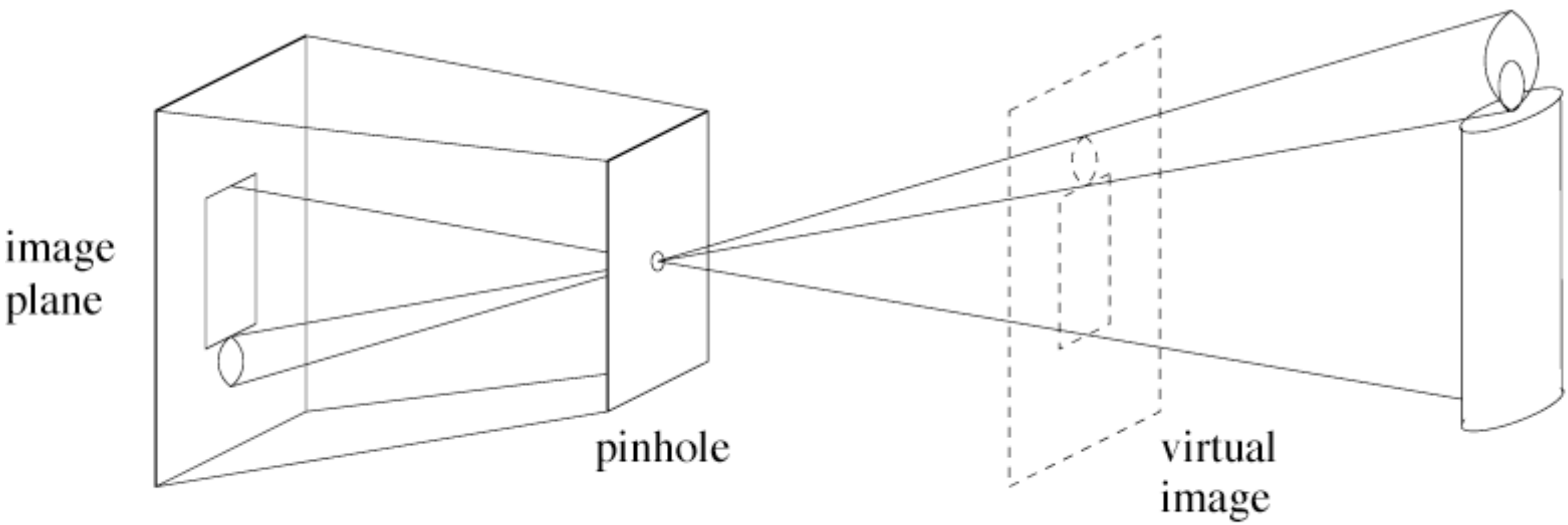
Two cameras: Basics

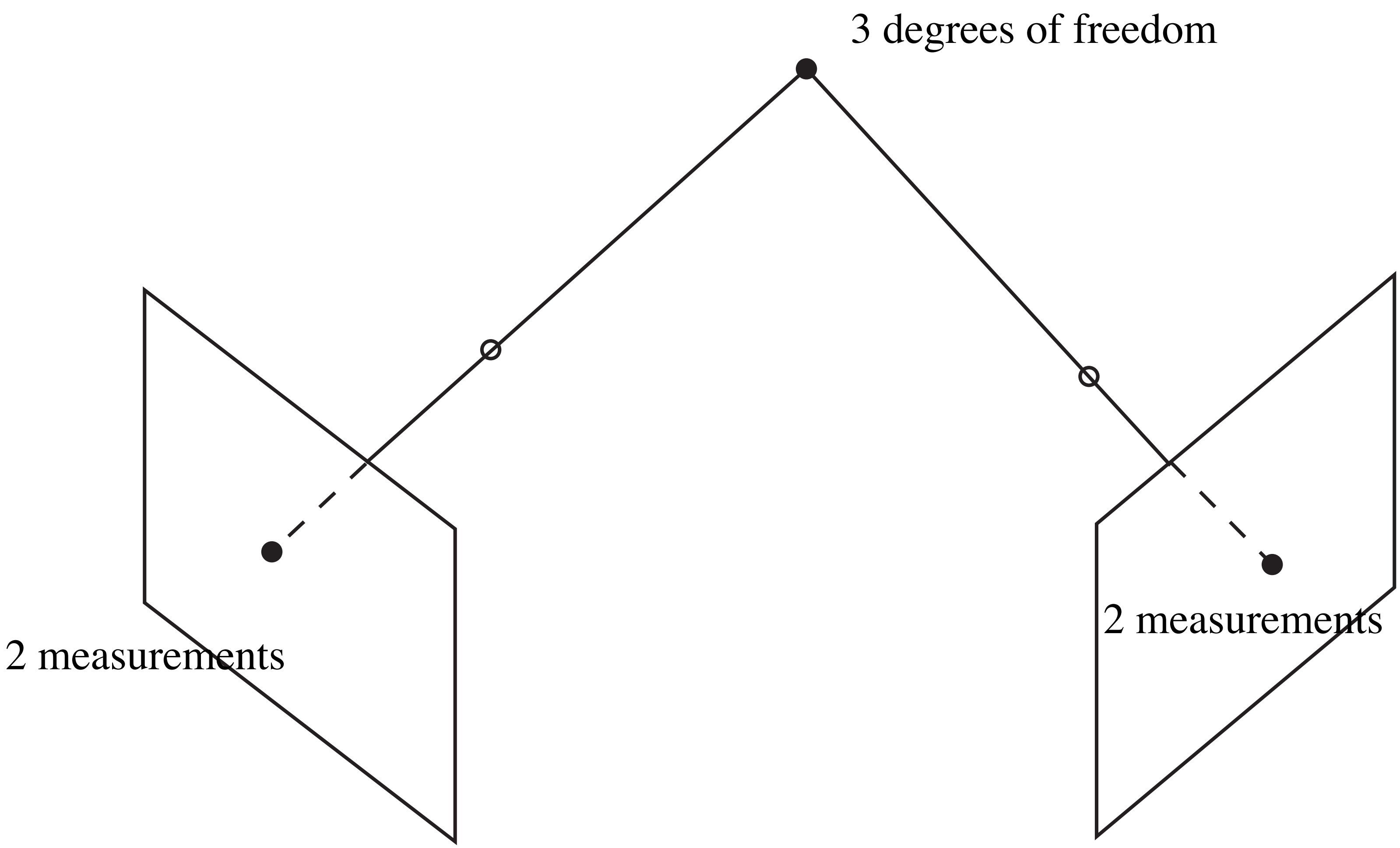
D.A. Forsyth, UIUC

How cameras work

Pinhole camera - an effective abstraction



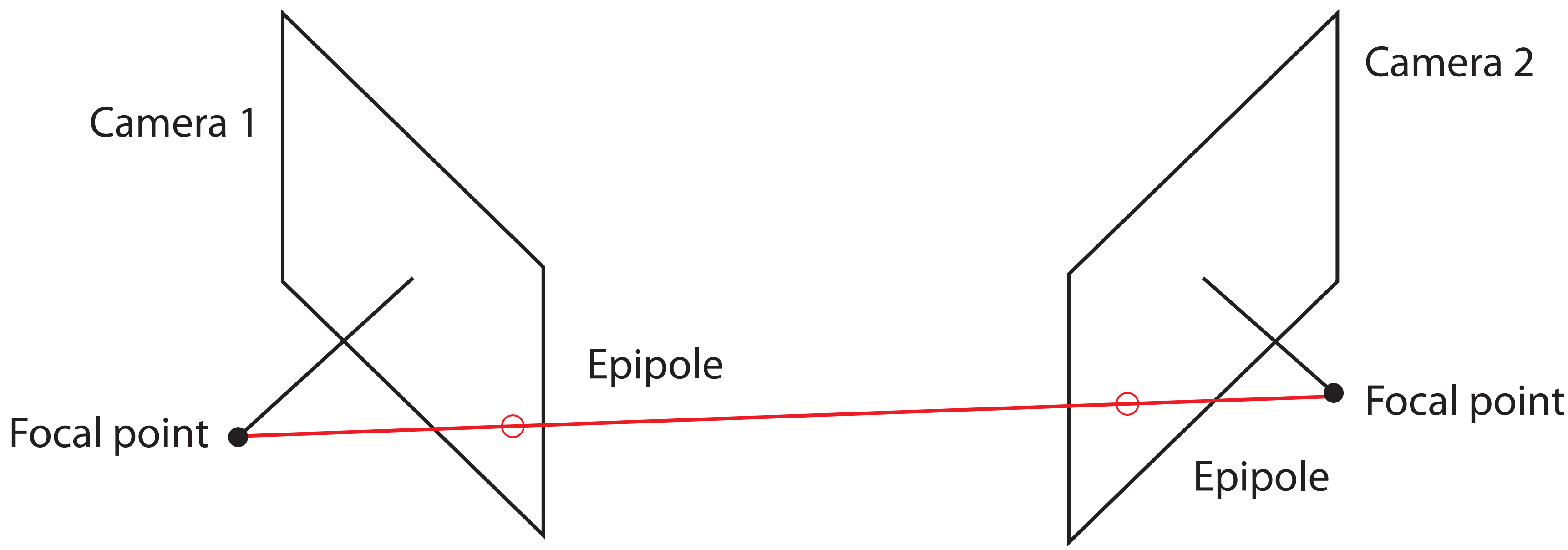
What happens in two views



