Geometry in Computer Vision

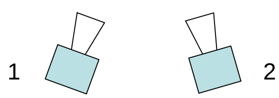
Spring 2010

Lecture 2
Epipolar Geometry

Epipolar geometry

- Epipolar geometry is the geometry related to how *two* cameras (stereo cameras) depict the same scene
- Three or more cameras:
  – Multi-view geometry
- Basic assumptions:
  – Pin-hole cameras
  – All images are taken from different positions
    $\Rightarrow$ The cameras have distinct camera centers

Possible camera configurations

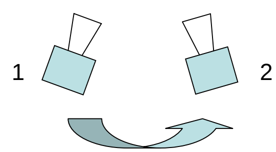


Two camera units

• Possibly with different internal parameters

• Possibly taking their images simultaneously

- Non-static scene is allowed



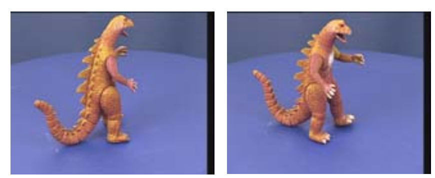
One camera unit that moves from position 1 to position 2

• Image are taken at different time points

- Scene must be static

Examples of camera motion patterns

The camera rotates around the scene



The camera moves along the principal axis

The camera moves “sideways”

Epipolar geometry

Two basic issues of epipolar geometry:

- The *correspondence problem*:
  How can we know if a point in image 1 is the same as some point in image 2?

- The *reconstruction problem*:
  Given that two image points correspond, which 3D point do they refer to?

Basic setup

- Let $\mathbf{C}_1$ and $\mathbf{C}_2$ be the camera matrices of the two cameras
- Let $\mathbf{x}$ be the homogeneous coordinates of a 3D point
- Let $\mathbf{y}_1$ and $\mathbf{y}_2$ be the homogeneous coordinates of the images of $\mathbf{x}$
- Let $\mathbf{n}_1$ and $\mathbf{n}_2$ be the homogeneous coordinates of the camera centers

$$\mathbf{y}_1 \sim \mathbf{C}_1\,\mathbf{x} \qquad \mathbf{y}_2 \sim \mathbf{C}_2\,\mathbf{x}$$

$$\mathbf{C}_1\,\mathbf{n}_1 = \mathbf{0} \qquad \mathbf{C}_2\,\mathbf{n}_2 = \mathbf{0}$$

Pseudo-inverse

- For an $n \times m$ matrix ($n < m$) $\mathbf{A}$ we define its pseudo-inverse $\mathbf{A}^+$ as

$$\mathbf{A}^+ = \mathbf{A}^T(\mathbf{A}\,\mathbf{A}^T)^{-1}$$

- $\mathbf{A}^+$ is $m \times n$ and satisfy $\mathbf{A}\,\mathbf{A}^+ = \mathbf{I}$
- Assumes $\mathbf{A}$ is of rank n

Reprojection line

- If $\mathbf{y}_1$ is known, what can be said about $\mathbf{x}$?
- We known that $\mathbf{x}$ lies somewhere on <u>a 3D line</u>
  - Passes through: $\mathbf{n}_1$
  - Passes through: $\mathbf{C}_1^+\,\mathbf{y}_1$

These two points are always distinct!

Reprojection line

- Parametric representation of the line:

$$(1-t)\,\mathbf{n}_1 + t\,\mathbf{C}_1^+\mathbf{y}_1$$

Here we assume $t \neq 0$

- Check:

$$\mathbf{C}_1\left[(1-t)\,\mathbf{n}_1 + t\,\mathbf{C}_1^+\mathbf{y}_1\right] = t\,\mathbf{y}_1 \sim \mathbf{y}_1$$

Reprojection line

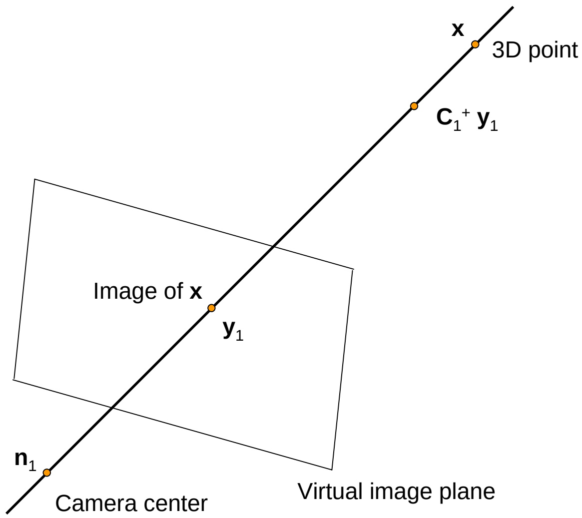


Image of a line

- What is the image of this line in camera 2?
- The parametric 3D point is mapped to $\mathbf{y}'_2(t)$ in image 2:

$$\mathbf{y}'_2(t) \sim \mathbf{C}_2 \left[(1-t)\, \mathbf{n}_1 + t\, \mathbf{C}_1^{+} \mathbf{y}_1 \right]$$

$$\mathbf{y}'_2(t) \sim \underline{(1-t)\, \mathbf{C}_2 \mathbf{n}_1} + \underline{t\, \mathbf{C}_2 \mathbf{C}_1^{+} \mathbf{y}_1}$$

A point in image 2

Another point in image 2

A parameterized line in image 2, passes through both points

Image of a line

- This is a general result:
  - The image of a 3D line is always a 2D line (why?)
- Form 2D line

