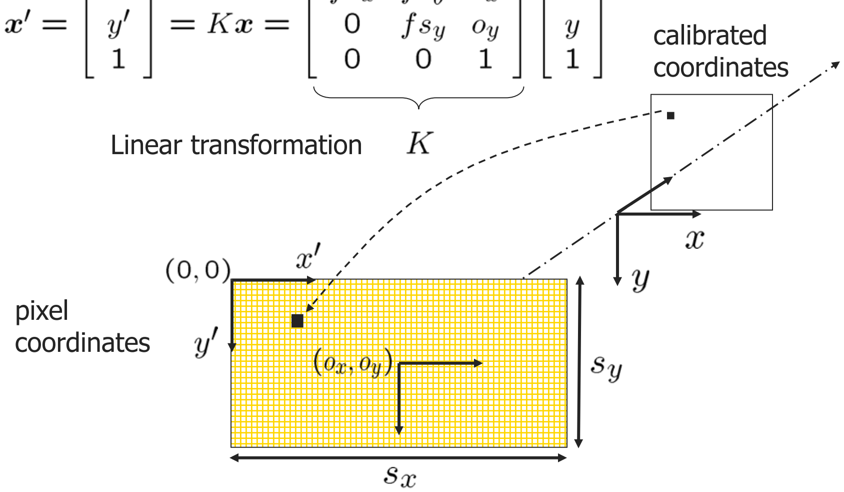
Uncalibrated Two-View Geometry

1

Uncalibrated Camera

$$\boldsymbol{x}' = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = K\boldsymbol{x} = \underbrace{\begin{bmatrix} fs_x & fs_\theta & o_x \\ 0 & fs_y & o_y \\ 0 & 0 & 1 \end{bmatrix}}_{K} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Linear transformation K



2

Overview

- Calibration with a rig
- Uncalibrated epipolar geometry
- Ambiguities in image formation
- Stratified reconstruction
- Autocalibration with partial scene knowledge

Calibration with a Rig

Use the fact that both 3-D and 2-D coordinates of feature points on a pre-fabricated object (e.g., a cube) are known.

Calibration with a Rig

- Given 3-D coordinates on known object

$$\lambda \mathbf{x}' = [KR, KT]\mathbf{X} \Rightarrow \lambda \mathbf{x}' = \Pi \mathbf{X}$$

$$\lambda \begin{bmatrix} x^i \\ y^i \\ 1 \end{bmatrix} = \begin{bmatrix} \pi_1^T \\ \pi_2^T \\ \pi_3^T \end{bmatrix} \begin{bmatrix} X^i \\ Y^i \\ Z^i \\ 1 \end{bmatrix}$$

- Eliminate unknown scales

$$x^i(\pi_3^T \mathbf{X}) = \pi_1^T \mathbf{X},$$
$$y^i(\pi_3^T \mathbf{X}) = \pi_2^T \mathbf{X}$$

- Recover projection matrix $\Pi = [KR, KT] = [R', T']$

$$\min \|M\Pi^s\|^2 \quad \text{subject to} \quad \|\Pi^s\|^2 = 1$$

$$\Pi^s = [\pi_{11}, \pi_{21}, \pi_{31}, \pi_{12}, \pi_{22}, \pi_{32}, \pi_{13}, \pi_{23}, \pi_{33}, \pi_{14}, \pi_{24}, \pi_{34}]^T$$

- Factor the *KR* into $R \in SO(3)$ and *K* using QR decomposition
- Solve for translation $T = K^{-1}T'$

More details

- Direct calibration by recovering and decomposing the projection matrix

$$\lambda \begin{bmatrix} x^i \\ y^i \\ 1 \end{bmatrix} = \begin{bmatrix} \pi_1^T \\ \pi_2^T \\ \pi_3^T \end{bmatrix} \begin{bmatrix} X^i \\ Y^i \\ Z^i \\ 1 \end{bmatrix} \rightarrow Z \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{13} & \pi_{14} \\ \pi_{21} & \pi_{22} & \pi_{23} & \pi_{24} \\ \pi_{31} & \pi_{32} & \pi_{33} & \pi_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$x_i = \frac{\pi_{11}X_i + \pi_{12}Y_i + \pi_{13}Z_i + \pi_{14}}{\pi_{31}X_i + \pi_{32}Y_i + \pi_{33}Z_i + \pi_{34}} \qquad y_i = \frac{\pi_{21}X_i + \pi_{22}Y_i + \pi_{23}Z_i + \pi_{24}}{\pi_{31}X_i + \pi_{32}Y_i + \pi_{33}Z_i + \pi_{34}}$$

$$x_i(\pi_{31}X_i + \pi_{32}Y_i + \pi_{33}Z_i + \pi_{34}) = \pi_{11}X_i + \pi_{12}Y_i + \pi_{13}Z_i + \pi_{14}$$
$$y_i(\pi_{31}X_i + \pi_{32}Y_i + \pi_{33}Z_i + \pi_{34}) = \pi_{21}X_i + \pi_{22}Y_i + \pi_{23}Z_i + \pi_{24}$$

$$x^i(\pi_3^T \mathbf{X}) = \pi_1^T \mathbf{X},$$
$$y^i(\pi_3^T \mathbf{X}) = \pi_2^T \mathbf{X}$$

