Image Analysis - Lecture 13

Computer Vision

Magnus Oskarsson

October 11, 2011

Lecture 13

Contents

- What is computer vision?
- Projective geometry
- Pinhole camera
- Two View Geometry

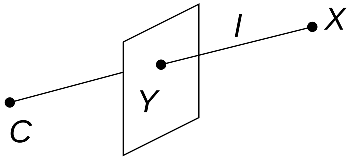
What is Computer Vision?

Can be seen as and attempt at mimicking human (biological) vision.

Some examples of research areas:

- Recognition
- Motion estimation
- 3D reconstruction from 2D images

Pinhole Camera



All points on the line l are mapped to the same point Y.
c is called **camera center**:
The transformation P: X to Y is called **perspective transformation**.
Note: P is not invertible. (The depth is unknown)
Special case: Planar object - P invertible.
Special case: c infinitely far away. P parallel projection (affine transformation)

a) Affine transformations (changes of coordinates) $\mathbb{R}^3 \rightarrow \mathbb{R}^3$ can be written

$$y = Ax + b \ ,$$

where

$$y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}, x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} .$$

This can also be written

$$y = \begin{bmatrix} A & b \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} \quad \Leftrightarrow \quad \begin{bmatrix} y \\ 1 \end{bmatrix} = \begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix}$$

Euclidean transformations

b) **Euclidean transformations** can be written

$$\begin{bmatrix} y \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} \quad ,$$

where R is an orthogonal matrix, i.e.

$$RR^T = R^T R = I \quad .$$

Projective transformations

c) **Projective transformations** are defined as

$$\lambda \begin{bmatrix} y \\ 1 \end{bmatrix} = \begin{bmatrix} A & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} \quad .$$

$$\lambda(y_1, y_2, y_3, 1) = (\lambda y_1, \lambda y_2, \lambda y_3, \lambda)$$

are called **homogeneous coordinates** for

$$(y_1, y_2, y_3) \quad .$$

In the **projective space**, $\mathcal{P}^3$, one identifies

$$(y_1, y_2, y_3, 1) \quad \text{and} \quad (\lambda y_1, \lambda y_2, \lambda y_3, \lambda).$$

These are considered as the same point.

Homogeneous coordinates

An image point (x, y) is represented in homogeneous coordinate as

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

or multiples of this vector, e.g.

$$\begin{pmatrix} 5x \\ 5y \\ 5 \end{pmatrix}$$

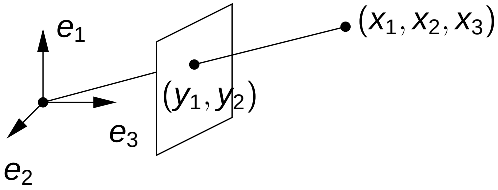
What homogenous coordinates does the image point $(127, 321)$ have? Answer: $\begin{pmatrix} 127 & 321 & 1 \end{pmatrix}^T$.

Which image point does the homogeneous coordinates $\begin{pmatrix}1 & 2 & 3\end{pmatrix}^T$ correspond to? Answer: Image point: $(1/3, 2/3)$. In the same way, a scene point (x, y, z) can be represented in homogeneous coordinates by

$$\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

(Note the similarity with line representations: $\begin{pmatrix}a & b & c\end{pmatrix}^T$. This is the homogeneous coordinate representation of the line $ax + by + c = 0$.)

Perspective transformation



$$y_1 = \frac{x_1}{x_3}, \quad y_2 = \frac{x_2}{x_3}$$

This is a *non linear expression*, which can be written

$$\lambda \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad ,$$

with $\lambda = x_3$.
λ is called the **depth**.

External calibration

Using homogeneous coordinates

$$\lambda \begin{bmatrix} y_1 \\ y_2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} I & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix}$$

Euclidean change of coordinates in the scene gives:

$$\lambda \begin{bmatrix} y_1 \\ y_2 \\ 1 \end{bmatrix} = \begin{bmatrix} I & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1' \\ x_2' \\ x_3' \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \end{bmatrix} \begin{bmatrix} x_1' \\ x_2' \\ x_3' \\ 1 \end{bmatrix}$$

R and t are called **external calibration**.
R describes camera orientation.
$-R^{-1}t$ is called the camera center.