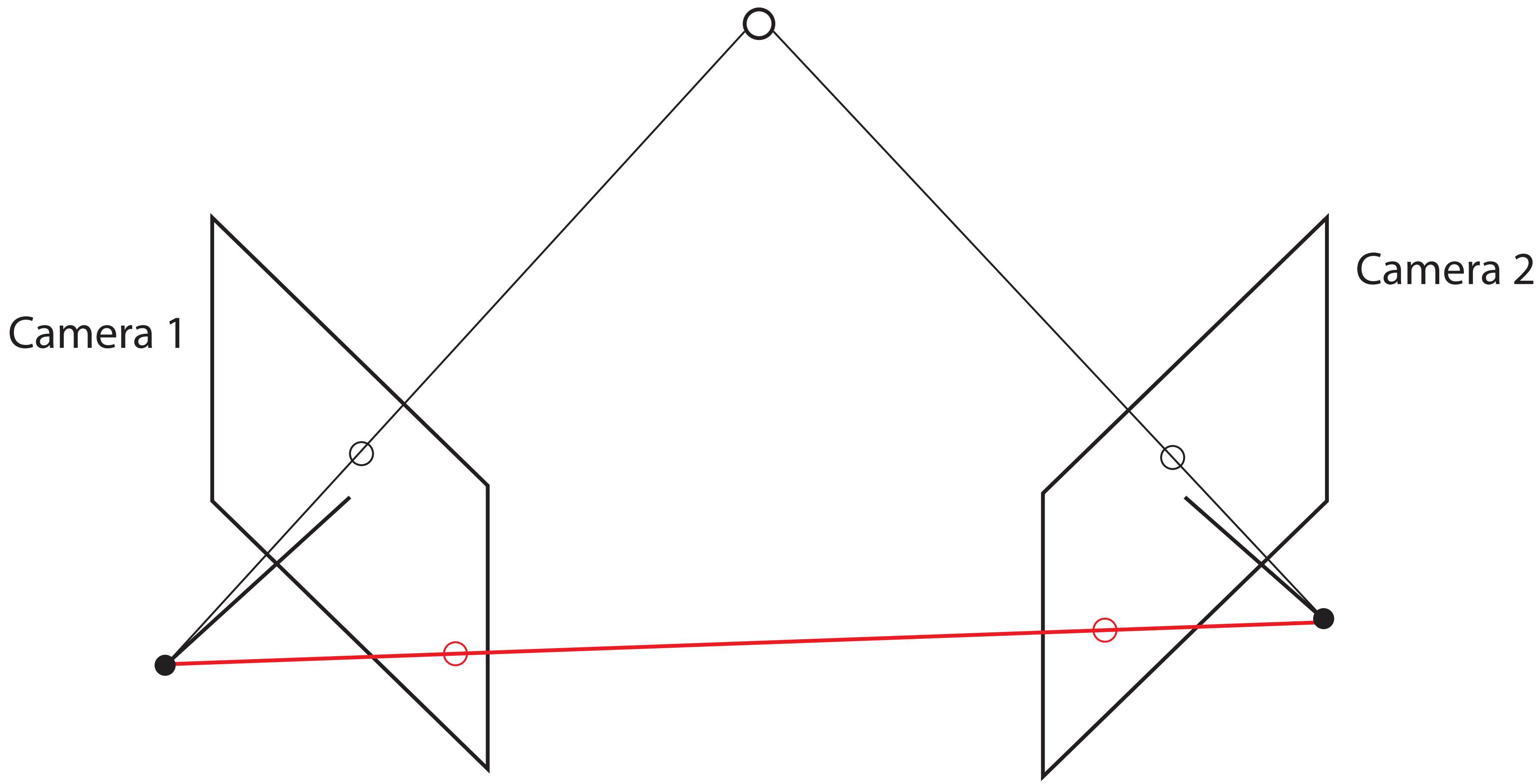
All of Camera Geometry

- From the picture
  - two views of a point give four measurements of three DOF
  - this means
    - correspondence is constrained
    - if we have enough points and enough pix we can recover
      - points
      - cameras