2 constraints per point

$$[X_i, Y_i, Z_i, 1, 0, 0, 0, 0, -x_iX_i, -x_iY_i, -x_iZ_i, -x_i]\Pi_s = 0$$
$$[0, 0, 0, 0, X_i, Y_i, Z_i, 1, -y_iX_i, -y_iY_i, -y_iZ_i, -y_i]\Pi_s = 0$$

$$\Pi_s = [\pi_{11}, \pi_{12}, \pi_{13}, \pi_{14}, \pi_{21}, \pi_{22}, \pi_{23}, \pi_{24}, \pi_{31}, \pi_{32}, \pi_{33}, \pi_{34}, \pi_{41}, \pi_{42}, \pi_{43}, \pi_{44}]^T$$

More details

- Recover projection matrix $\Pi = [KR, KT] = [R', T']$

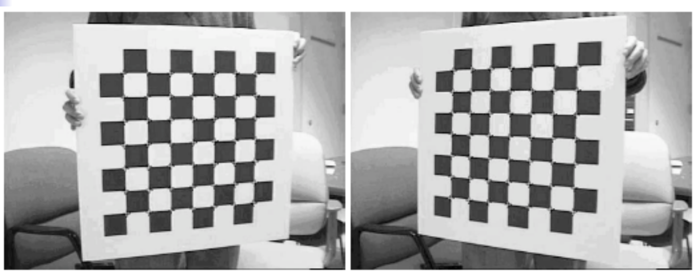
$$\min \|M\Pi^s\|^2 \quad \text{subject to} \quad \|\Pi^s\|^2 = 1$$

$$\Pi^s = [\pi_{11}, \pi_{21}, \pi_{31}, \pi_{12}, \pi_{22}, \pi_{32}, \pi_{13}, \pi_{23}, \pi_{33}, \pi_{14}, \pi_{24}, \pi_{34}]^T$$

- Collect the constraints from all N points into matrix M (2N x 12)
- Solution eigenvector associated with the smallest eigenvalue $M^T M$
  [u,s,v] = svd(M) take v(:,12)
- Unstack the solution and decompose into rotation and translation
- Factor the R' into $R \in SO(3)$ and K using QR decomposition (qr matlab function)
- Solve for translation $T = K^{-1}T'$

7

Calibration with a planar pattern



$$H \doteq K[r_1, r_2, T] \quad \in \mathbb{R}^{3\times 3} \qquad \lambda \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = K[r_1, r_2, T] \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix},$$

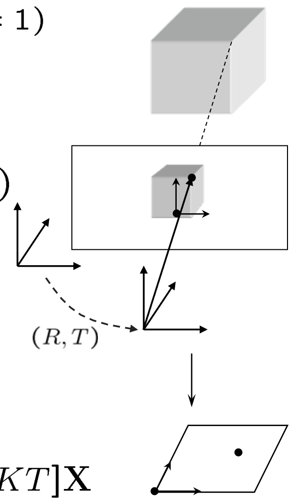
To eliminate unknown depth, multiply both sides by $\widehat{x'}$

$$\widehat{x'} H [X, Y, 1]^T = 0.$$

8

Uncalibrated Camera

$$\mathbf{X} = [X, Y, Z, W]^T \in \mathbb{R}^4, \quad (W = 1)$$



Calibrated camera

- Image plane coordinates $\mathbf{x} = [x, y, 1]^T$
- Camera extrinsic parameters $g = (R, T)$
- Perspective projection $\lambda \mathbf{x} = [R, T]\mathbf{X}$

Uncalibrated camera

- Pixel coordinates $\mathbf{x}' = K\mathbf{x}$
- Projection matrix $\lambda \mathbf{x}' = \Pi \mathbf{X} = [KR, KT]\mathbf{X}$

9

Taxonomy on Uncalibrated Reconstruction

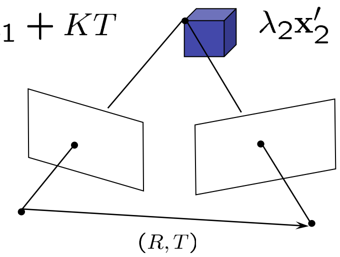
$$\lambda \mathbf{x}' = [KR, KT]\mathbf{X}$$

- K is known, back to calibrated case $\mathbf{x} = K^{-1}\mathbf{x}'$
- K is unknown
  - Calibration with complete scene knowledge (a rig) – estimate
  - Uncalibrated reconstruction despite the lack of knowledge of
  - Autocalibration (recover from uncalibrated images)
- Use partial knowledge
  - Parallel lines, vanishing points, planar motion, constant intrinsic
- Ambiguities, stratification (multiple views)