Internal calibration

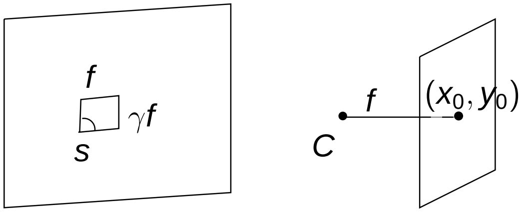
Modeling of the camera gives:

$$\lambda \begin{bmatrix} y_1' \\ y_2' \\ 1 \end{bmatrix} = K \begin{bmatrix} R & t \end{bmatrix} \begin{bmatrix} x_1' \\ x_2' \\ x_3' \\ 1 \end{bmatrix}, \quad K = \begin{bmatrix} f & sf & x_0 \\ 0 & \gamma f & y_0 \\ 0 & 0 & 1 \end{bmatrix}$$

The parameters in K are called **internal parameters**:

- f: **focal length**
- (x_0, y_0): **principal point** (orthogonal projection of the camera center in the image plane.)
- γ: **aspect ratio** (rectangular light sensitive elements instead of quadratic)
- s: **skew** (skewed light sensitive elements)

Calibrated vs uncalibrated camera



- Calibrated camera = K known
- Uncalibrated camera = K unknown
- External orientation = R and t
- Internal orientation = K

If both K, R and t are unknown, then you can rewrite $\lambda \mathbf{y} = P\mathbf{x}$, where P is a 3×4 matrix. Such a matrix P is called a **camera matrix**. This can be written as (in homogeneous coordinates) $\mathbf{y} \sim P\mathbf{x}$.

Camera matrix

The camera matrix P is a 3×4 matrix, that describes camera position, orientation, zoom, etc.
Assume that the camera matrix is

$$P = \begin{bmatrix} 5 & -14 & 2 & 17 \\ -10 & -5 & -10 & 50 \\ 10 & 2 & -11 & 19 \end{bmatrix}$$

and that a 3D-point with coordinates $(0, 3, 2)$ are given. What is the image point?

Camera matrix

The camera matrix P is a 3×4 matrix, that describes camera position, orientation, zoom, etc.
Assume that the camera matrix is

$$P = \begin{bmatrix} 5 & -14 & 2 & 17 \\ -10 & -5 & -10 & 50 \\ 10 & 2 & -11 & 19 \end{bmatrix}$$

and that a 3D-point with coordinates $(0, 3, 2)$ are given. What is the image point?
Answer:The homogeneous coordinates of the 3D point is $X = \begin{bmatrix} 0 & 3 & 2 & 1 \end{bmatrix}^T$. The product PX is $PX = \begin{bmatrix} -21 & 15 & 3 \end{bmatrix}^T$.
This corresponds to the image point $(-21/3, 15/3) = (-7, 5)$.

Camera center

The homogeneous coordinates of the camera center (focal point) C can be calculated by

$$PC = 0.$$

What is the 3D-coordinates of the camera center for the camera matrix above?

Camera center

The homogeneous coordinates of the camera center (focal point) C can be calculated by

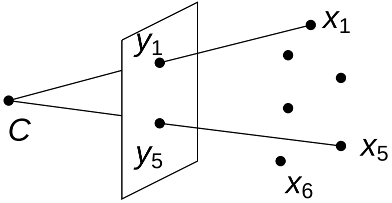
$$PC = 0.$$

What is the 3D-coordinates of the camera center for the camera matrix above?
Answer: The null space of P is given by $C = t\begin{bmatrix}2 & 4 & 6 & 2\end{bmatrix}^T$.
The scene coordinates for the camera center is $(1, 2, 3)$.

Calibration

Example

Assume that we have measured 6 3D points, x_i. By taking image of these points we obtain 6 image points, y_i. How can we calculate K?



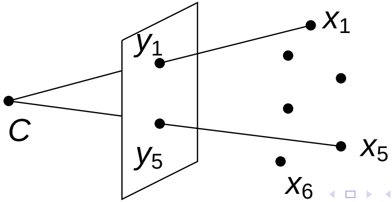
Calibration

Example

Assume that we have measured 6 3D points, x_i. By taking image of these points we obtain 6 image points, y_i. How can we calculate K?

$$\lambda_i \mathbf{y_i} = P\mathbf{x_i}$$

gives 3 equations for every point (linear in λ_i and p_{ij}). We get 18 equations in 18 unknown parameters (12 for P and one for every point). Solve these equations and factorize P as $P = K[\,R\,|\,t\,]$. ■