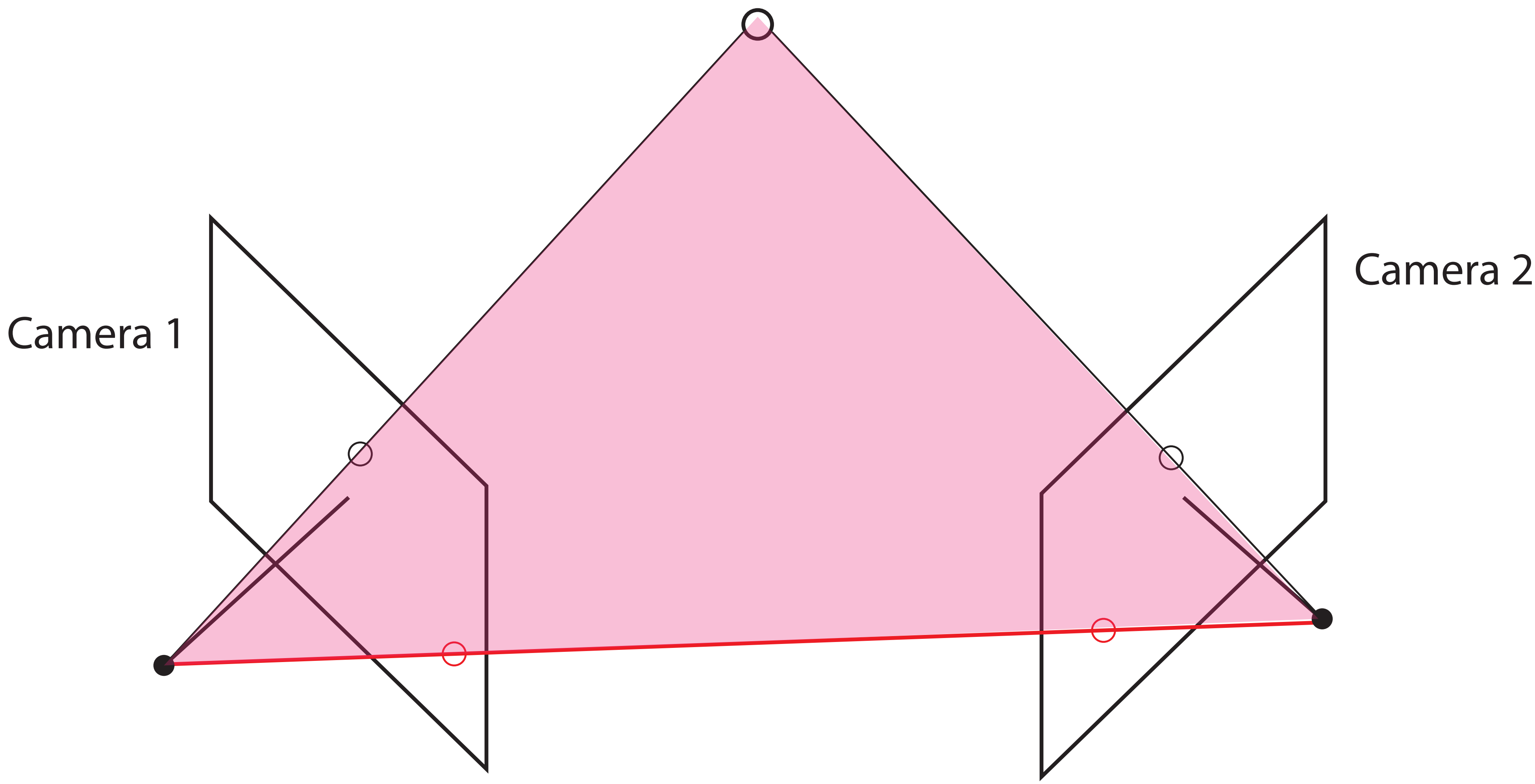
The Epipoles



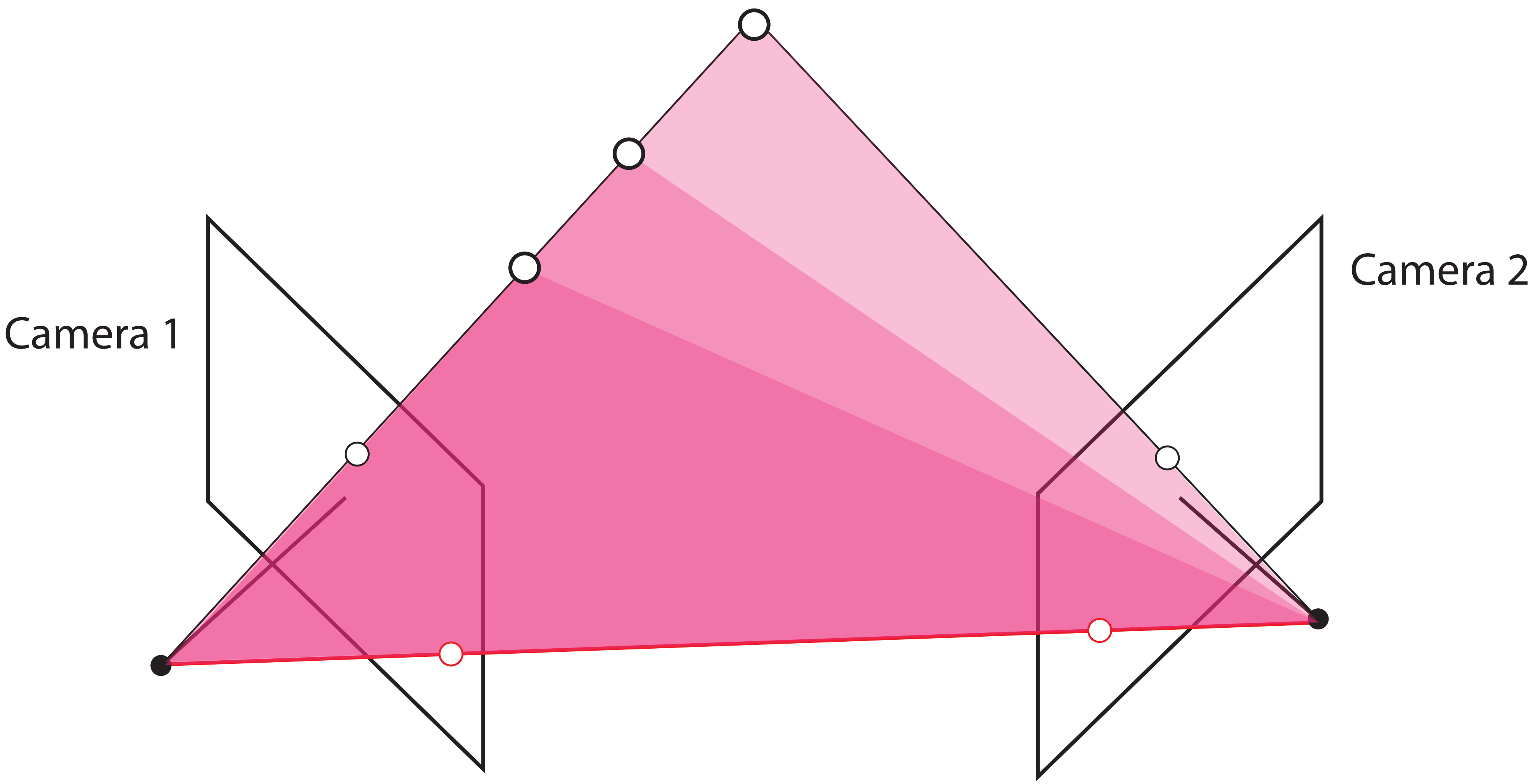
Constraints on correspondence



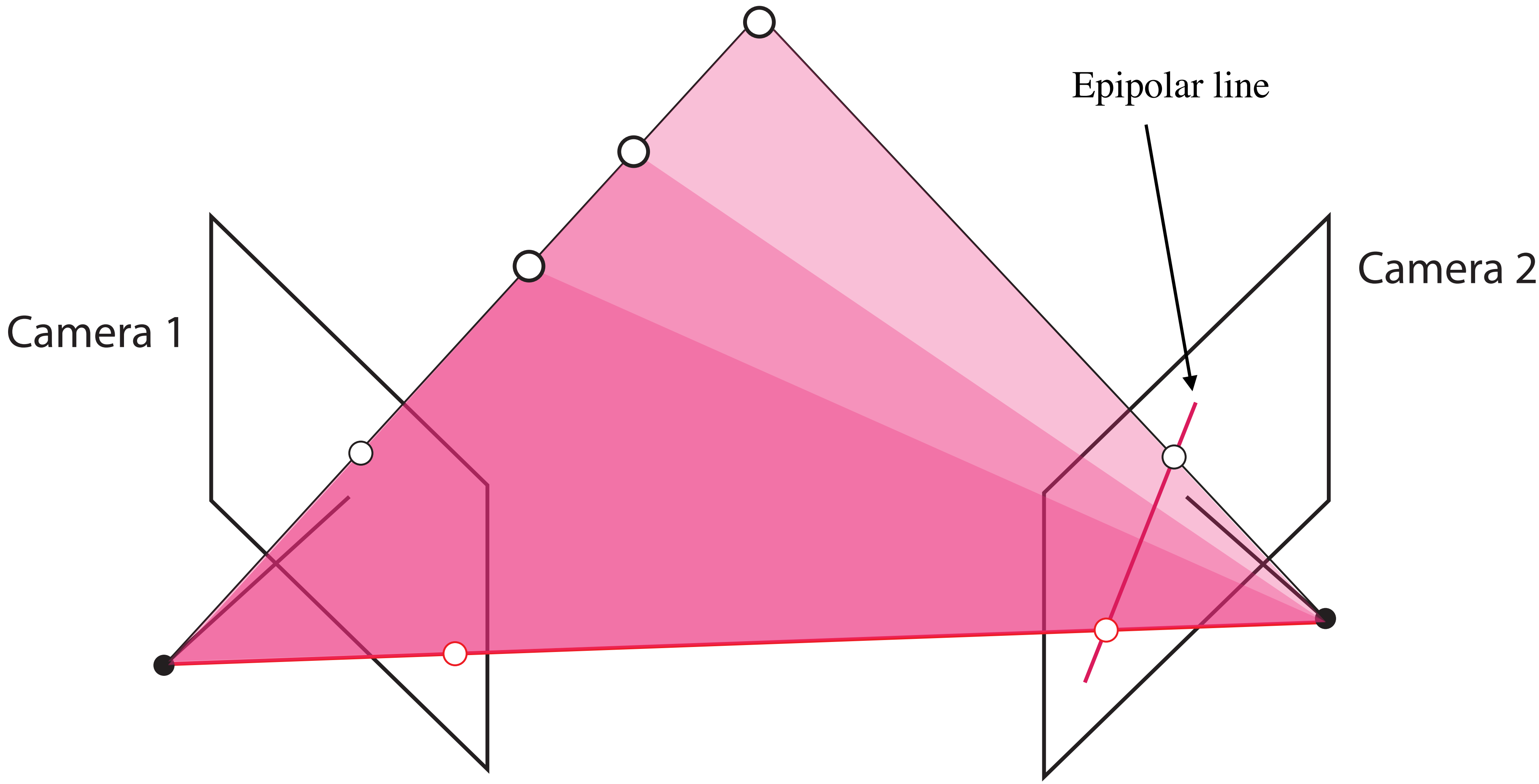