10

Uncalibrated Epipolar Geometry

$\lambda_2 K \mathbf{x}_2 = K R \lambda_1 \mathbf{x}_1 + K T$ $\qquad$ $\lambda_2 \mathbf{x}'_2 = K R K^{-1} \lambda_1 \mathbf{x}'_1 + T'$

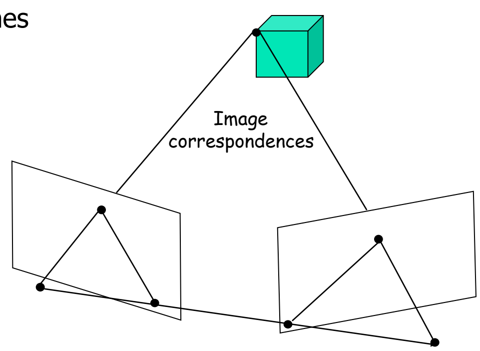


- Epipolar constraint $\quad \mathbf{x'}_2^T \underbrace{K^{-T} \widehat{T} R K^{-1}} \mathbf{x'}_1 = 0$
- Fundamental matrix $\quad F = K^{-T} \widehat{T} R K^{-1}$
- Equivalent forms of $\quad F = K^{-T} \widehat{T} R K^{-1} = \widehat{T}' K R K^{-1}$

Properties of the Fundamental Matrix

$$\mathbf{x'}_2^T F \mathbf{x'}_1 = 0$$

- Epipolar lines
- Epipoles



$l_1 \sim F^T \mathbf{x}'_2$ $\qquad$ $l_i^T \mathbf{x}'_i = 0$ $\qquad$ $l_2 \sim F \mathbf{x}'_1$

$F \mathbf{e}_1 = 0$ $\qquad$ $l_i^T \mathbf{e}_i = 0$ $\qquad$ $\mathbf{e}_2^T F = 0$

Properties of the Fundamental Matrix

A nonzero matrix $F \in \mathbb{R}^{3\times 3}$ is a fundamental matrix if has a singular value decomposition (SVD) $F = U\Sigma V^T$ with

$$\Sigma = \text{diag}\{\sigma_1, \sigma_2, 0\}$$

for some $\sigma_1, \sigma_2 \in \mathbb{R}_+$.

There is little structure in the matrix except that

$$\det(F) = 0$$

13

Estimating Fundamental Matrix

- Find such F that the epipolar error is minimized

$$\min_F \sum_{j=1}^{n} (\mathbf{x}'^{jT}_2 F \mathbf{x}'^{j}_1)^2 \quad \longleftarrow \text{ Pixel coordinates}$$

- Fundamental matrix can be estimated up to scale
- Denote $\mathbf{a} = \mathbf{x}'_1 \otimes \mathbf{x}'_2$

$$\mathbf{a} = [x_1x_2, x_1y_2, x_1z_2, y_1x_2, y_1y_2, y_1z_2, z_1x_2, z_1y_2, z_1z_2]^T$$

$$F^s = [f_1, f_4, f_7, f_2, f_5, f_8, f_3, f_6, f_9]^T$$

- Rewrite $\mathbf{a}^T F^s = 0$
- Collect constraints from all points

$$\chi F^s = 0$$

$$\min_F \sum_{j=1}^{n} (\mathbf{x}'^{jT}_2 F \mathbf{x}'^{j}_1)^2 \quad \Rightarrow \quad min_{F^s} \|\chi F^s\|^2$$

14

Two view linear algorithm – 8-point algorithm

- Solve the LLSE problem:

$$\min_F \sum_{j=1}^{n} (\mathbf{x'}_2^{jT} F \mathbf{x'}_1^{j})^2 \quad \Rightarrow \quad \chi F^s = 0$$

- Solution eigenvector associated with smallest eigenvalue of $\chi^T \chi$

- Compute SVD of F recovered from data

$$F = U \Sigma V^T \quad \Sigma = diag(\sigma_1, \sigma_2, \sigma_3)$$

- Project onto the essential manifold:

$$\Sigma' = diag(\sigma_1, \sigma_2, 0) \quad F = U \Sigma' V^T$$

- cannot be unambiguously decomposed into pose and calibration

$$F = K^{-T} \hat{T} R K^{-1}$$

15

What Does F Tell Us?