  $\mathbf{l}_2 = (\mathbf{C}_2\mathbf{n}_1) \times (\mathbf{C}_2\mathbf{C}_1^{+}\, \mathbf{y}_1)$
- Follows: the points $\mathbf{y'}_2(t)$ lie on the 2D line $\mathbf{l}_2$
- Follows: $\mathbf{y'}_2(t) \cdot \mathbf{l}_2 = 0$ for all t
- $\mathbf{y}_2 = \mathbf{y'}_2(t)$ for some $t \Rightarrow \mathbf{y}_2 \cdot \mathbf{l}_2 = 0$

$\mathbf{y}_2$ is the image of $\mathbf{x}$

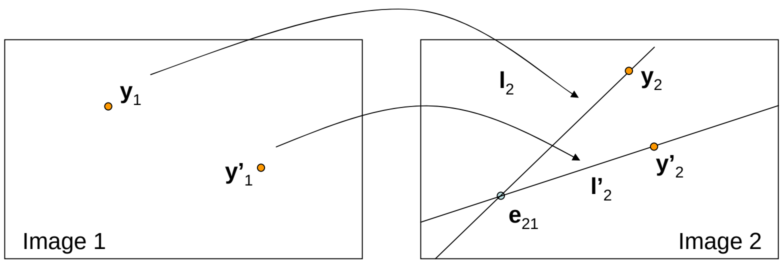
Conclusions

- Given that $\mathbf{y}_1$ in image 1 is known, we know that $\mathbf{y}_2$ lies on a line $\mathbf{l}_2$ in image 2
- The line $\mathbf{l}_2$ depends on $\mathbf{y}_1$
- $\mathbf{l}_2$ is called an *epipolar line*
- All epipolar lines in image 2 intersect the point $\mathbf{e}_{21} = \mathbf{C}_2\mathbf{n}_1$ (why?)
- $\mathbf{e}_{21}$ is called *epipolar point* or just *epipole*
- Symmetry between image 1 and image 2

The image of camera center 1 in image 2

Epipolar lines and points

y_1 and **y_2** correspond to the same 3D point **x**
y'_1 and **y'_2** correspond to the same 3D point **x'**



y_1 generates epipolar line **l_2** in image 2
y'_1 generates epipolar line **l'_2** in image 2

Both epipolar lines intersect at epipolar point **e_{21}**

y_2 lies on **l_2** and **y'_2** lies on **l'_2**

More conclusions

- The mapping from a point $\mathbf{y}_1$ to a line $\mathbf{l}_2$:

$$\mathbf{l}_2 = (\mathbf{C}_2\mathbf{n}_1) \times (\mathbf{C}_2\mathbf{C}_1^+ \mathbf{y}_1)$$

$$\mathbf{l}_2 = [\mathbf{e}_{21}]_\times \mathbf{C}_2\mathbf{C}_1^+ \mathbf{y}_1$$

The cross product operator, see previous lecture

$\mathbf{l}_2$ is given by a "linear mapping" on $\mathbf{y}_1$!

The fundamental matrix

- This mapping is called the *fundamental matrix*, denoted **F**.
- **F** is 3×3

$$\mathbf{l}_2 = \mathbf{F}\, \mathbf{y}_1$$

$$\mathbf{F} = [\mathbf{e}_{21}]_\times \mathbf{C}_2\mathbf{C}_1^+$$

F depends only on the camera matrices $\mathbf{C}_1$ and $\mathbf{C}_2$ ($\mathbf{e}_{21}$ depends on $\mathbf{C}_1$ and $\mathbf{C}_2$)

The epipolar constraint

- If $\mathbf{y}_1$ and $\mathbf{y}_2$ correspond to the same 3D point **x**:

$$\mathbf{y}_2^T\, \mathbf{l}_2 = 0$$

$$\mathbf{y}_2^T\, \mathbf{F}\, \mathbf{y}_1 = 0$$

This relation must always be satisfied for points $\mathbf{y}_1$ and $\mathbf{y}_2$ if they correspond to the same 3D point

Epipolar constraint

The epipolar constraint

- The epipolar constraint is necessary for correspondence (but not sufficient!)

$\mathbf{y}_1$ and $\mathbf{y}_2$ correspond to the same 3D point $\mathbf{x}$

$$\mathbf{y}_2^T \mathbf{F} \mathbf{y}_1 = 0$$

(why not sufficient?)

Summary so far

- Given that $\mathbf{C}_1$ and $\mathbf{C}_2$ are known, **F** can be computed
- Given that **F** is known, we can test if a point in image 1 and a point in image 2 correspond to the same 3D point
- Given a point $\mathbf{y}_1$ in image 1, the corresponding point $\mathbf{y}_2$ lies on an epipolar line $\mathbf{l}_2$ in image 2
- All epipolar lines in image 2 intersect with the epipolar point $\mathbf{e}_{21}$
- $\mathbf{l}_2$ is given by $\mathbf{F}\,\mathbf{y}_1$

Symmetry

- In the previous derivation we started with a point in image 1 that defines an epipolar line in image 2
- Due to symmetry, we can instead start with a point in image 2 and define an epipolar line in image 1

$$\mathbf{l}_1 = \mathbf{F}^T \mathbf{y}_2$$

$$\mathbf{e}_{12} = \mathbf{C}_1 \mathbf{n}_2$$

$$\mathbf{F}^T = [\mathbf{e}_{12}]_\times \mathbf{C}_1 \mathbf{C}_2^+$$

Properties of F

- From $\mathbf{F} = [\mathbf{e}_{21}]_\times \mathbf{C}_2 \mathbf{C}_1^+$

$$\Rightarrow \mathbf{e}_{21}^T \mathbf{F} = \mathbf{e}_{21}^T [\mathbf{e}_{21}]_\times \mathbf{C}_2 \mathbf{C}_1^+ = \mathbf{0}$$

- From symmetry: $\mathbf{F}\,\mathbf{e}_{12} = \mathbf{0}$
- Follows: rank **F** = 2 and det **F** = 0

Internal constraint on **F**

- The epipoles define the left and right null spaces of **F**, respectively
- **F** and α**F** determine the same constraint if $\alpha \neq 0$
  - **F** can be seen as an element of $P^8 = P(R^9)$
- **F** has 7 degrees of freedom (why?)