Constraints on CSP -II



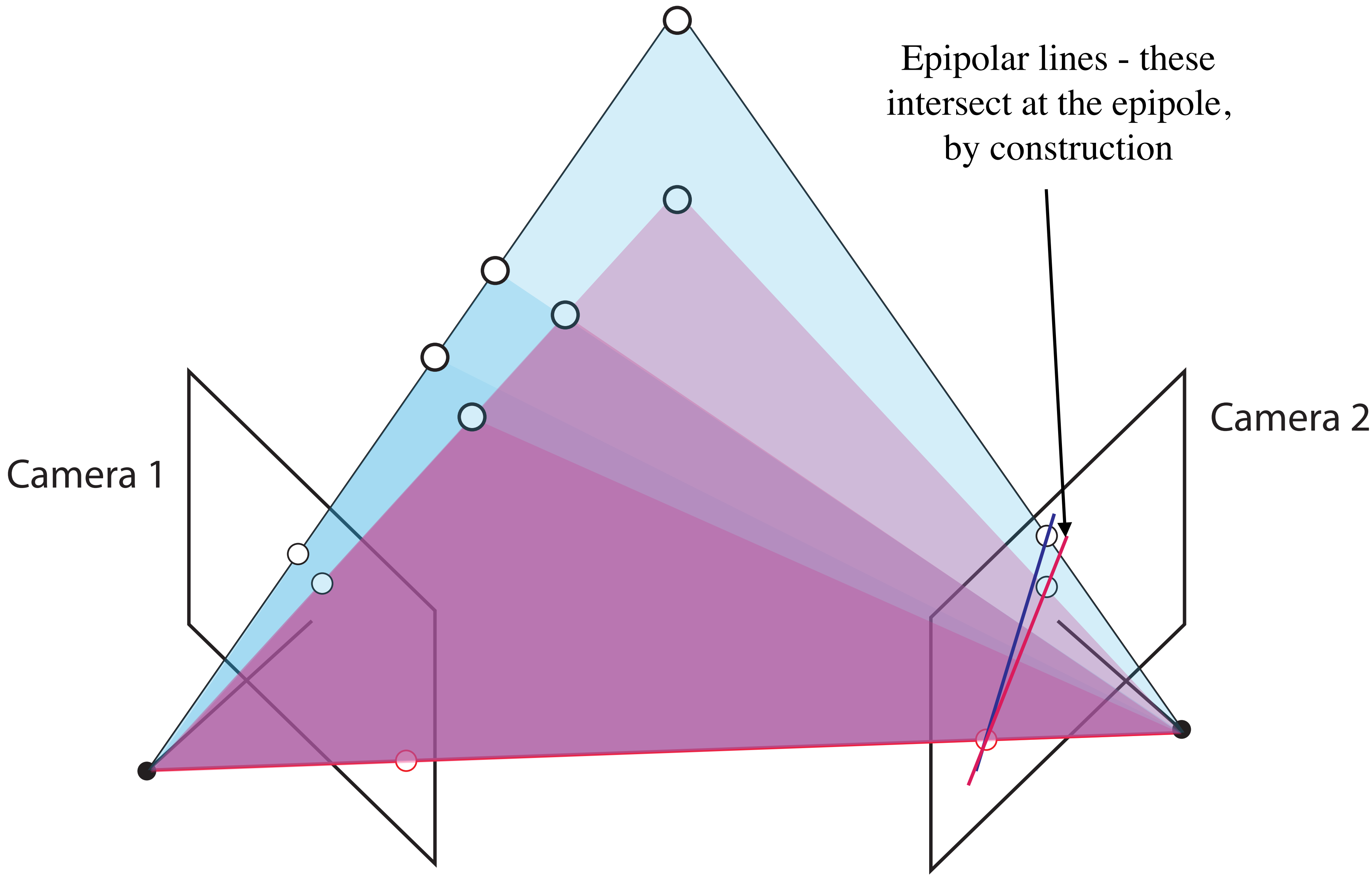
Constraints on CSP - III



Constraints on CSP - III

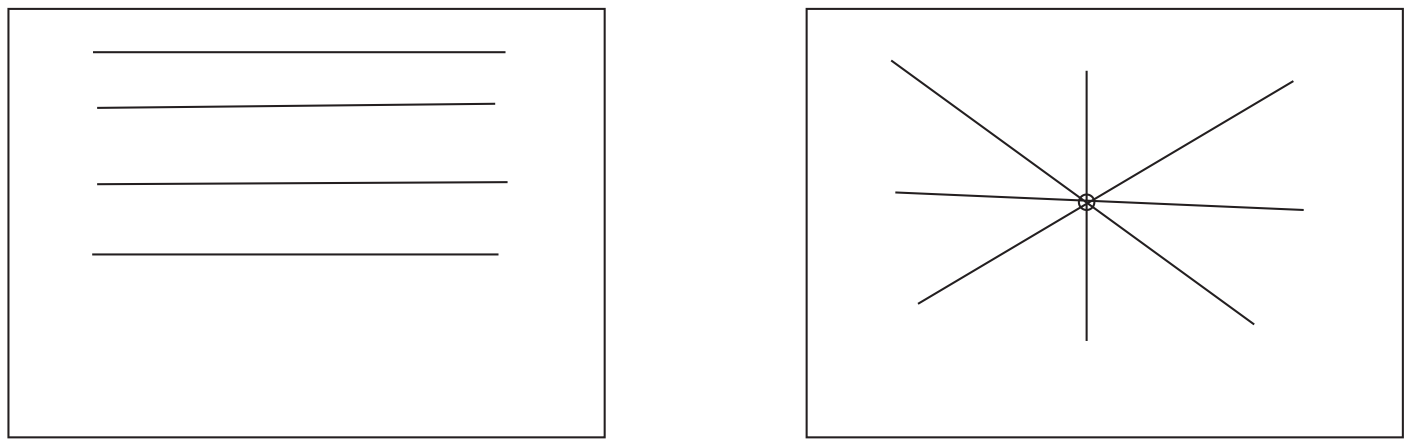


Constraints on CSP - IV



Epipoles (resp. epipolar lines)

- Informative



Epipole and epipolar lines in camera 1 - where is camera 2?