- F can be inferred from point matches (eight-point algorithm)
- Cannot extract motion, structure and calibration from one fundamental matrix (two views)
- F allows reconstruction up to a projective transformation (as we will see soon)
- F encodes all the geometric information among two views when no additional information is available
- F is often used in robust matching for establishing correspondences (one cannot recover R,T from single F)

16

Projective Reconstruction

- From points, extract , followed by computation of projection matrices Π_{1p}, Π_{2p} and structure $\mathbf{X}_p$
- Canonical decomposition

$$F \quad \mapsto \quad \Pi_{1p} = [I,\ 0],\ \Pi_{2p} = [(\widehat{T'})^T F,\ T']$$

- Given projection matrices – recover structure

$$\begin{aligned} \lambda_1 \mathbf{x}'_1 &= \Pi_{1p}\mathbf{X}_p = [I, 0]\mathbf{X}_p, \\ \lambda_2 \mathbf{x}'_2 &= \Pi_{2p}\mathbf{X}_p = [(\widehat{T'})^T F, T']\mathbf{X}_p. \end{aligned}$$

- Projective ambiguity – non-singular 4x4 matrix

$$\lambda_i \mathbf{x}'_i = \boxed{\Pi_{ip} H^{-1} H \mathbf{X}_p}$$

$$\lambda_i \mathbf{x}'_i = \tilde{\Pi}_{1p} \tilde{\mathbf{X}}_p$$

Both $\mathbf{X}_p$ and Π_{1p}, Π_{1p} are consistent with the epipolar geometry – give the same fundamental matrix

17

Projective Reconstruction

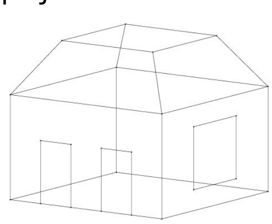
- Given projection matrices recover projective structure

$$\begin{aligned} (x_1 \pi_1^{3T})\mathbf{X}_p &= \pi_1^{1T}\mathbf{X}_p, & (y_1 \pi_1^{3T})\mathbf{X}_p &= \pi_1^{2T}\mathbf{X}_p, \\ (x_2 \pi_2^{3T})\mathbf{X}_p &= \pi_2^{1T}\mathbf{X}_p, & (y_2 \pi_2^{3T})\mathbf{X}_p &= \pi_2^{2T}\mathbf{X}_p, \end{aligned}$$

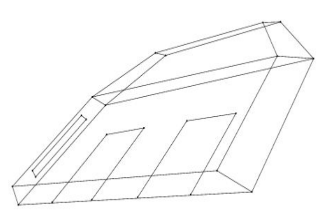
- This is a linear problem and can be solve using linear least-squares

$$M\mathbf{X}_p = 0$$

- Projective reconstruction – projective camera matrices and projective structure



$$\mathbf{X}_e = H\mathbf{X}_p$$



Euclidean Structure

Projective Structure

18

Euclidean vs Projective reconstruction

- Euclidean reconstruction – true metric properties of objects lenghts (distances), angles, parallelism are preserved
- Unchanged under rigid body transformations
- => Euclidean Geometry – properties of rigid bodies under rigid body transformations, similarity transformation

- Projective reconstruction – lengths, angles, parallelism are NOT preserved – we get distorted images of objects – their distorted 3D counterparts --> 3D projective reconstruction
- => Projective Geometry

19

Euclidean Geometry

- Describes shapes as they are
- Properties of objects that are unchanged by Rigid Body Transformation
- - lengths
- - angles
- - parallelism

20

Projective Geometry

- Describes things as they are
- Lengths, angles become distorted
- When we look at the objects
- Mathematical model how the images of the world are formed

Examples – corner of the room
- railroad tracks

Example – parallax – displacement of objects due to the change of viewpoints

Homogeneous Coordinates (RBM)

3-D coordinates are related by: $\boldsymbol{X}_c = R\boldsymbol{X}_w + T,$

Homogeneous coordinates:

$$\boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad \rightarrow \quad \boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \in \mathbb{R}^4,$$

Homogeneous coordinates are related by:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

Homogenous and Projective Coordinates

- Homogenous coordinates in 3D before – attach 1 as the last coordinate – render the transformation as linear transformation
- Before 4th coordinate cannot be zero 0
- Projective coordinates – all points are equivalent up to a scale

$$\boldsymbol{X} = \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \approx \boldsymbol{X} = \begin{bmatrix} WX \\ WY \\ W \end{bmatrix} \in \mathbb{R}^3 \qquad \boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \approx \boldsymbol{X} = \begin{bmatrix} WX \\ WY \\ WZ \\ W \end{bmatrix} \in \mathbb{R}^4$$

2D- projective plane

3D- projective space

Each point on the plane is represented by a ray in projective space

Homogenous and Projective Coordinates

- Ideal points – last coordinate is 0 – ray parallel to the image plane point at infinity – never intersects the image plane

$$\boldsymbol{X} = \begin{bmatrix} X \\ Y \\ 0 \end{bmatrix} \qquad \boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \\ 0 \end{bmatrix} \in \mathbb{R}^4$$

Vanishing points

Representation of a 3-D line – in homogeneous coordinates

$$\boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} X_o \\ Y_o \\ Z_o \\ 1 \end{bmatrix} + \lambda \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ 0 \end{bmatrix}, \quad \mu \in \mathbb{R}$$

When λ -> infinity - vanishing points – last coordinate -> 0

$$\boldsymbol{X} = \begin{bmatrix} X_o + \lambda V_1 \\ Y_o + \lambda V_2 \\ Z_o + \lambda V_3 \\ 1 \end{bmatrix} \qquad \boldsymbol{X} = \begin{bmatrix} X_o/\lambda + V_1 \\ Y_o/\lambda + V_2 \\ Z_o/\lambda + V_3 \\ 1/\lambda \end{bmatrix} \qquad \boldsymbol{X} = \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ 0 \end{bmatrix}$$

