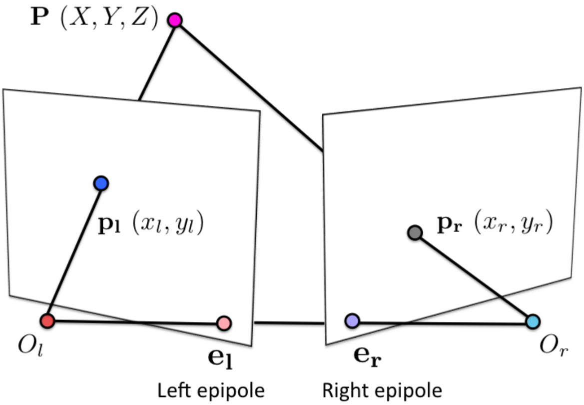
Stereo – General Case

Stereo: General Calibrated Cameras

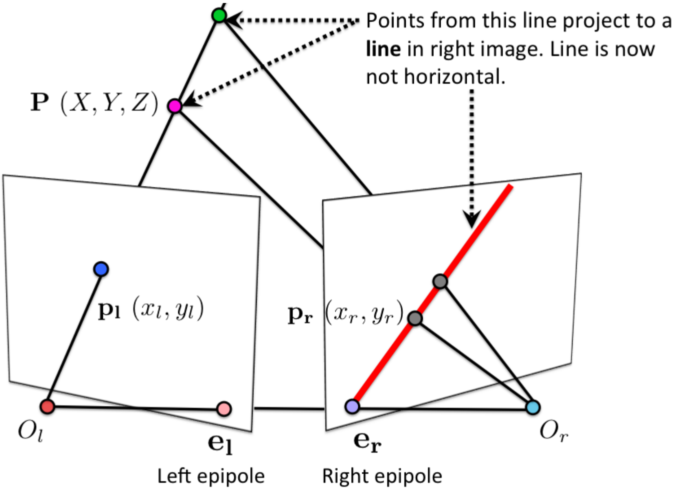
- Some notation: the **left** and **right epipole**



Where line $O_l O_r$ intersects the image planes

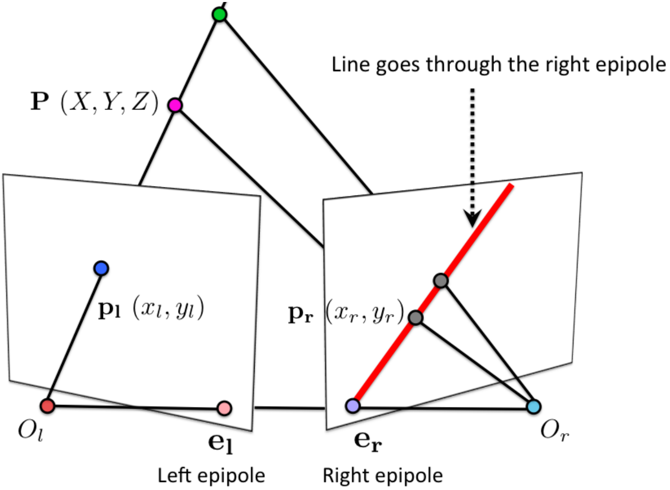
Stereo: General Calibrated Cameras

- All points from the projective line $O_l p_l$ project to a line on the right image plane. This time the line is not (necessarily) horizontal.



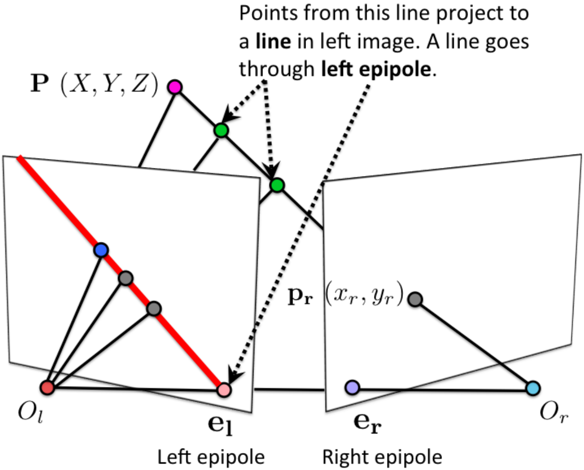
Stereo: General Calibrated Cameras

- The line goes through the right epipole.