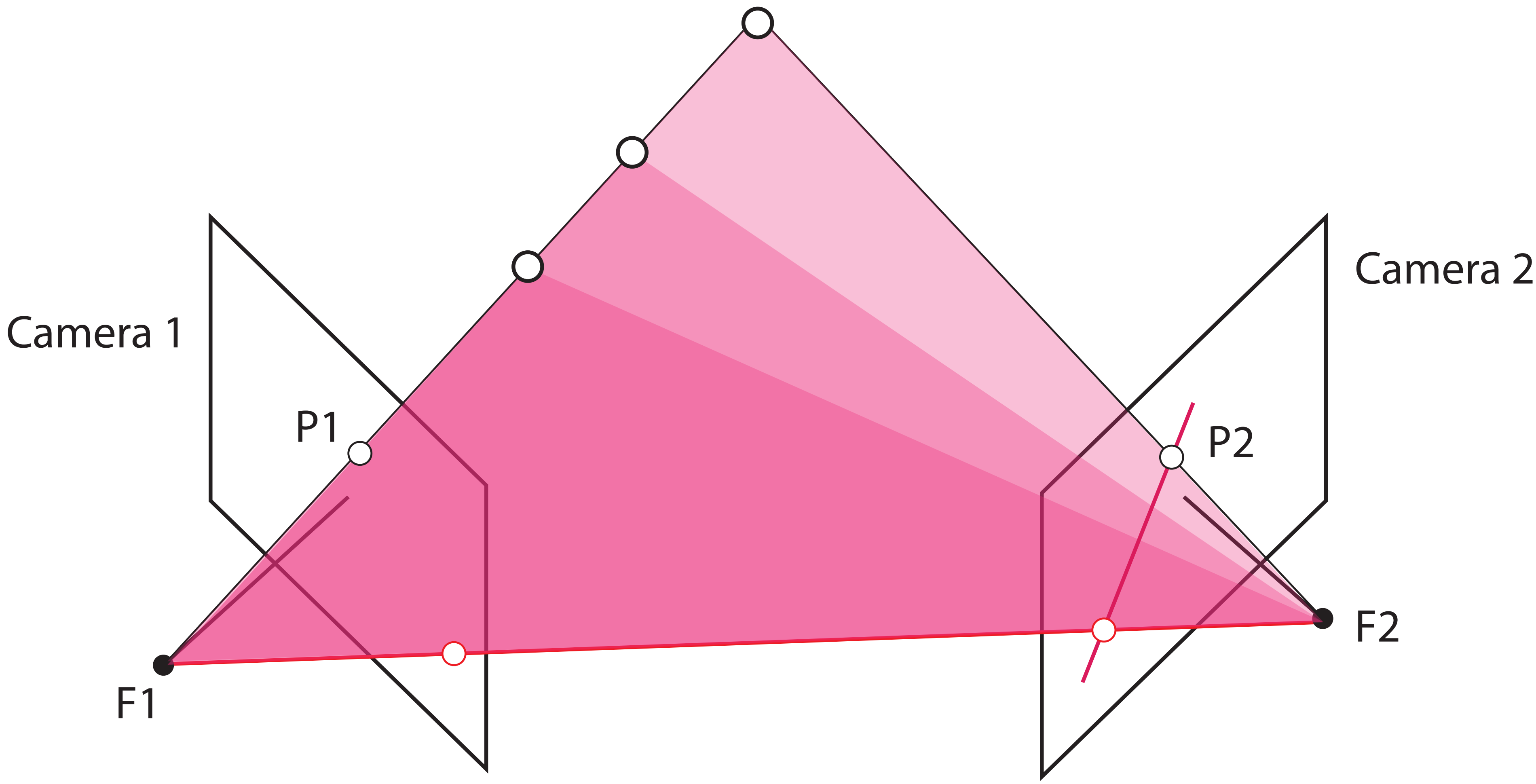
This means:

- A point in camera 1 identifies a line in camera 2
  - of all possible corresponding points in camera 2
- Equivalently, there is a map
  - from points in camera 1 (resp 2)
  - to lines in camera 2 (resp 1)
- Q: what is the form of the map?

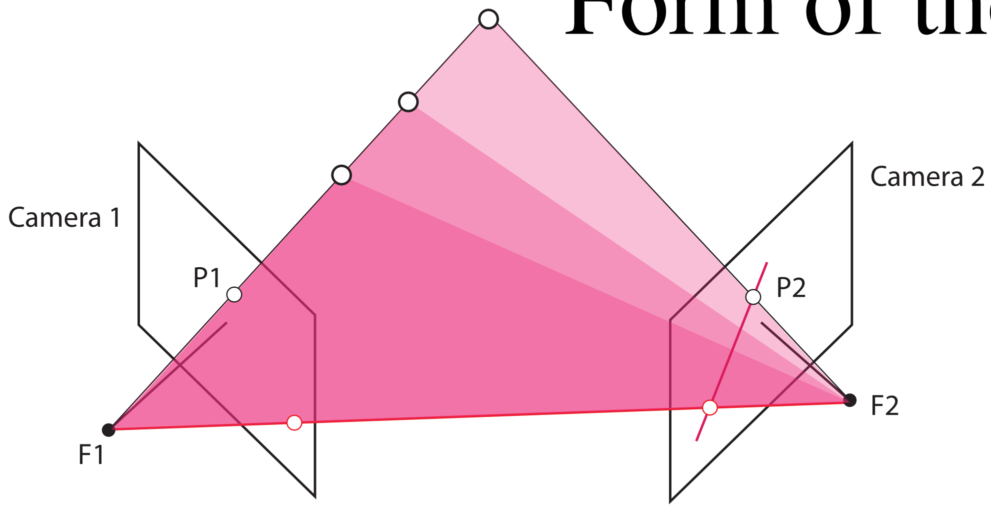
Planes in HCs

- Assume four points P1, P2, P3, P4 are coplanar
- Then
  - determinant([P1, P2, P3, P4])=0
- Trick:
  - equation of plane through three points?
  - determinant([P1, P2, P3, X])=0

Form of the map - notation



Form of the map - II



- 3D coordinates of P1 are linear in image coordinates (p1)
- 3D coordinates of P2 are linear in image coordinates (p2)
- so

$$\det\left([P_1, P_2, F_1, F_2]\right) = 0$$

In HC's

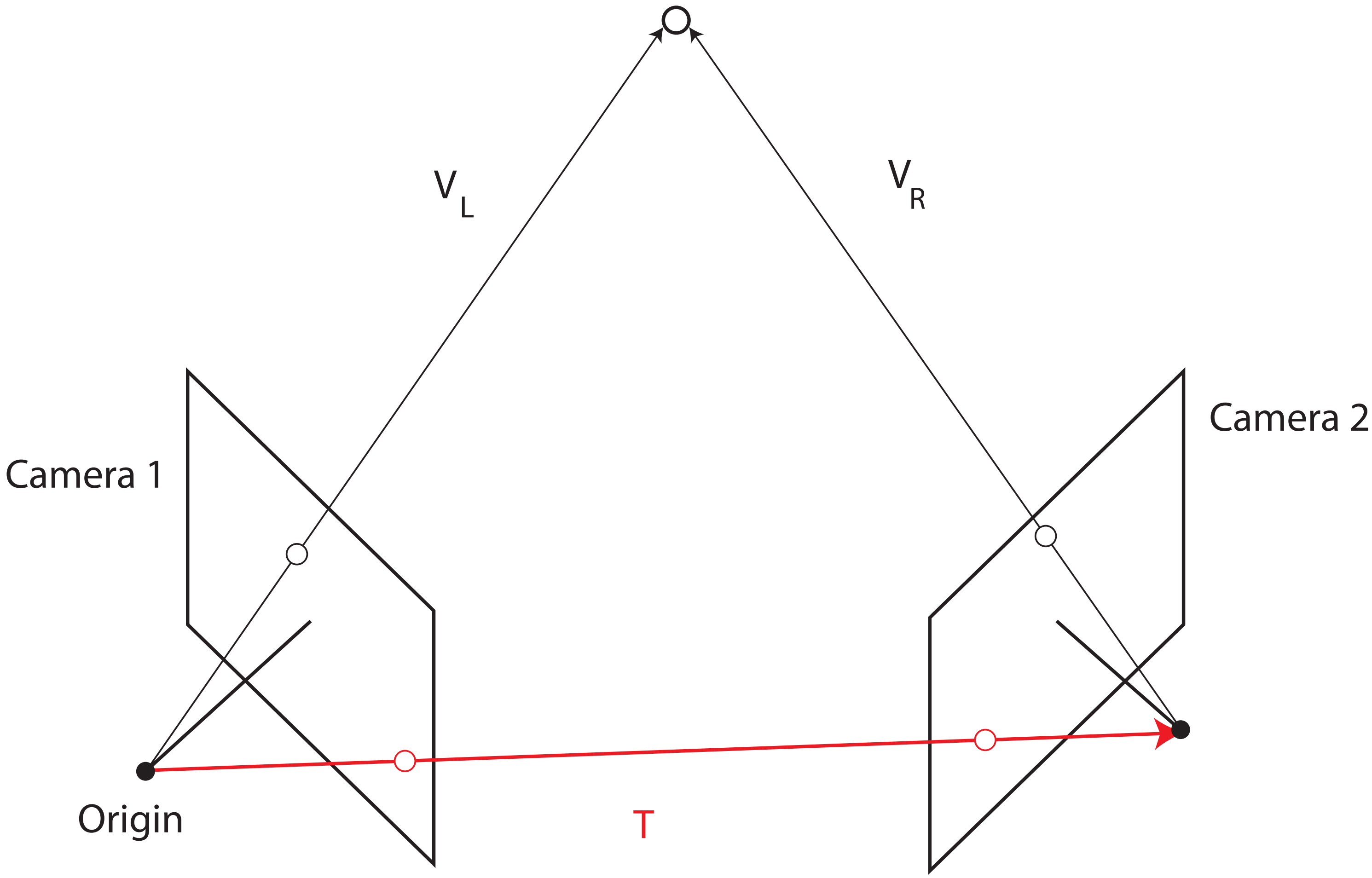
  - linear in p1; linear in p2
- so there is some matrix F (function of cameras) so that