Epipolar degeneracies

- If the two camera centers coincide

$\mathbf{F} = [\mathbf{e}_{21}]_\times \mathbf{C}_2 \mathbf{C}_1^+ = \mathbf{0}$ (why?)

- From previous lecture we know that in this case

$\mathbf{y}_2 = \mathbf{H}\, \mathbf{y}_1$, where $\mathbf{H}$ is a homography

Epipolar degeneracies

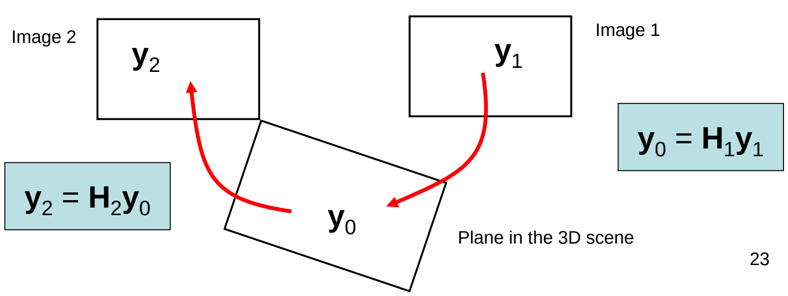
- Follows:

$\mathbf{0} = [\mathbf{y}_2]_\times \mathbf{H}\, \mathbf{y}_1$

- 3 constraints on the two image coordinates!
- They are linearly independent (why?)
- Conclusion:
  - In this case the fundamental matrix is not unique
  - This is flagged by $\mathbf{F}=\mathbf{0}$ when computed from $\mathbf{C}_1$ and $\mathbf{C}_2$
  - $\mathbf{F}$ lies in a 3-dim space of possible solutions to the epipolar constraint

Epipolar degeneracies

- A similar situations occurs when the 3D scene consists points in a plane
- All observations of image points $\mathbf{y}_1$ and $\mathbf{y}_2$ can be written $\mathbf{y}_2 = \mathbf{H}\, \mathbf{y}_1 = \mathbf{H}_2\, \mathbf{H}_1\, \mathbf{y}_1$



Two cases of determining **F**

- The calibrated case:
  - $\mathbf{F}$ is computed from $\mathbf{C}_1$ and $\mathbf{C}_2$
- The uncalibrated case:
  - Given a set of N corresponding image points, $\mathbf{y}_{1k}$ in image 1 and $\mathbf{y}_{2k}$ in image 2, it is possible to determine $\mathbf{F}$ from the constraints:

$$\mathbf{y}_{2k}^T\, \mathbf{F}\, \mathbf{y}_{1k} = 0, \quad k = 1, \ldots, N$$

The uncalibrated case

- No camera matrices need to be known
  - We estimte **F** from image coordinates only
- Image coordinates can only be determined up to a certain accuracy
  - lens distortion
  - quantization to integer pixel coordinates
  - detection inaccuracy
- This accuracy affects the estimation of **F**

Estimation of F

- Let **y** and **y**' be corresponding points in image 1 and image 2 (no image noise!)

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \quad \mathbf{y}' = \begin{pmatrix} y'_1 \\ y'_2 \\ y'_3 \end{pmatrix} \quad \mathbf{F} = \begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix}$$

This can be $\mathbf{y}_1$

This can be $\mathbf{y}_2$

Estimation of F

- The epipolar constraint: $\mathbf{y}'^{T}\,\mathbf{F}\,\mathbf{y} = 0$

$$\mathbf{y}'^{T}\mathbf{F}\,\mathbf{y} = y'_1 y_1 f_{11} + y'_2 y_1 f_{21} + y'_3 y_1 f_{31} + y'_1 y_2 f_{12} + y'_2 y_2 f_{22} + y'_3 y_2 f_{32} + y'_1 y_3 f_{13} + y'_2 y_3 f_{23} + y'_3 y_3 f_{33}$$

Estimation of F

- The epipolar constraint: $\mathbf{Y} \cdot \mathbf{F}_{vec} = 0$

$$\mathbf{Y} = \begin{pmatrix} y'_1 y_1 \\ y'_2 y_1 \\ y'_3 y_1 \\ y'_1 y_2 \\ y'_2 y_2 \\ y'_3 y_2 \\ y'_1 y_3 \\ y'_2 y_3 \\ y'_3 y_3 \end{pmatrix} \quad \mathbf{F}_{vec} = \begin{pmatrix} f_{11} \\ f_{21} \\ f_{31} \\ f_{12} \\ f_{22} \\ f_{32} \\ f_{13} \\ f_{23} \\ f_{33} \end{pmatrix}$$

Vector representation of the 3 × 3 fundamental matrix

The mapping from **F** to $\mathbf{F}_{vec}$ is one-to-one !

