Epipolar Geometry

CSE 6367: Computer Vision

Instructor: William J. Beksi

Introduction

- The **epipolar geometry** is the intrinsic projective geometry between two views
- It is independent of scene structure, and only depends on the camera's internal parameters and relative pose

Introduction

- The fundamental matrix F encapsulates this intrinsic geometry
- It is a 3×3 matrix of rank 2
- If a point in 3-space $\mathbf{X}$ is imaged as $\mathbf{x}$ in the first view, and $\mathbf{x}'$ in the second, then the image points satisfy the relation $\mathbf{x}'^T F \mathbf{x} = 0$

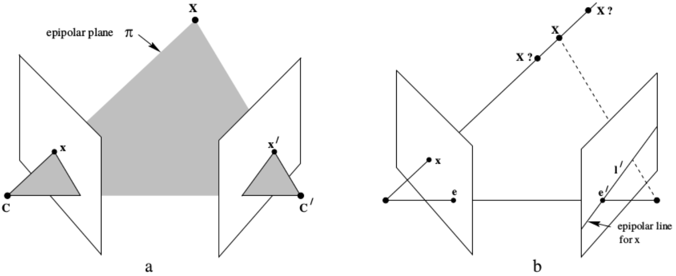
Epipolar Lines and Planes

- The epipolar geometry between two views is essentially the geometry at the intersection of the image planes with the pencil of the planes having the baseline (the line joining the camera centers) as the axis
- This geometry is usually motivated by considering the search for corresponding points in stereo matching

Epipolar Lines and Planes

- Suppose a point $\mathbf{X}$ in 3-space is imaged in two views, at $\mathbf{x}$ in the first and $\mathbf{x}'$ in the second
- What is the relation between the corresponding image points $\mathbf{x}$ and $\mathbf{x}'$?

Epipolar Lines and Planes



- Point correspondence geometry: (a) Two cameras are indicated by their centers **C** and **C**′ and image planes; The camera centers, **X**, and its images **x** and **x**′ lie in a common plane π; (b) **x** back-projects to a ray in 3-space defined by **C** and **x**, this ray is imaged as a line **l**′ in the second view

Epipolar Lines and Planes

- The image points $\mathbf{x}$ and $\mathbf{x}'$, space point $\mathbf{X}$, and camera centers are coplanar and lie on the plane π
- The rays back-projected from $\mathbf{x}$ and $\mathbf{x}'$ intersect at $\mathbf{X}$, and the rays are coplanar in π which is of significance in searching for a correspondence

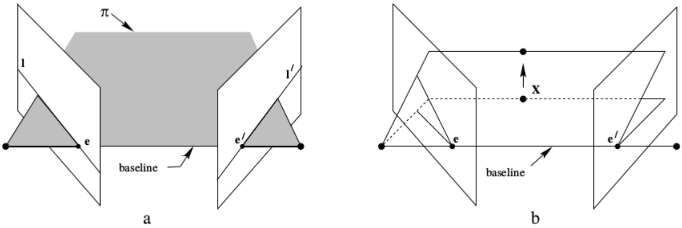
Epipolar Lines and Planes

- Suppose that we only know $\mathbf{x}$, we may ask how the corresponding point $\mathbf{x}'$ is constrained
- The plane π is determined by the baseline and the ray defined by $\mathbf{x}$
- We know that the ray corresponding to the (unknown) point $\mathbf{x}'$ lies in π, therefore the point $\mathbf{x}'$ lies on the line of intersection $\mathbf{l}'$ of π with the second image plane

Epipolar Lines and Planes

- This line $\mathbf{l}'$ is the image in the second view of the ray back-projected from $\mathbf{x}$, and it is called the epipolar line corresponding to $\mathbf{x}$
- In terms of a stereo correspondence algorithm, the benefit is that the search for the point corresponding to $\mathbf{x}$ need not cover the entire image plane but can be restricted to the line $\mathbf{l}'$