$$\mathbf{p_1}^T \mathcal{F} \mathbf{p_2} = 0$$

The Fundamental Matrix

$$\mathbf{p_1}^T \mathcal{F} \mathbf{p_2} = 0$$

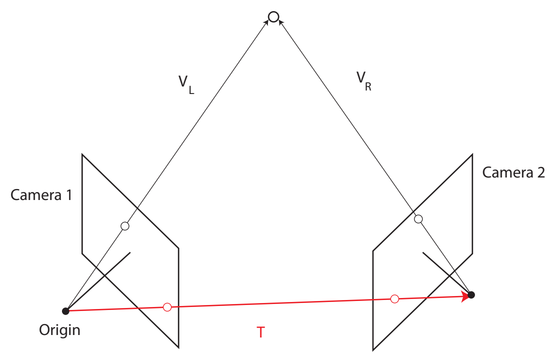
- Easy closed form expression exists
  - in terms of rot, trans between cameras, intrinsics
  - following slides
- Can be fit a pair of images using feature correspondences
  - 8 point algorithm
  - robustness is an important issue
  - we'll do this



Camera translation

$$\mathbf{V}_R = \mathcal{R}(\mathbf{V}_L - \mathbf{T})$$

Camera rotation



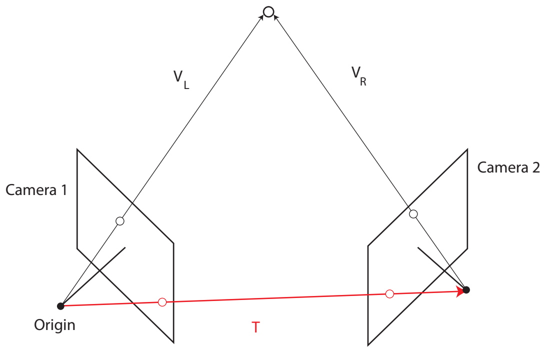
Camera translation

$$\mathbf{V}_R = \mathcal{R}(\mathbf{V}_L - \mathbf{T})$$

Camera rotation

$$\mathcal{S} = \begin{pmatrix} 0 & -T_z & T_y \\ T_z & 0 & -T_x \\ -T_y & T_x & 0 \end{pmatrix}$$

$$\mathbf{T}^T \mathcal{S} = \mathbf{0}$$



$$\mathbf{V}_R = \mathcal{R}(\mathbf{V}_L - \mathbf{T})$$

$$\mathbf{V}_R^T \mathcal{R} \mathcal{S} \mathbf{V}_L = (\mathbf{V}_L - \mathbf{T})^T \mathcal{R}^T \mathcal{R} \mathcal{S} \mathbf{V}_L = \mathbf{V}_L^T \mathcal{S} \mathbf{V}_L = 0$$

RECALL: The camera matrix - II

- Turn previous expression into HC's
  - HC's for 3D point are (X,Y,Z,T)
  - HC's for point in image are (U,V,W)

Transforms points from object coordinates into world coordinates most likely a rotation and translation

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \mathcal{C} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \mathcal{W} \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix}$$

Transforms camera coordinates (f is hidden in here)

So…

$$\mathcal{F} = k\mathcal{C}_L^{-T}\mathcal{R}\mathcal{S}\mathcal{C}_R^{-1}$$

If we know these

we can recover info about R, T from F

Fundamental matrix and epipolar lines

- In homogenous coordinates, line in plane is:

$$aX + bY + cZ = 0$$

- can write:

$$\mathbf{a}^T \mathbf{x} = 0$$

- But look at

$$\mathbf{p_1}^T \mathcal{F} \mathbf{p_2} = 0$$

  - which can be written

Coefficients of a line in image 2
created by F and p_1

$$\mathbf{p}_1^T \mathcal{F} \mathbf{p}_2 = (\mathcal{F}^T \mathbf{p}_1)^T \mathbf{p}_2 = 0$$

The Fundamental Matrix

$$\mathbf{p_1}^T \mathcal{F} \mathbf{p_2} = 0$$

- A map from
  - point in 1 (resp. 2) to line in 2 (resp. 1)
- This is the algebraic version of the picture
  - but the picture tells us more
- Any point in 1 maps to a line through the epipole
  - MOST lines in 2 are NOT in the image of the map
  - only a 1 parameter family of lines IS (the ones through the epipole)
- F has rank 2!
- Left (resp. right) kernel of F is left (resp. right) epipole