Estimation of F

- Conclusion: each pair of corresponding points $\mathbf{y}_{1k}$, $\mathbf{y}_{2k}$ in the two images represents <u>*one linear & homogeneous equation*</u> in $\mathbf{F}_{vec}$

$$\mathbf{Y}_k \cdot \mathbf{F}_{vec} = 0$$

$$\mathbf{Y}_k^T \mathbf{F}_{vec} = 0$$

Estimation of F

- Conclusion: $\mathbf{F}_{vec}$ must satisfy the linear homogeneous equation

$$\mathbf{A}\,\mathbf{F}_{vec} = \mathbf{0} \quad \Rightarrow \quad \mathbf{A}^T\mathbf{A}\,\mathbf{F}_{vec} = \mathbf{0}$$

  where **A** is an $N \times 9$ matrix that contains $\mathbf{Y}_k^T$ for $k = 1, \ldots, N$ in its rows

- $\mathbf{F}_{vec}$ is a right singular vector of **A**, of singular value zero
- Alt: $\mathbf{F}_{vec}$ is an eigenvector of $\mathbf{A}^T\mathbf{A}$, of eigenvalue zero

The basic 8-point algorithm

Given N pairs of corresponding points $\mathbf{y}_{1k}$, $\mathbf{y}_{2k}$

1. Form $\mathbf{Y}_k$ from these pairs for $k = 1, \ldots, N$ and then **A** from all $\mathbf{Y}_k$
2. $\mathbf{F}_{vec}$ = right singular vector of **A**, of singular value zero
3. Reshape $\mathbf{F}_{vec}$ back to a 3×3 matrix **F**.
   This **F** is an estimate of the fundamental matrix

[Longuet-Higgins, *Nature*, 1981]

The 8-point algorithm, practice

- In practice the image coordinates will be perturbed by noise
  - Geometric distortion
  - Coordinate quantization
  - Estimation noise
- Corresponding image coordinates do not satisfy the epipolar constraint exactly

ˇ E d g w k l q j v can happen
  - The estimated **F** may not satisfy the int. contr.
    - ⇒ Epipolar points are not well-defined
    - ⇒ The epipolar geometry is not well-defined

Enforcement of the internal constraint

- If det **F** ≠ 0, we can enforce its internal constraint:
  - Make the smallest possible change in **F** to $\mathbf{F}_0$ (in Frobenius norm) such that det $\mathbf{F}_0 = 0$
- Ho to do this:

An SVD of **F** gives: $\mathbf{F} = \mathbf{U}\,\mathbf{S}\,\mathbf{V}^T$

det **F** = $\pm\, \sigma_1 \cdot \sigma_2 \cdot \sigma_3$

S is diagonal, holding the singular values $\sigma_1 \geq \sigma_2 \geq \sigma_3 > 0$

Set $\mathbf{S}_0 = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ and $\mathbf{F}_0 = \mathbf{U}\,\mathbf{S}_0\,\mathbf{V}^T$

The smallest singular value is set to zero

The 8-point algorithm, practice

If *N*=8, then **F** is well-defined from **A** $\mathbf{F}_{vec}$ = **0**

This is why it is called the 8-point algorithm

Assumes that we don't have degeneracies

- This F satisfies the epipolar constraint for the 8 corresponding point pairs.
- However, for *N* > 8 and noisy image points **A** $\mathbf{F}_{vec}$ = **0** does not have a well-defined solution

The 8-point algorithm, practice

- We can, for example, obtain a total least squares estimate:

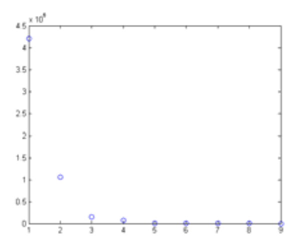
Get **F** from the $\mathbf{F}_{vec}$ that is the right singular vector of **A** corresponding to the <u>smallest</u> singular value of **A**

Equivalently: find $\mathbf{F}_{vec}$ with $||\mathbf{F}_{vec}||=1$ that minimizes $||\mathbf{A}\,\mathbf{F}_{vec}||$

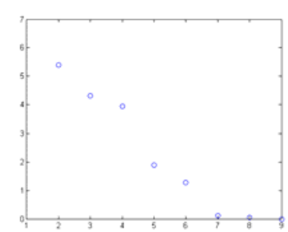
(why?)

Hartley normalization

Distribution of singular values from **A**
An example based on real data



Without Hartley-normalization



With Hartley-normalization

Small perturbations in image coordinates are likely to cause large changes in the singular vectors corresponding to small singular values

Hartley normalization

- Hartley analyzed the numerical sensitivity of the 8-point algorithm and devised a solution: *Hartley-normalization* of the image coordinates
- Transform the coordinate system of each image independently such that
  – The origin is the centroid of the image points
  – The mean distance to the origin = $2^{1/2}$

(why?)

[Hartley, *In defense of the 8-point algorithm*, PAMI, 1997]

Hartley normalization

Consequently:

Whenever we want estimate geometric objects based on total least squares:

1. Transform all image point to Hartley-normalized coordinates (translation and scaling)
2. Estimate geometric object (e.g. **F**)
3. Transform the object back to standard coordinates

HZ: Hartley-normalization is not optional, it is often required to get useful results

The normalized 8-point algorithm

Putting all this into one single algorithm gives:

1. Start with $N \geq 8$ corresponding points in the two images: $\mathbf{y}_{1k}$ and $\mathbf{y}_{2k}$ with $k = 1, \ldots, N$
2. In each image: transform the coordinates to Hartley normalized form: $\mathbf{y}'_{1k} = \mathbf{H}_1\, \mathbf{y}_{1k}$ and $\mathbf{y}'_{2k} = \mathbf{H}_2\, \mathbf{y}_{2k}$
3. Build the $N \times 9$ data matrix **A**' from $\mathbf{y}'_{1k}$ and $\mathbf{y}'_{2k}$
4. Find $\mathbf{F}'_{vec}$ as the singular vector of smallest singular value relative **A'**
5. Reshape $\mathbf{F}'_{vec}$ to the 3×3 matrix **F**'
6. Enforce the internal constraint on **F**' to get $\mathbf{F}'_0$
7. Transform back to original coordinate system:

$$\mathbf{F} = \mathbf{H}_2^T\, \mathbf{F}'_0\, \mathbf{H}_1$$

(why is this correct?)