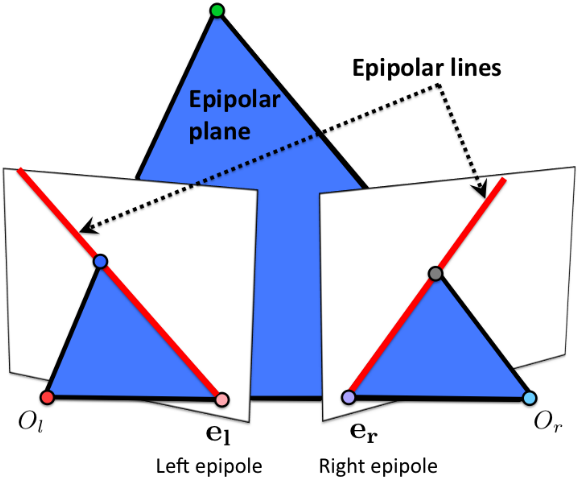
Stereo: General Calibrated Cameras

- Similarly, All points from the projective line $O_r p_r$ project to a line on the left image plane. This line goes through the left epipole.



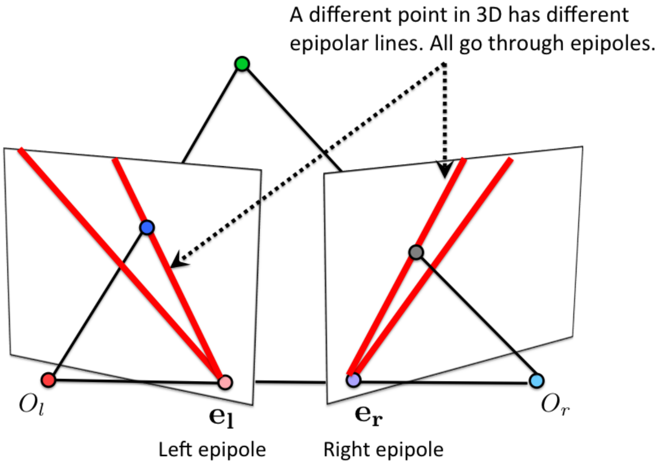
Stereo: General Calibrated Cameras

- The reason for all this is simple: points $\mathrm{O_l}$, $\mathrm{O_r}$, and a point P in 3D lie on a plane. We call this the **epipolar plane**. This plane intersects each image plane in a line. We call these lines **epipolar lines**.



Stereo: General Calibrated Cameras

- Obviously a different point in 3D will form a different epipolar plane and therefore different epipolar lines. But these epipolar lines go through epipoles as well.

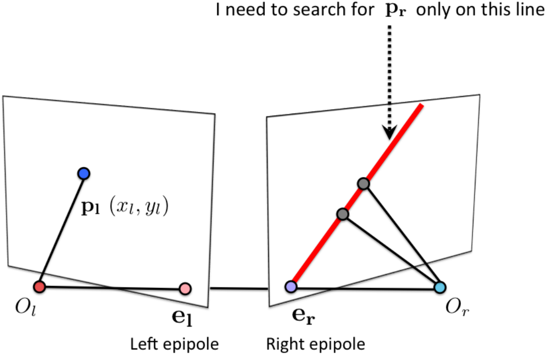


Stereo: General Calibrated Cameras

- Why are we even dumping all this notation? Are epipolar lines, epipoles, etc somehow useful?

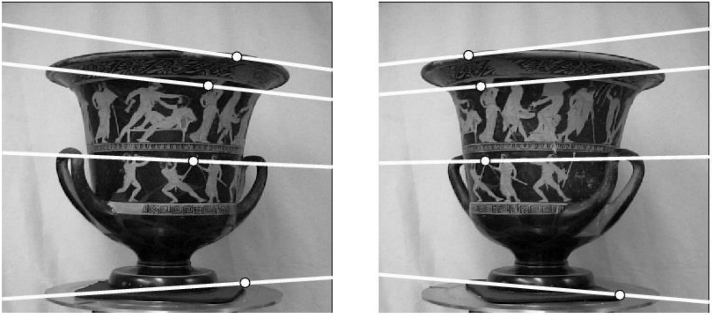
Stereo: General Calibrated Cameras

- Remember what we did for parallel cameras? We were matching points in the left and right image, giving us a point in 3D. We want the same now.
- Epipolar geometry is useful because it constrains our search for the matches:
  - For each point p_l we need to search for p_r only on a epipolar line (much simpler than if I need to search in the full image)
  - All matches lie on lines that intersect in epipoles. This gives another constraint.