The 8 point algorithm

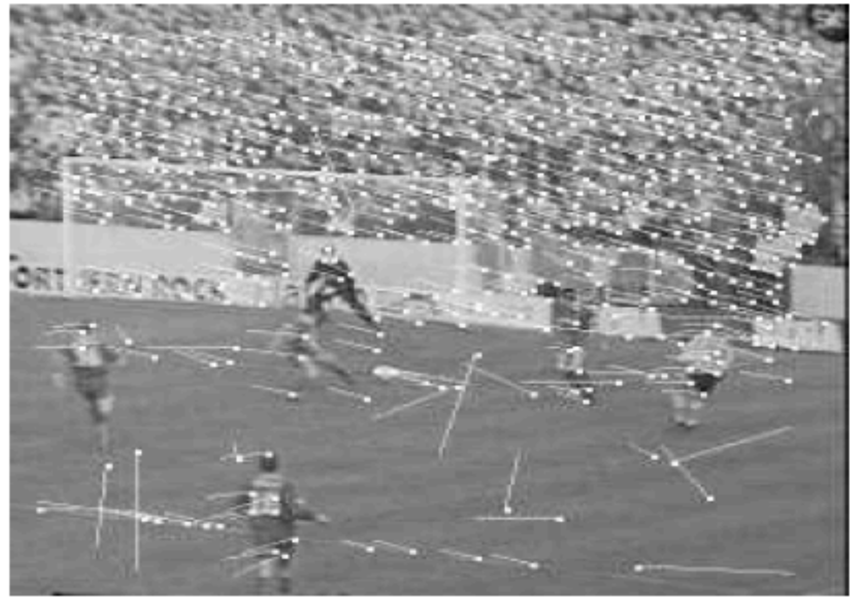
$$\mathbf{p_1}^T \mathcal{F} \mathbf{p_2} = 0$$

- Find 8 p1, p2 pairs
  - this gives 8 homogeneous linear equations in F coefficients
  - solve these
- Improvements
  - you can do it with seven points and solving a cubic (rank deficient)
  - the image coordinate system really matters for the quality of estimate
  - this requires robust estimation to work well
    - RANSAC

RANSAC (outline)

- Repeat many times
  - Find 8 pairs (p1, p2)
  - Fit F using 8 point
  - record the number of inlying pairs
    - pairs p1, p2 where:
      - p2 is "close" to (F p1)^T
      - p1 is "close" to (F p2)^T
      - there's an appearance match
- Take F with most inlying pairs
  - fit to all inliers
    - using perpendicular distance from point to line
- Q: repeat how many times?
  - A: often enough that you have high prob of seeing 8 inlying pairs

(a)



(b)

(c)

(d)

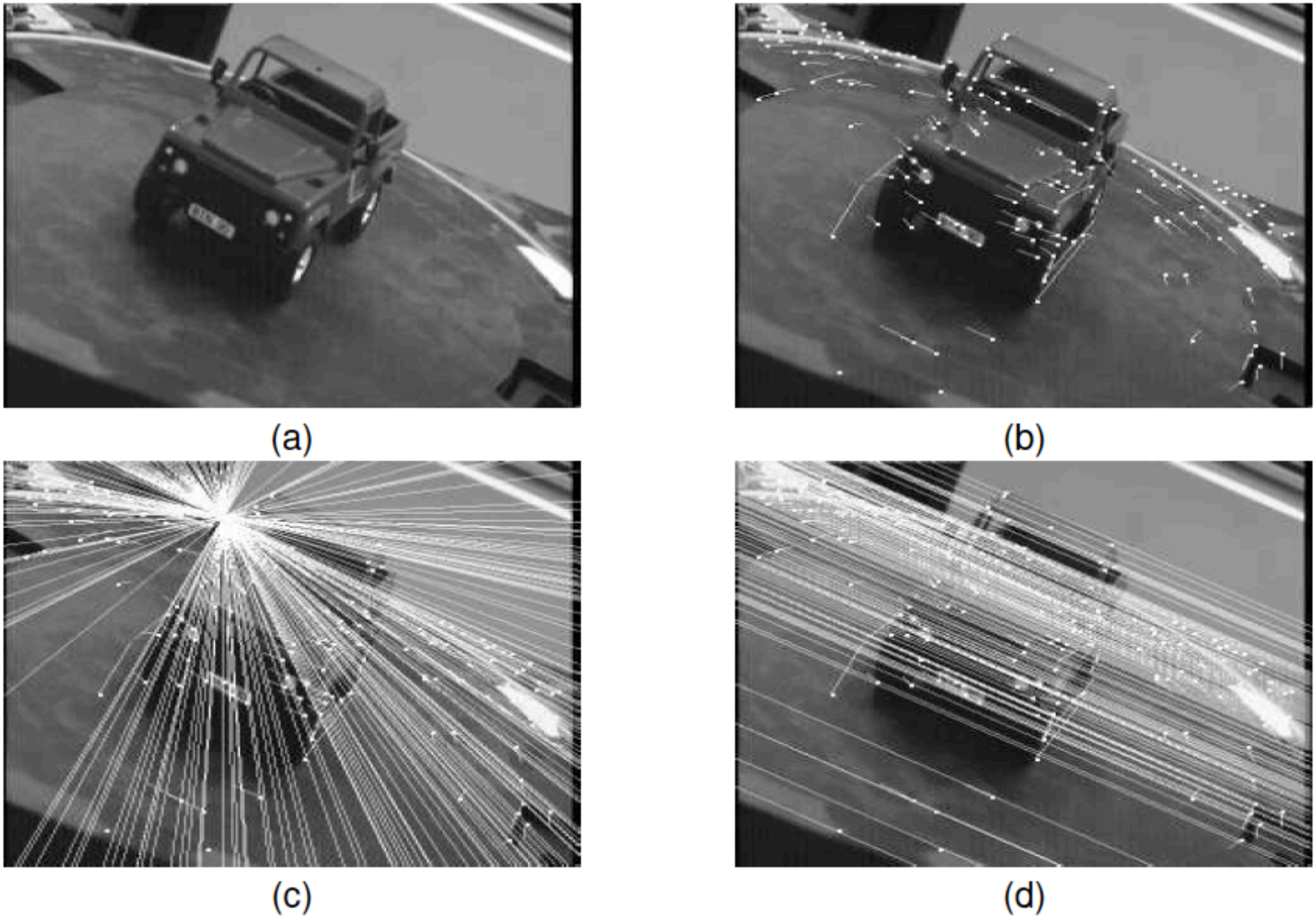


Fig. 18. In (a) (b) two consecutive images of a buggy rotating on a turntable. (b) has 167 matches superimposed on the second image. (c) (d) show two epipolar geometries generated by two distinct fundamental matrices, 139 correspondences are consistent with the fundamental matrix in (a), 131 are consistent with the fundamental matrix in (b) yet the two epipolar geometries obviously differ.