Estimation of **F**: Algebraic minimization

- When F is estimated from the normalized 8-point algorithm:
  – The initial estimate is guaranteed to minimize the *algebraic error* $||\mathbf{A}\,\mathbf{F}_{vec}||$ with $||\mathbf{F}_{vec}|| = 1$
  – We then enforce the internal constraint
  – This, in general, increases the algebraic error
- Can we find **F** that satisfies its internal constraint and minmizes the algebraic error?
- An iterative algoritm exists for doing this (HZ)
- Uses **F** from N8PA as initial estimate
- In general, gives a better estimate for **F**

The 7-point algorithm

- N8PA is based on using $N \geq 8$ epipolar constraints to estimate **F**
- We may also use the internal constraint + 7 epipolar constraints to determine **F**

1. $\mathbf{A}\,\mathbf{F}_{vec} = \mathbf{0} \Rightarrow$ A 2-dim solution space for $\mathbf{F}_{vec}$
2. determine up to 3 unique solutions for **F** in this solution space using the internal constraint (how?)

• Only 7 point correspondences are needed to determine **F**
• They meet the internal constraint automatically

• Up to 3 possible solutions, but only 1 is correct
• All 3 solutions must be treated as correct

Break

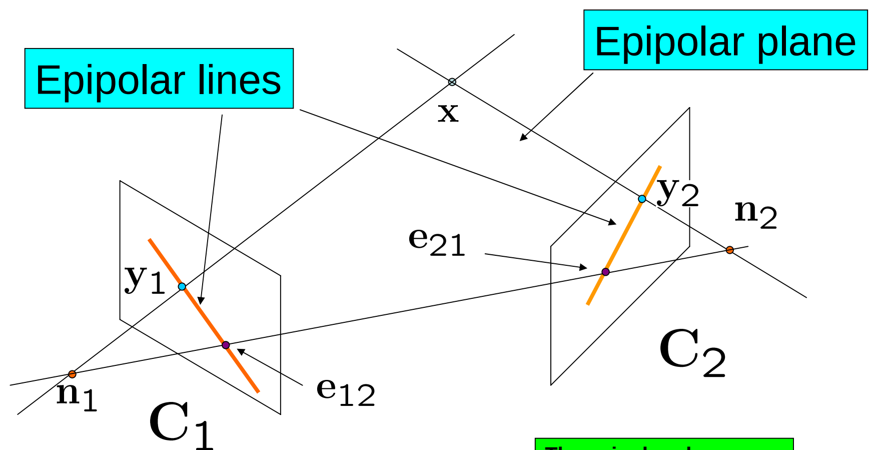
Epipolar line transfer

- In epipolar geometry we cannot map $\mathbf{y}_1$ directly to its corresponding point $\mathbf{y}_2$
- We can however map $\mathbf{y}_1$ to an epipolar line $\mathbf{l}_2$, that intersects $\mathbf{y}_2$ (or $1 \leftrightarrow 2$)
- All points $\mathbf{y}_1$ that are mapped to the same epipolar line $\mathbf{l}_2$ lie on the same epipolar line $\mathbf{l}_1$ (why?)
- $[\mathbf{e}_{12}]_\times \mathbf{l}_1$ is a point on epipolar line $\mathbf{l}_1$ (why?)
- Then $\mathbf{F}^T[\mathbf{e}_{12}]_\times \mathbf{l}_1$ is the corresponding epipolar line $\mathbf{l}_2$ (previous result!)

$\mathbf{l}_2 = \mathbf{F}^T[\mathbf{e}_{12}]_\times \mathbf{l}_1$

$\mathbf{l}_1 = \mathbf{F}\,[\mathbf{e}_{21}]_\times \mathbf{l}_2$

Epipolar lines and plane



The epipolar plane is define from either of the two epipolar lines

Special cases of **F**

- In some practical cases the two cameras $\mathbf{C}_1$ and $\mathbf{C}_2$ are, in fact, the same camera that moves in 3D space
- Special cases of the camera motion corresponds to special cases of F
  - Pure translation
  - Planar motion
- Both cases assume that internal camera parameters are constant!

Pure translation

- As long as the translation is ≠ 0, the two epipoles are well-defined
  - But may be points at infinity
- In the case of pure translation

$\mathbf{e}_{12} \sim \mathbf{e}_{21}$

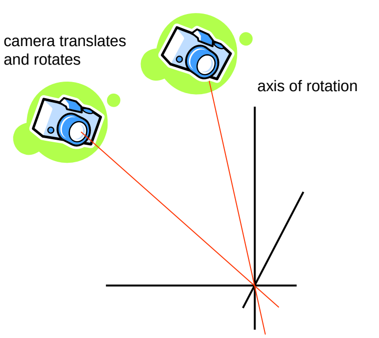
$\mathbf{F} = [\mathbf{e}_{12}]_x = [\mathbf{e}_{21}]_x$ (why?)

F has 2 d.o.f.

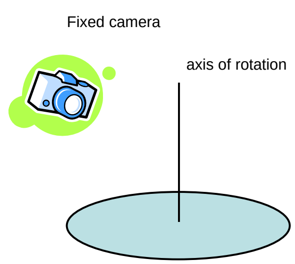
- Example:
  "horizontal" translation $\mathbf{F} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$

Planar motion

- The camera translation is <u>perpendicular</u> to the rotation axis



Case 1: the camera moves



Case 2: rotation table

Planar motion

- Both cases are equivalent
- The rotation axis is *invariant*:
  - The image of a point on this axis must be the same in the two images
  - The image of the rotation axis is a line **l** in both images
- $\mathbf{F} = [\mathbf{e}_{12}]_\times [\mathbf{l}]_\times [\mathbf{e}_{21}]_\times$ (why?)