Epipolar geometry: Examples

- Example of epipolar lines for converging cameras. How did they get outside of the image?



[Source: J. Hays, pic from Hartley & Zisserman]

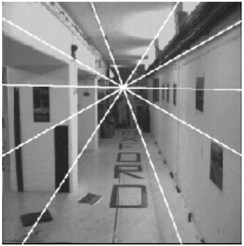
Epipolar geometry: Examples

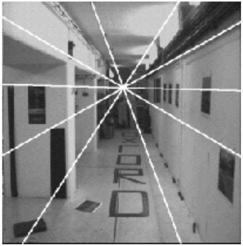
- How would epipolar lines look like if the camera moves directly forward?

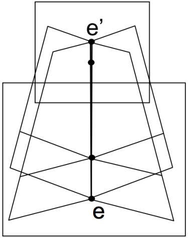
[Source: J. Hays]

Epipolar geometry: Examples

- Example of epipolar lines for **forward motion**







Epipole has same coordinates in both images.
Points move along lines radiating from e: "Focus of expansion"

[Source: J. Hays, pic from Hartley & Zisserman]

Stereo for General Cameras

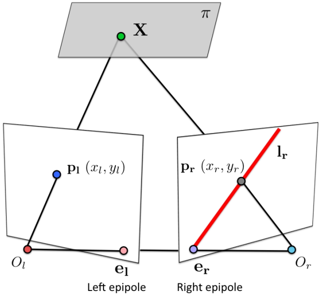
How we'll get 3D:

- We first need to figure out on which line we need to search for the matches for each p_l
- Each point in left image maps to a line in right image. We will see that this mapping can be described by a single 3×3 matrix F, called the **fundamental matrix**
- Given F, you can **rectify** the images such that the epipolar lines are horizontal
- And we know how to take it from there

The Fundamental Matrix

- The fundamental matrix F is defined as $l_r = Fp_l$, where l_r is the right epipolar line corresponding to p_l.
- F is a 3×3 matrix
- For any point p_l its epipolar line is defined by the **same matrix** F.

The Fundamental Matrix



- Extend the line $\mathsf{O_l p_l}$ until you hit a plane π (arbitrary)
- Find the image $\mathsf{p_r}$ of X in the right camera
- Get epipolar line $\mathsf{l_r}$ from $\mathsf{e_r}$ to $\mathsf{p_r}$: $\mathsf{l_r} = \mathsf{e_r} \times \mathsf{p_r}$
- Points $\mathsf{p_l}$ and $\mathsf{p_r}$ are related via homography: $\mathsf{p_r} = H_\pi \mathsf{p_l}$
- Then: $\mathsf{l_r} = \mathsf{e_r} \times \mathsf{p_r} = \mathsf{e_r} \times H_\pi \mathsf{p_l} = \mathsf{F} \mathsf{p_l}$
- The fundamental matrix F is defined $\mathsf{l_r} = \mathsf{F} \mathsf{p_l}$

[Adopted from: R. Urtasun]

The Fundamental Matrix

- The fundamental matrix F is defined as $\mathsf{l_r} = \mathsf{F}\mathsf{p_l}$, where $\mathsf{l_r}$ is the right epipolar line corresponding to $\mathsf{p_l}$.
- F is a 3×3 matrix
- For any point $\mathsf{p_l}$ its epipolar line is defined by the **same matrix** F.
- Do a trick:

$$\mathsf{p_r}^T \cdot \mathsf{l_r} = \mathsf{p_r}^T \mathsf{F} \mathsf{p_l}$$

The Fundamental Matrix

- The fundamental matrix F is defined as $l_r = Fp_l$, where l_r is the right epipolar line corresponding to p_l.
- F is a 3×3 matrix
- For any point p_l its epipolar line is defined by the **same matrix** F.
- Do a trick:

$$\underbrace{{p_r}^T \cdot l_r}_{=0,\text{ because } p_r \text{ lies on a line } l_r} = {p_r}^T F p_l$$