Projection of a line - line in the image plane

$$x = \frac{X_o + \lambda V_1}{Z_o + \lambda V_3}$$
$$y = \frac{Y_o + \lambda V_2}{Z_o + \lambda V_3}$$

Calibration using vanishing points

- Vanishing points – intersections of the parallel lines

$$v_i = l_1 \times l_2 = \widehat{l_1} l_2$$

- Vanishing points of three orthogonal directions

$$\mathbf{v}_1 = KRe_1, \quad \mathbf{v}_2 = KRe_2, \quad \mathbf{v}_3 = KRe_1$$

- Orthogonal directions – inner product is zero

$$\mathbf{v}_i^T S \mathbf{v}_j = \mathbf{v}_i^T K^{-T} K^{-1} \mathbf{v}_j = e_i^T R^T R e_j = e_i^T e_j = 0, \quad i \neq j,$$

- Provide directly constraints on matrix $S = K^{-T} K^{-1}$
- S – has 5 degrees of freedom, 3 vanishing points – 3 constraints (need additional assumption about K)
- Assume zero skew and aspect ratio = 1

Applications of projective geometry

Vermeer's *Music Lesson*

27

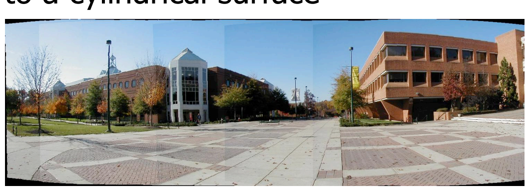
Rotation Only - Calibrated Case

- Calibrated Two views related by rotation only

$$\lambda_2 \mathbf{x}_2 = R\lambda_1 \mathbf{x}_1 \qquad \widehat{\mathbf{x}_2} R \mathbf{x}_1 = 0$$

- Mapping to a reference view – rotation can be estimated

- Mapping to a cylindrical surface



28

Rotation Only - Uncalibrated Case

- Calibrated Two views related by rotation only

$$\lambda_2 K \mathbf{x}_2 = \lambda_1 K R K^{-1} K \mathbf{x}_1 \qquad \widehat{\mathbf{x}'_2} K R K^{-1} \mathbf{x}'_1 = 0$$

$$C = K R K^{-1}$$

- Given three rotations around linearly independent axes – S, K can be estimated using linear techniques
- Applications – image mosaics

29

Projective transformations in 2D

And what remains invariant

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$$

Projective transformation 8-DOF
(collinearity, cross ratios)

$$H = \begin{bmatrix} a_1 & a_2 & d_1 \\ a_3 & a_4 & d_2 \\ 0 & 0 & 1 \end{bmatrix}$$

Affine transformation 6-DOF
(parallelism, ration of areas, length ratios)

$$H = \begin{bmatrix} sr_{11} & sr_{12} & t_1 \\ sr_{21} & sr_{22} & t_2 \\ 0 & 0 & 1 \end{bmatrix}$$

Similarity transformation 4-DOF
(angles, length ratios)

$$H = \begin{bmatrix} r_{11} & r_{12} & t_1 \\ r_{21} & r_{22} & t_2 \\ 0 & 0 & 1 \end{bmatrix}$$

Rigid Body Motion 3-DOF
(angles, lengths, areas)

30

Example

Images of planes (+ rectification)

$$\lambda_2 \mathbf{x}'_2 = H \lambda_1 \mathbf{x}'_1 \qquad \lambda_2 \mathbf{x}'_2 = H \mathbf{X} \qquad \mathbf{X} = [X, Y, 1]^T$$

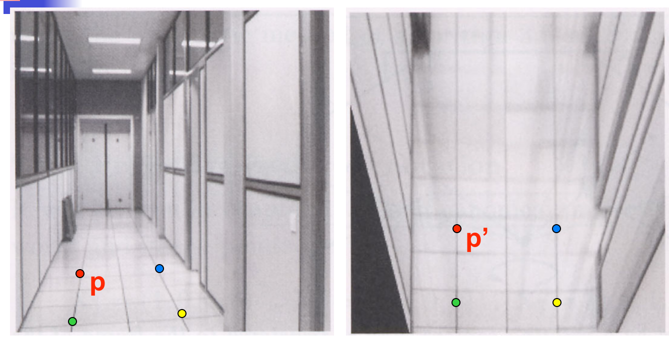
- There is one-to-one mapping between two images of a plane
- or between image plane and world plane

- 2D projective transformation H – homography (3x3 matrix)
- Estimation of homography from point correspondences
  1. eliminate unknown depth

$$\widehat{\mathbf{x}'_2} H \mathbf{x}'_1 = 0$$

  2. get two independent constraints per point – (9-1) unknowns
  3. need at least 4 points to estimate H
  4. H is can be estimated up to a scale factor