F has 6 d.o.f.

Cameras from **F**

- Given that $\mathbf{C}_1$ and $\mathbf{C}_2$ are known, **F** can be determined
- What about the outer way around?
- $\mathbf{C}_1$ and $\mathbf{C}_2$ can be determined but not uniquely

- With **F** known $\Rightarrow$ $\mathbf{e}_{12}$ and $\mathbf{e}_{21}$ are known

Canonical cameras from **F**

- It is straight-forward to show that

$$\mathbf{C}_1 = \left(\, \mathbf{I} \;\middle|\; \mathbf{0} \,\right)$$

$$\mathbf{C}_2 = \left(\, [\mathbf{e}_{12}]_\times \mathbf{F} + \mathbf{e}_{12}\mathbf{v}^T \;\middle|\; \lambda \mathbf{e}_{12} \,\right)$$

satisfy $\mathbf{F} = [\mathbf{e}_{21}]_\times \, \mathbf{C}_2 \mathbf{C}_1^+$

for arbitrary $\mathbf{v} \in R^3$, $\lambda \in R$, $\lambda \neq 0$

General cameras from **F**

- However, these cameras are not unique:
- Take $\mathbf{C}_1$ and $\mathbf{C}_2$ such that

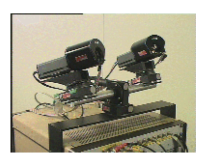
$$\mathbf{F} = [\mathbf{e}_{21}]_\times \, \mathbf{C}_2 \mathbf{C}_1^+$$

- Then $\mathbf{C}'_1 = \mathbf{C}_1\mathbf{H}$ and $\mathbf{C}'_2 = \mathbf{C}_2\mathbf{H}$ also give the same **F** for any 3D homography transformation **H** (why?)

Stereo rig

- A general stereo rig consists of two cameras with
  - distinct camera centers
  - general orientations of the camera principal axes (although often toward a common scene!)

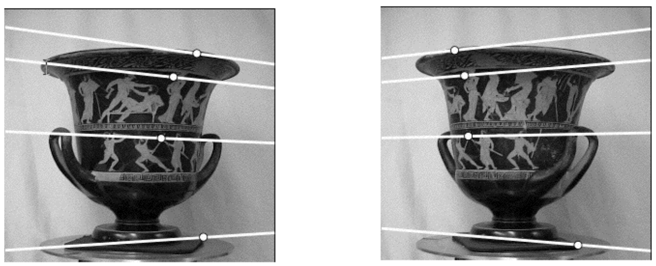
Research stereo rig, Aalborg University



Point Grey, Bumblebee stereo cameras

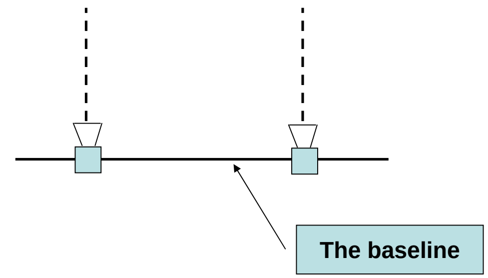
For a general stereo rig

- In each image: the epipolar lines may not be parallel
  - Instead they intersect at the epipolar point that is a real point



Rectified stereo rig

- In a *rectified stereo rig*, the principal directions of the cameras are parallel and orthogonal to the baseline and the cameras are identical



Rectified stereo images

- The rectified stereo rig produces images where
  - The epipolar lines are parallel
  - The epipolar points are points at infinity
- More precisely:

$$\mathbf{e}_{12} = \mathbf{e}_{21} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

This is a point infinitely far away on the horizontal axis

In a coordinate system where first coordinate: right second coordinate: down

Rectified stereo images

- The corresponding fundamental matrix is

$$\mathbf{F}_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

- Note that $\mathbf{y}_2^T \mathbf{F}_0 \mathbf{y}_1 = 0$ for

$$\mathbf{y}_1 = \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \qquad \mathbf{y}_2 = \begin{pmatrix} u + d \\ v \\ 1 \end{pmatrix}$$

Some point in image 1

Same point in image 2, displaced horizontally by d

Rectified stereo rig

- Although a rectified stereo rig can be accomplished by means of accurate measurements, cameras, and mechanics
  - It is difficult and expensive to accomplish the necessary mechanical accuracy
- At best we can set up an approximate rectified stereo rig

Rectified stereo rig

We know

- All cameras that have the same camera center are equivalent ⇒
- If a camera rotates around its camera center, the image transforms according to a homography **H**

Rectified images

Consequence:

- If the principal axis of a camera is not exactly pointing in the right direction, this can be compensated for by applying a suitable homography **H** on the image coordinates
  - This makes the epipolar lines parallel
  - Independent **H** in each image
- The result are *rectified images*

Image rectification

- How do we determine $\mathbf{H}_1$ for image 1 and $\mathbf{H}_2$ for image 2 so that both images are rectified ($\mathbf{H}_1$, $\mathbf{H}_2$ are homographies)?
- Estimate **F** from corresponding points in the two images
  - The 8-point algorithm
- Find $\mathbf{H}_1$, $\mathbf{H}_2$ such that $\underline{(\mathbf{H}_2^{-1})^T \mathbf{F} \mathbf{H}_1^{-1}} \sim \mathbf{F}_0$