Internal Calibration

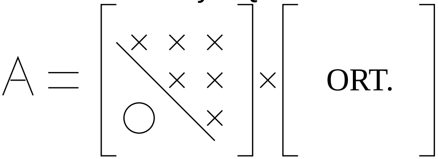
Use:

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = P \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

for 6 points and calculate $P = [\, A \,|\, b \,]$.
P is written as

$$P = K [\, Q \,|\, t \,] = [\, KQ \,|\, Kt \,]$$

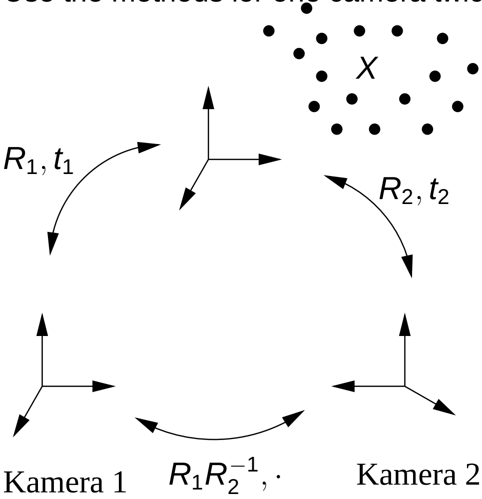
K is obtained by RQ-factorization of A.



Two views: Absolute orientation

Determine the camera matrix using a known object.

Use the methods for one camera twice!



Relative Orientation

Determine the cameras relative motion. Not using a known scene.

Note: Corresponding points in the two images are still needed.

Problem

Relative orientation. Assume that only image correspondences (x_1, x_2) *are given in the two images. Use the equations:*

$$\lambda_1 x_1 = P_1 X$$

$$\lambda_2 x_2 = P_2 X$$

to solve for camera matrices P_1, P_2 *and for 3D coordinates of the point* X.

Fundamental (Essential) matrix

Two projections of the same point:

$$\lambda_1 x_1 = P_1 X$$

$$\lambda_2 x_2 = P_2 X$$

Rewrite as one coordinate system:

$$\underbrace{\begin{pmatrix} P_1 & x_1 & 0 \\ P_2 & 0 & x_2 \end{pmatrix}}_{M} \begin{pmatrix} X \\ -\lambda_1 \\ -\lambda_2 \end{pmatrix} = 0$$

The matrix M has $det(M) = 0$. Discussion: Why?
Expand the determinant (It is linear in each column) in x_1 and x_2 to obtain $det(M) = x_1^T F x_2 = 0$ for some matrix F.

Camera matrix -> fundamental matrix

The elements of F depends on the camera matrices (P_1, P_2). By expanding the determinant $det(M)$ and identifying with $x_1^T F x_2 = 0$, one can show that the elements of F can be written as determinants of matrices that contains some of the rows of the two camera matrices
As an example, element

$$F_{11} = \det \begin{pmatrix} P_{1,21} & P_{1,22} & P_{1,23} & P_{1,24} \\ P_{1,31} & P_{1,32} & P_{1,33} & P_{1,34} \\ P_{2,21} & P_{2,22} & P_{2,23} & P_{2,24} \\ P_{2,31} & P_{2,32} & P_{2,33} & P_{2,34} \end{pmatrix}$$

The mapping $(P_1, P_2) \mapsto F$ is thus relatively simple.

Fundamental matrix -> camera matrix

There are 24 elements in the camera matrices (P_1, P_2).
The scale is arbitrary (-2)
There is freedom in choosing a projective coordinate system (-15)
Thus there remains 7 degrees of freedom
There are 9 elements in the fundamental matrix F.
The scale is arbitrary (-1).
There is an extra constraint $\det(F) = 0$ (-1).
Thus there remains 7 degrees of freedom in F.
It turns out that one can computer the relative motion (P_1, P_2) from F. The mapping is 'almost' unique.
The mapping $F \mapsto (P_1, P_2)$ is thus known.

Solution to the relative orientation problem

Each point correspondence (x_1, x_2) gives a linear constraint on F.

$$x_1^T F x_2 = 0$$

Given at least 7 such correspondences one can calculate F.
>From F one can calculate relative orientation (P_1, P_2).
Using image points (x_1, x_2) and camera matrices (P_1, P_2) one can calculate X (and (λ_1, λ_2)).
Thus one has solved for all unknowns in the equations

$$\lambda_1 x_1 = P_1 X$$

$$\lambda_2 x_2 = P_2 X$$

Essential matrix

A particular case of the fundamental matrix is the case of camera matrices that have been corrected for internal calibration.
There is then 2 additional constraints on the fundamental matrix. Such a matrix is often called essential matrix.