Epipolar Lines and Planes

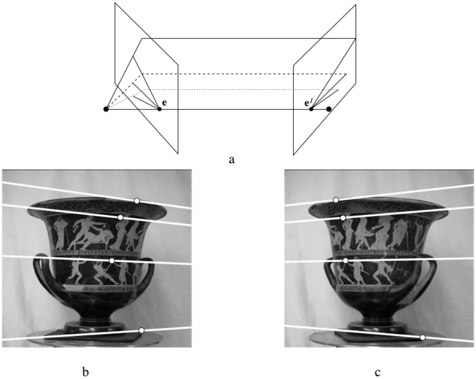


- Epipolar geometry: (a) The camera baseline intersects each image plane at the epipoles $\mathbf{e}$ and $\mathbf{e}'$; any plane π containing the baseline is an epipolar plane, and intersects the image planes in $\mathbf{l}$ and $\mathbf{l}'$; (b) As the position of $\mathbf{X}$ varies, the epipolar planes "rotate" about the baseline

Epipolar Lines and Planes

- The geometric entities involved in epipolar geometry are the following:
    - The **epipole** is the *point* of intersection of the line joining the camera centers (the baseline) with the image plane, it is also the vanishing point of the baseline (translation) direction
    - An **epipolar plane** is a plane containing the baseline, there is a one-parameter family (a pencil) of epipolar planes
    - An **epipolar line** is the intersection of an epipolar plane with the image plane (all epipolar lines intersect at the epipole) and an epipolar plane intersects the left and right image planes in epipolar lines thus defining the correspondence between the lines

Epipolar Lines and Planes



- (a) Epipolar geometry for converging cameras; (b) and (c) A pair of images with superimposed corresponding points and their epipolar lines (in white); The motion between the views is a translation and rotation

The Fundamental Matrix F

- The **fundamental matrix** is the algebraic representation of epipolar geometry
- Given a pair of images, for each point $\mathbf{x}$ in one image, there exists a corresponding epipolar line $\mathbf{l}'$ in the other image
- Any point $\mathbf{x}'$ in the second image matching the point $\mathbf{x}$ must lie on the epipolar line $\mathbf{l}'$

The Fundamental Matrix F

- The epipolar line is the projection in the second image of the ray from the point $\mathbf{x}$ through the camera center $\mathbf{C}$
- Thus, there is a map

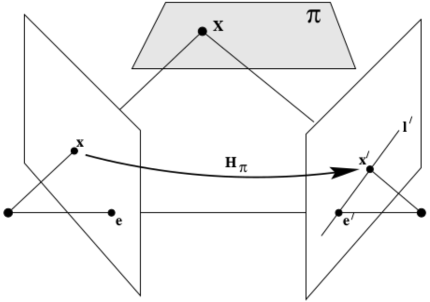
$$\mathbf{x} \mapsto \mathbf{l}'$$

from a point in one image to its corresponding epipolar line in the other image

Geometric Derivation of F

- The mapping from a point in one image to a corresponding epipolar line in the other image may be decomposed into two steps
- In the first step, the point $\mathbf{x}$ is mapped to some point $\mathbf{x}'$ in the other image lying on the epipolar line $\mathbf{l}'$ (this point $\mathbf{x}'$ is a potential match for the point $\mathbf{x}$)
- In the second step, the epipolar line $\mathbf{l}'$ is obtained as the line joining $\mathbf{x}'$ to the epipole $\mathbf{e}'$

Geometric Derivation of F



- A point $\mathbf{x}$ in one image is transferred via the plane $\boldsymbol{\pi}$ to a matching point $\mathbf{x}'$ in the second image, the epipolar line through $\mathbf{x}$ is obtained by joining $\mathbf{x}$ to the epipole $\mathbf{e}$

Step 1: Point Transfer via a Plane

- Consider a plane $\boldsymbol{\pi}$ in space not passing through either of the two camera centers
- The ray through the first camera center corresponding to $\mathbf{x}$ meets $\boldsymbol{\pi}$ at $\mathbf{X}$, then $\mathbf{X}$ is projected to $\mathbf{x}'$ in the second image
- This procedure is known as transfer via the plane $\boldsymbol{\pi}$