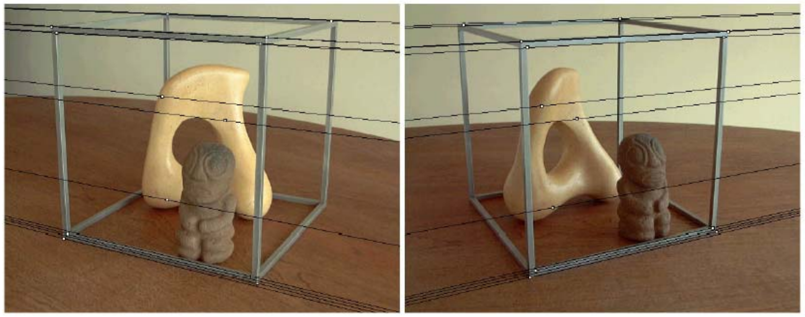
This is the fundamental matrix after transformation of both coordinate systems

Image rectification

- This relation in $\mathbf{H}_1$ and $\mathbf{H}_2$ has multiple solutions, some of which are unwanted
  - Ex: horizontal mirroring
  - Extreme geometric distortion
- Several methods for determining "good" $\mathbf{H}_1$ and $\mathbf{H}_2$ from $\mathbf{F}$ exist, for example:
  - Loop & Zhang, ICPR 1999
  - Determines $\mathbf{H}_1$ and $\mathbf{H}_2$ based on minimization of geometric distortion
  - A similar idea is explored in HZ

Image rectification

Example of a stereo image pair



Black lines are epipolar lines. Not parallel

From Loop & Zhang

Image rectification

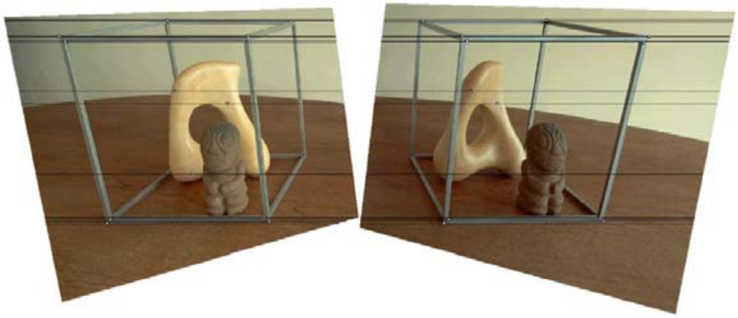
Example of rectification

Epipolar lines are parallel and aligned!

From Loop & Zhang

Image rectification

Another example, less geometric distortion than previous one



Epipolar lines are parallel and aligned!

From Loop & Zhang

Stereo image rectification, summary

- A stereo image pair that are approximately rectified
  - the principal axes are parallel and perpendicular to the baseline
- can be rectified by homography transformations such that
  - corresponding points are found on the same row
- Multiple solutions to the rectification exist

Reconstruction

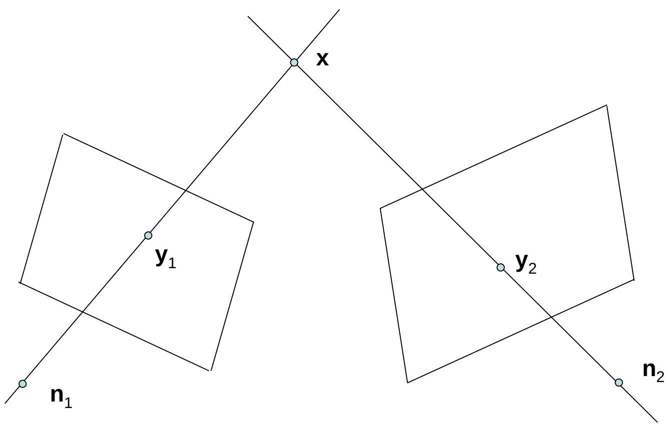
- Given a pair of corresponding image points $\mathbf{y}_1$ and $\mathbf{y}_2$

$$\mathbf{y}_1 \sim \mathbf{C}_1\,\mathbf{x}$$
$$\mathbf{y}_2 \sim \mathbf{C}_2\,\mathbf{x}$$

we know that: $\mathbf{y}_2^T\,\mathbf{F}\,\mathbf{y}_1 = 0$

- What about $\mathbf{x}$? Can $\mathbf{x}$ be determined?
- This problem is called *triangulation* or *reconstruction*

Reconstruction



Epipolar constraint satisfied ⇔ Reprojection lines intersect

In this case: there is a well-defined x that projects to $\mathbf{y}_1$ and $\mathbf{y}_2$

Reconstruction

- In reality, the image points $\mathbf{y}_1$ and $\mathbf{y}_2$ don't satisfy $\mathbf{y}_2^T\,\mathbf{F}\,\mathbf{y}_1 = 0$ exactly
  - Lens distortion
  - Coordinate quantization
  - Estimation inaccuracy
- The two reprojection lines don't intersect

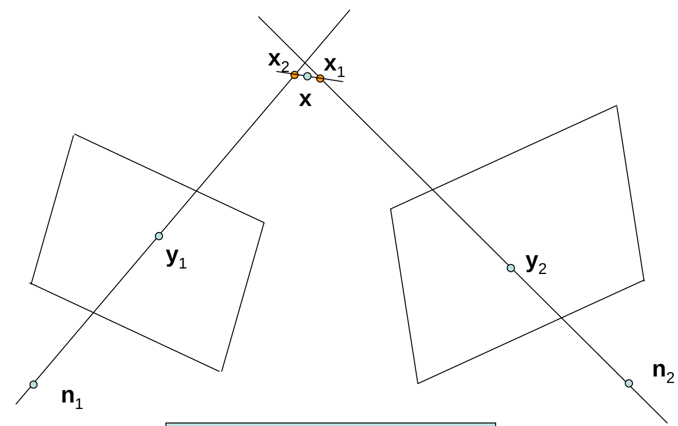
  In this case: $\mathbf{x}$ is not well-defined