Using H



- Image based rectification (given some points in 3D world) compute H which would map them into a square
- Use H to rectify the entire image
- In calibrated case inter-image homography $H = (R + \frac{1}{d} T n^T)$

33

Solving for homographies

$$\begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} \cong \begin{bmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$

$$x'_i = \frac{h_{00}x_i + h_{01}y_i + h_{02}}{h_{20}x_i + h_{21}y_i + h_{22}}$$

$$y'_i = \frac{h_{10}x_i + h_{11}y_i + h_{12}}{h_{20}x_i + h_{21}y_i + h_{22}}$$

$$x'_i(h_{20}x_i + h_{21}y_i + h_{22}) = h_{00}x_i + h_{01}y_i + h_{02}$$
$$y'_i(h_{20}x_i + h_{21}y_i + h_{22}) = h_{10}x_i + h_{11}y_i + h_{12}$$

$$\begin{bmatrix} x_i & y_i & 1 & 0 & 0 & 0 & -x'_i x_i & -x'_i y_i & -x'_i \\ 0 & 0 & 0 & x_i & y_i & 1 & -y'_i x_i & -y'_i y_i & -y'_i \end{bmatrix} \begin{bmatrix} h_{00} \\ h_{01} \\ h_{02} \\ h_{10} \\ h_{11} \\ h_{12} \\ h_{20} \\ h_{21} \\ h_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

34

Solving for homographies

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x'_1 x_1 & -x'_1 y_1 & -x'_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -y'_1 x_1 & -y'_1 y_1 & -y'_1 \\ & & & & \vdots & & & & \\ x_n & y_n & 1 & 0 & 0 & 0 & -x'_n x_n & -x'_n y_n & -x'_n \\ 0 & 0 & 0 & x_n & y_n & 1 & -y'_n x_n & -y'_n y_n & -y'_n \end{bmatrix} \begin{bmatrix} h_{00} \\ h_{01} \\ h_{02} \\ h_{10} \\ h_{11} \\ h_{12} \\ h_{20} \\ h_{21} \\ h_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

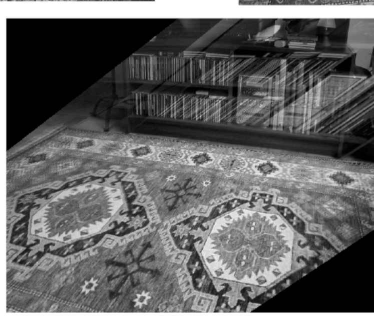
$$||\mathbf{A}\hat{\mathbf{h}}||^2 = (\mathbf{A}\hat{\mathbf{h}})^T \mathbf{A}\hat{\mathbf{h}} = \hat{\mathbf{h}}^T \mathbf{A}^T \mathbf{A}\hat{\mathbf{h}}$$

- LLS solution to solving a system of homogeneous equations
- Solution – eigenvector associated with the smallest eigenvalue of A^TA

Example

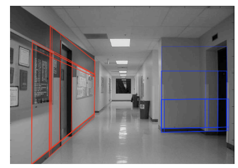
1st view

2nd view



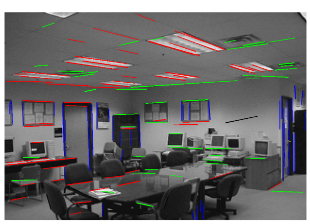
2nd view warped
by the planar homography
between two views

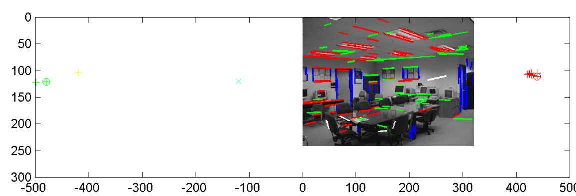
Extraction of rectangular structures and pose recovery



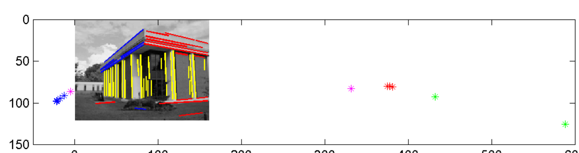
37

Examples





Given a set of line segments, group them based on which vanishing direction then belong to and estimate vanishing points



38

Camera pose recovery

- Assume partially calibrated camera

$$K = \begin{bmatrix} f & \alpha_\theta & o_x \\ 0 & kf & o_y \\ 0 & 0 & 1 \end{bmatrix} \qquad K_f = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- Explicitely parametrize the homography

$$\boldsymbol{x} \sim H\boldsymbol{X}$$

$$\lambda \mathbf{x} = \begin{bmatrix} fr_{11} & fr_{12} & ft_x \\ fr_{21} & fr_{22} & ft_y \\ r_{31} & r_{32} & t_z \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = H \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}.$$

- Explicitely parametrize the unknown structure

$$\mathbf{S} = \begin{bmatrix} 0 & 0 & \alpha b & \alpha b \\ 0 & b & b & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Homography Estimation