The Fundamental Matrix

- The fundamental matrix F is defined as $l_r = Fp_l$, where l_r is the right epipolar line corresponding to p_l.
- F is a 3×3 matrix
- For any point p_l its epipolar line is defined by the **same matrix** F.
- So:

$$\boxed{{p_r}^T F p_l = 0}$$

  for any match (p_l, p_r) (main thing to remember)!!
- We can compute F from a few correspondences. How do we get these correspondences?
- By finding reliable matches across two images without any constraints. We know how to do this from our DVD matching example (e.g. SIFT).
- We get a linear system.

The Fundamental Matrix

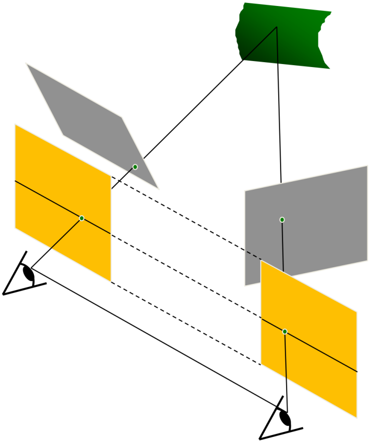
- Let's say that you found a few matching points in both images: $(x_{l,1}, y_{l,1}) \leftrightarrow (x_{r,1}, y_{r,1})$, ..., $(x_{l,n}, y_{l,n}) \leftrightarrow (x_{r,n}, y_{r,n})$
- Then you can get the parameters $\mathrm{f} := [F_{11}, F_{12}, \ldots, F_{33}]$ by solving:

$$\begin{bmatrix} x_{r,1}\, x_{l,1} & x_{r,1}\, y_{l,1} & x_{r,1} & y_{r,1}\, x_{l,1} & y_{r,1}\, y_{l,1} & y_{r,1} & x_{l,1} & y_{l,1} & 1 \\ & & & & \vdots & & & & \\ x_{r,n}\, x_{l,n} & x_{r,n}\, y_{l,n} & x_{r,n} & y_{r,n}\, x_{l,n} & y_{r,n}\, y_{l,n} & y_{r,n} & x_{l,n} & y_{l,n} & 1 \end{bmatrix} \mathrm{f} = 0$$

- How many correspondences do we need?
- We can estimate F with 8 correspondences. Of course, the more the better (why?).
- See Zisserman & Hartley's book for details.

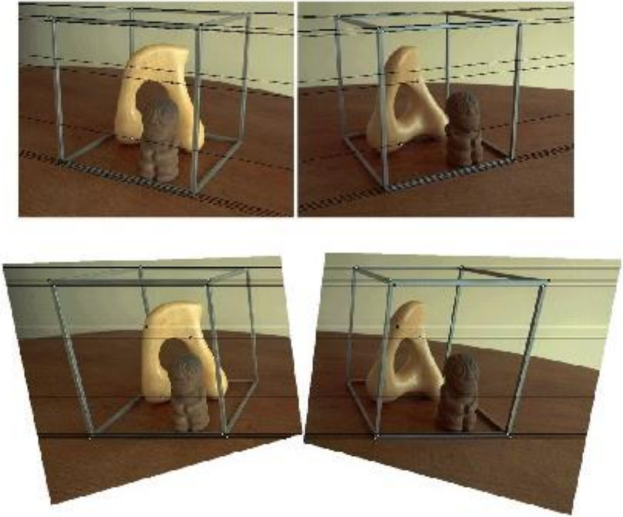
Rectification

- Once we have F we can compute homographies that transform each image plane such that they are parallel (see Zisserman & Hartley's book)
- Once they are parallel, we know how to proceed (matching, etc)



[Source: J. Hays]

Rectification Example



[Source: J. Hays]

Birdseye View on What We Learned So Far

Problem	Detection	Description	Matching
Find Planar Distinctive Objects	Scale Invariant Interest Points	Local feature: SIFT	All features to all features + Affine / Homography
Panorama Stitching	Scale Invariant Interest Points	Local feature: SIFT	All features to all features + Homography
Stereo	Compute in every point	Intensity or Gradient patch	For each point search on epipolar line