  ⇒ It somehow has to be approximated

The mid-point method

- Find the unique points $\mathbf{x}_1$ and $\mathbf{x}_2$ on each reprojection line that is closest to the other line
- Draw a line between $\mathbf{x}_1$ and $\mathbf{x}_2$
- Set $\mathbf{x}$ = the mid-point between $\mathbf{x}_1$ and $\mathbf{x}_2$ on this line
- $\mathbf{x}_1$ and $\mathbf{x}_2$ are identical $\Leftrightarrow$ $\mathbf{y}_2^T \mathbf{F} \mathbf{y}_1 = 0$

The mid-point method



$\mathbf{y}_1$ and $\mathbf{y}_2$ approximately satisfy the epipolar constraint

Linear methods

From

$$\mathbf{y}_1 \sim \mathbf{C}_1\,\mathbf{x}$$
$$\mathbf{y}_2 \sim \mathbf{C}_2\,\mathbf{x}$$

follows

$$0 = \mathbf{y}_1 \times \mathbf{C}_1\,\mathbf{x}$$
$$0 = \mathbf{y}_2 \times \mathbf{C}_2\,\mathbf{x}$$

or

$$0 = [\mathbf{y}_1]_\times \mathbf{C}_1\,\mathbf{x}$$
$$0 = [\mathbf{y}_2]_\times \mathbf{C}_2\,\mathbf{x}$$

3 linear homogeneous equations in x

3 more linear homogeneous equations in x

Linear methods

Since $[\mathbf{y}_1]_\times$ has rank 2: one of the 3 equations is linearly dependent to the other two:

$$0 = [\mathbf{y}_1]_\times \mathbf{C}_1\,\mathbf{x}$$
$$0 = [\mathbf{y}_2]_\times \mathbf{C}_2\,\mathbf{x}$$

In total: 4 linear independent homogeneous equations in x

This can be written

$\mathbf{B}\,\mathbf{x} = \mathbf{0}$

$\mathbf{B}$ is a 6×4 matrix

Linear methods

- In theory **B** has rank 3 and **x** is well-defined
- In practice (with noise) **B x** = **0** cannot be solved exactly
- Solution (for example): determine **x** that minimizes

||**B x**|| with ||**x**||=1

Total least squares minimzation

- **x** = The right singular vector of **B** with smallest singular value
- What about Hartley-normalization?
- *Homogeneous* triangulation

Linear methods

Alternatively: we know that $\mathrm{x} = \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = \begin{pmatrix} \bar{\mathrm{x}} \\ 1 \end{pmatrix}$

B x = **0** can then be rewritten as $\mathrm{B}_1 \, \bar{\mathrm{x}} = \mathrm{b}_0$

Which is solved using standard methods

- *Inhomogeneous* triangulation

Properties of triangulation methods

- In the ideal case any triangulation method gives the same **x** for any $\mathbf{y}_1$, $\mathbf{y}_2$ that satisfy the epipolar constraint, but
- If the constraint is not satisfy the results differ
- The metods have slightly different computational complexity (SVD, iterative, etc)
- Singularities (e.g. the inhomogeneous method fails for 3D points at infinity) (why?)

Invariance to 3D transformations

- Does the resulting **x** change if we change the 3D coordinate system?
  - the mid-point method is only invariant to translations, rotations, and scalings
  - The inhomogeneous method is only invariant to affine transformations
  - The homogeneous method is invariant only to 3D homographies $\in$ SO(4)

Optimal triangulation

- Assume that **F** is known (or estimated)
- Assume that $\mathbf{y}_1$ and $\mathbf{y}_2$ have been perturbed by noise of isotropic distribution
- Find $\mathbf{y'}_1$ and $\mathbf{y'}_2$ such that

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

d is the Euclidean distance in the image (in pixels)

$d(\mathbf{y}_1,\mathbf{y'}_1)^2 + d(\mathbf{y}_2,\mathbf{y'}_2)^2$ is minimal

Optimal triangulation Maximum Likelihood Estimation

- These $\mathbf{y'}_1$ and $\mathbf{y'}_2$ are then *Maximum Likelihood estimates* of $\mathbf{y}_1$ and $\mathbf{y}_2$ that also satisfy the epipolar constraint
- Once $\mathbf{y'}_1$ and $\mathbf{y'}_2$ are determined: use any of the previous methods to determine **x**
- A computational method exist for finding $\mathbf{y'}_1$ and $\mathbf{y'}_2$
  - Involves solving a 6th order polynomial
  - All 6 roots must be evaluated
- Invariant to any 3D homography transformation

- [Hartley & Sturm, *Optimal Triangulation*, CVIU 1997]

The triangulation tensor

- It is also possible to compute x as

$$\mathbf{x} \sim \mathbf{K}\,\mathbf{Y}$$

where $\mathbf{Y} = \mathbf{y}_1\mathbf{y}_2^T$ reshaped to a 9-dim vector

- **K** is a 4×9 matrix (or $4 \times 3 \times 3$ tensor), the *triangulation tensor*

The triangulation tensor

- Low computational complexity
- Invariant to 3D homography transformations
- **K** can be estimated from 3D+2D+2D correspondences
  - No need for camera matrices
- Can then be optimized relative to arbitrary error measures (in 2D, in 3D, L_1, L_2)
- Has a singularity on an arbitrary plane that intersects the camera centers, the *blind plane*

- [Nordberg, *The triangulation tensor*, CVIU, 2009]