- Decouple known and unknown structure

$$\mathbf{S} = \mathbf{S}_\alpha \mathbf{S}_s = \begin{bmatrix} \alpha b & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$\lambda \mathbf{x} = H_\alpha \mathbf{S}_s$$

- Estimate the unknown homography

$$H_\alpha = \begin{bmatrix} \alpha b f r_{11} & b f r_{12} & f t_x \\ \alpha b f r_{21} & b f r_{22} & f t_y \\ \alpha b r_{31} & b r_{32} & t_z \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$$

40

Homography Factorization

- Exploiting orthogonality constraints

$$r_1^T r_2 = 0, \|r_1\| = \|r_2\| = 1$$

• Directly estimate the focal length

$$\frac{1}{\gamma^2}\frac{1}{\alpha}\frac{1}{b^2}\left(\frac{h_{11}h_{12}+h_{21}h_{22}}{f^2} + h_{31}h_{32}\right) = 0.$$

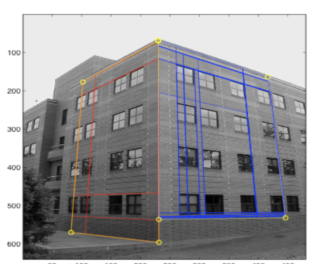
$$\hat{f} = \sqrt{\frac{-h_{31}h_{32}}{h_{11}h_{12}+h_{21}h_{22}}}$$

• remaining parameters and final pose

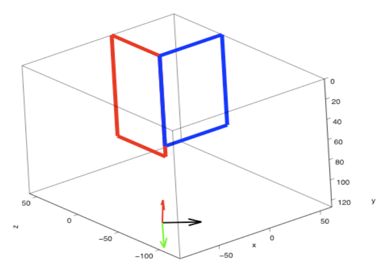
$$H' = \gamma \begin{bmatrix} \alpha b r_{11} & b r_{12} & t_x \\ \alpha b r_{21} & b r_{22} & t_y \\ \alpha b r_{31} & b r_{32} & t_z \end{bmatrix} = \gamma[h'_1, h'_2, h'_3]. \qquad g = \begin{bmatrix} r_{11} & r_{12} & \frac{t_x}{b} \\ r_{21} & r_{22} & \frac{t_y}{b} \\ r_{31} & r_{32} & \frac{t_z}{b} \end{bmatrix}$$

41

Example



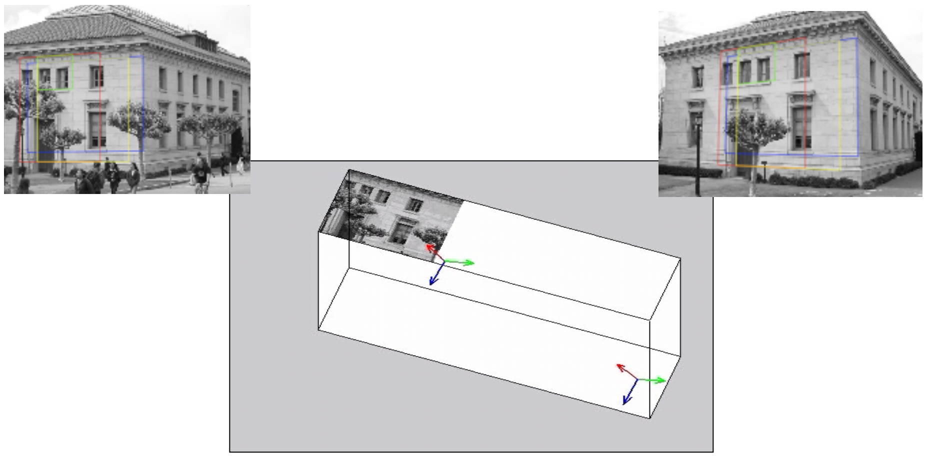
$(R_l, T_l), (R_r, T_l)$



• With shared segment the pose can be reconciled and we obtain single consistent pose recovery up to scale and error ~ 3 degrees

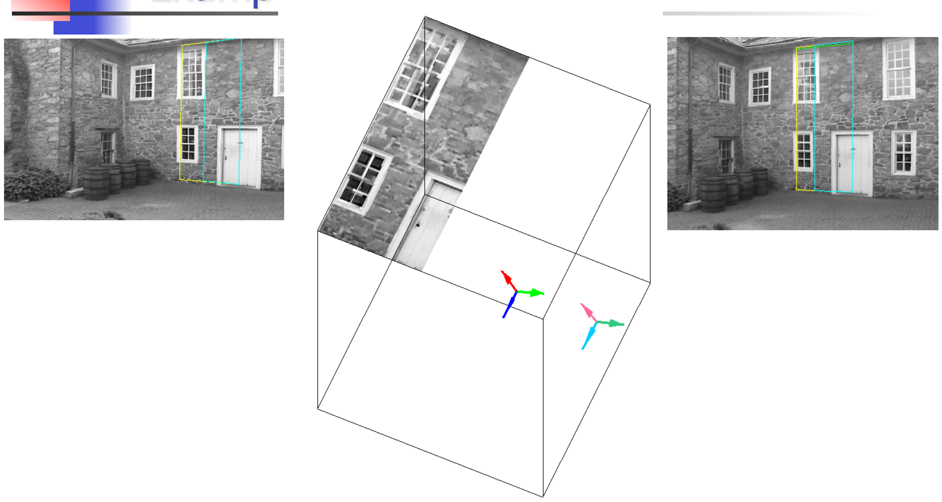
42

Example



- Recovery of the camera displacement from a planar structure 43

Example



- Recovery of the camera displacement from a planar structure
- Dominant plane is decomposed to