Theorem

An essential matrix E can be written in the form

$$E = R^T T_{\bar{r}_0}$$

and has singular value decomposition $E = U\Sigma V$, with

$$\Sigma = \begin{bmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Epipoles

Definition

The fundamental matrix F defines the **epipolar geometry**. ■

Definition

The image point $e_l = (e_{lx}, e_{ly}, 1)$ with

$$e_l^T F = 0$$

is called the **left epipole**. Similarly the point $e_r = (e_{rx}, e_{ry}, 1)$ with

$$F e_r = 0$$

is called the **right epipole**. ■

Proposition

The left epipole is the projection of the camera center of the right camera and vice versa.

Epipolar lines

Given a point $(x_l', y_l', 1)$, in the left image and the epipolar matrix F.
The epipolar constraint

$$\begin{bmatrix} x_l' & y_l' & 1 \end{bmatrix} F \begin{bmatrix} x_r' \\ y_r' \\ 1 \end{bmatrix} = l^T \begin{bmatrix} x_r' \\ y_r' \\ 1 \end{bmatrix} = 0$$

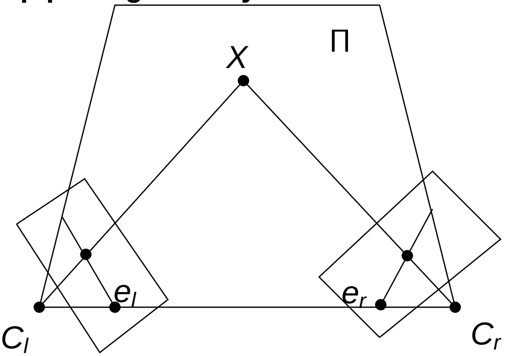
gives that the corresponding point in the right image must lie on a line.

Definition

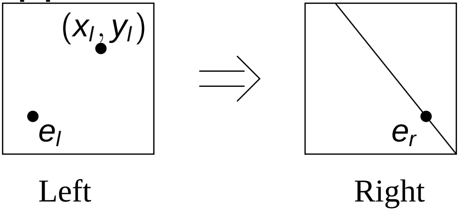
This line is called the **epipolar line** to the point $(x_l', y_l', 1)$. ■

Illustrations

Epipolar geometry:



Epipolar lines:



Fundamental matrix

Assume that we have two images and that their relative orientation is given by the fundamental matrix

$$F = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

Assume that a point in image 1 has coordinates $(2, -3)$. What coordinates can corresponding point have in image 2? Can the point $(-2, 1)$ be the corresponding point in image 2? Where are the left and right epipoles?

Fundamental matrix

Assume that we have two images and that their relative orientation is given by the fundamental matrix

$$F = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

Assume that a point in image 1 has coordinates $(2, -3)$. What coordinates can corresponding point have in image 2? Can the point $(-2, 1)$ be the corresponding point in image 2? Where are the left and right epipoles?

Answer: Homogeneous coordinates for the points are $x_1 = \begin{pmatrix} 2 & -3 & 1 \end{pmatrix}^T$ and $x_2 = \begin{pmatrix} -2 & 1 & 1 \end{pmatrix}^T$. Since the points fulfill the epipolar constraint $x_1^T F x_2 = 0$ they could be corresponding points.

In fact all points in image 2 that fulfill

$$\begin{pmatrix} 2 & -3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = -3x - 3y - 3 = 0$$

i.e. that lie on the line $x + y + 1 = 0$ could be corresponding points.
The epipoles are given by $Fe_2 = 0$ and $e_1^T F = 0$. Calculations give $e_1 = \begin{pmatrix} 1 & -2 & 1 \end{pmatrix}^T$ and $e_2 = \begin{pmatrix} 1 & -2 & 1 \end{pmatrix}^T$

Masters thesis suggestion of the day:

The general problem of calculating 3D shape from images can be simplified considerably e.g. if the images are taken of an object on a rotating plane. The goal of the masters thesis is to see to what extent this process can be automated and what degree of model complexity can be obtained from such a system.

Image sequence -> VRML model

Review - Lecture 11

- Euclidean, affine and projective transformations
- External orientation
- Internal orientation
- Absolute orientation
- Relative orientation
- Fundamental matrix (epipoles, epipolar lines)
- Essential matrix