = -N_1$ $\frac{1}{d} T_3 = -\frac{1}{d} T_1$	$R_2 = W_2 U_2^T$ $N_2 = \widehat{v}_2 u_2$ $\frac{1}{d} T_2 = (H - R_2) N_2$	$R_4 = R_2$ $N_4 = -N_2$ $\frac{1}{d} T_4 = -\frac{1}{d} T_2$
---	---	---	---

$$H^T H = V \Sigma V^T \quad V = [v_1, v_2, v_3] \quad \Sigma = \text{diag}(\sigma_1^2, \sigma_2^2, \sigma_3^2)$$

$$u_1 \doteq \frac{\sqrt{1-\sigma_3^2} v_1 + \sqrt{\sigma_1^2-1} v_3}{\sqrt{\sigma_1^2-\sigma_3^2}} \quad u_2 \doteq \frac{\sqrt{1-\sigma_3^2} v_1 - \sqrt{\sigma_1^2-1} v_3}{\sqrt{\sigma_1^2-\sigma_3^2}}$$

$$U_1 = [v_2, u_1, \widehat{v}_2 u_1], \quad W_1 = [H v_2, H u_1, H v_2 H u_1];$$

$$U_2 = [v_2, u_2, \widehat{v}_2 u_2], \quad W_2 = [H v_2, H u_2, \widehat{H v}_2 H u_2].$$



Motion and pose recovery for planar scene

- Given at least 4 point correspondences $\widehat{\mathbf{x}}_2^j H \mathbf{x}_1^j = 0$
- Compute an approximation of the homography matrix H_l^s
- As nullspace of χ
 $\chi H_l^s = 0$ the rows of χ are $\mathbf{a} = \mathbf{x}_1^j \otimes \widehat{\mathbf{x}}_2^j$
- Normalize the homography matrix

$$H = H_L / \sigma_3$$

- Decompose the homography matrix

$$H^T H = V \Sigma V^T$$

- Select two physically possible solutions imposing positive depth constraint



Example





Special Rotation Case

- Two view related by rotation only $\lambda_2 \mathbf{x}_2 = R \lambda_1 \mathbf{x}_1$

$$\widehat{\mathbf{x}}_2 R \mathbf{x}_1 = 0$$

- Mapping to a reference view



- Mapping to a cylindrical surface





Motion and Structure Recovery – Two Views

- Two views – general motion, general structure
 1. Estimate essential matrix
 2. Decompose the essential matrix
 3. Impose positive depth constraint
 4. Recover 3D structure
- Two views – general motion, planar structure
 1. Estimate planar homography
 2. Normalize and decompose H
 3. Recover 3D structure and camera pose