44

EM (Expectation Maximization)

Brief tutorial by example:

EM well known statistical technique for estimation of models from data

Set up: Given set of datapoints which were generated by multiple models estimate the parameters of the models and assignment of the data points to the models

Here: set of points in the plane with coordinates *(x,y), two lines* with parameters *(a1,b1)* and *(a2,b2)*

1. Guess the line parameters and estimate error of each point wrt to current model $r_1(i) = a_1 x_i + b_1 - y_i$
2. Estimate Expectation (weight for each point)

$$w_1(i) = \frac{e^{-r_1^2(i)/\sigma^2}}{e^{-r_1^2(i)/\sigma^2} + e^{-r_2^2(i)/\sigma^2}}$$

$$w_2(i) = \frac{e^{-r_2^2(i)/\sigma^2}}{e^{-r_1^2(i)/\sigma^2} + e^{-r_2^2(i)/\sigma^2}}$$

45

EM

Maximization step:

Traditional least squares:

$$\begin{pmatrix} \sum_i x_i^2 & \sum_i x_i \\ \sum_i x_i & \sum_i 1 \end{pmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_i x_i y_i \\ \sum_i y_i \end{bmatrix}$$

Here weighted least squares:

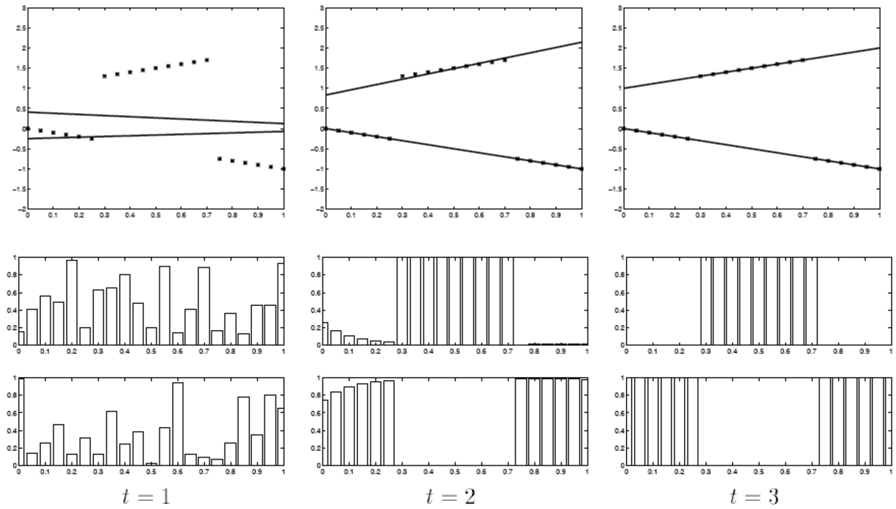
$$\begin{pmatrix} \sum_i w_i x_i^2 & \sum_i w_i x_i \\ \sum_i w_i x_i & \sum_i w_i 1 \end{pmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_i w_i x_i y_i \\ \sum_i w_i y_i \end{bmatrix}$$

Iterate until no change

Problems: local minima, how many models ?

46

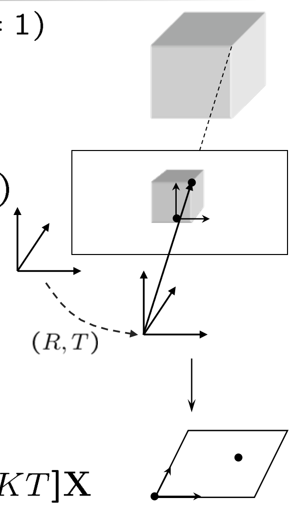
EM - example



47

Uncalibrated Camera

$$\mathbf{X} = [X, Y, Z, W]^T \in \mathbb{R}^4, \quad (W = 1)$$



Calibrated camera

- Image plane coordinates $\mathbf{x} = [x, y, 1]^T$
- Camera extrinsic parameters $g = (R, T)$
- Perspective projection $\lambda \mathbf{x} = [R, T]\mathbf{X}$

Uncalibrated camera

- Pixel coordinates $\mathbf{x}' = K\mathbf{x}$
- Projection matrix $\lambda \mathbf{x}' = \Pi \mathbf{X} = [KR, KT]\mathbf{X}$

48