Step 1: Point Transfer via a Plane

- Since $\mathbf{X}$ lies on the ray corresponding to $\mathbf{x}$, the projected point $\mathbf{x}'$ must lie on the epipolar line $\mathbf{l}'$ corresponding to the image of this ray
- Points $\mathbf{x}$ and $\mathbf{x}'$ are both images of the 3D point $\mathbf{X}$ lying on a plane

Step 1: Point Transfer via a Plane

- The set of all such points $\mathbf{x}_i$ in the first image and the corresponding points $\mathbf{x'}_i$ in the second image are projectively equivalent (since they are each projectively equivalent to the planar point set $\mathbf{X}_i$)
- Therefore, there is a 2D homography H_π mapping each $\mathbf{x}_i$ to $\mathbf{x'}_i$

Step 2: Constructing the Epipolar Line

- Given the point $\mathbf{x}'$ the epipolar line $\mathbf{l}'$ passing through $\mathbf{x}'$ and the epipole $\mathbf{e}'$ can be written as $\mathbf{l}' = \mathbf{e}' \times \mathbf{x}' = [\mathbf{e}']_\times \mathbf{x}'$
- Since $\mathbf{x}'$ may be written as $\mathbf{x}' = H_\pi \mathbf{x}$, we have

$$\mathbf{l}' = [\mathbf{e}']_\times H_\pi \mathbf{x} = F\mathbf{x}$$

where we define $F = [\mathbf{e}']_\times H_\pi$ as the fundamental matrix

Step 2: Constructing the Epipolar Line

- Geometrically, F represents a mapping from the 2D projective plane $\mathbb{P}^2$ of the first image to the pencil of epipolar lines through the epipole $\mathbf{e}'$
- Thus, it represents a mapping from a 2D onto a 1D projective space, and therefore must have rank 2

Algebraic Derivation of F

- The form of the fundamental matrix in terms of the two camera projection matrices, P and P', can be derived algebraically
- The ray back-projected from $\mathbf{x}$ by P is obtained by solving $P\mathbf{X} = \mathbf{x}$
- The one parameter family of solutions is given as

$$\mathbf{X}(\lambda) = P^{\dagger}\mathbf{x} + \lambda\mathbf{C}$$

where $P^{\dagger}$ is the psuedoinverse of P, the null vector of $\mathbf{C}$, namely the camera center, is defined by $P\mathbf{C} = \mathbf{0}$, and the ray is parameterized by the scalar λ

Algebraic Derivation of F

- Two points on the ray are $P^{\dagger}\mathbf{x}$ (at $\lambda = 0$) and the first camera center $\mathbf{C}$ (at $\lambda = \infty$)
- These two points are imaged by the second camera P' at $P'P^{\dagger}\mathbf{x}$ and $P'\mathbf{C}$ respectively in the second view
- The epipolar line is the line joining these two projected points, namely $\mathbf{l}' = (P'\mathbf{C}) \times (P'P^{\dagger}\mathbf{x})$

Algebraic Derivation of F

- The point $P'\mathbf{C}$ is the epipole in the second image, namely the projection of the first camera center, and may be denoted by $\mathbf{e}'$
- Therefore, $\mathbf{l}' = [\mathbf{e}']_\times (P'P^\dagger)\mathbf{x} = F\mathbf{x}$, where F is the matrix

$$F = [\mathbf{e}']_\times (P'P^\dagger)$$

- This is essentially the same formula for the fundamental matrix as given by the geometric derivation with the homography, H_π, having the explicit form $H_\pi = P'P^\dagger$ in terms of the two camera matrices

Example: Finding the Fundamental Matrix

- Suppose the camera matrices are those of a calibrated stereo rig with the world origin at the first camera

$$P = K[I \mid \mathbf{0}], \qquad P' = K'[R \mid \mathbf{t}]$$

Example: Finding the Fundamental Matrix

- Then

$$P^{\dagger} = \begin{bmatrix} K^{-1} \\ \mathbf{0}^T \end{bmatrix}, \qquad \mathbf{C} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}$$

and

$$\begin{aligned} F &= [P'\mathbf{C}]_\times P'P^{\dagger} \\ &= [K'\mathbf{t}]_\times K'RK^{-1} \\ &= K'^{-T}[\mathbf{t}]_\times RK^{-1} \\ &= K'^{-T}R[R^T\mathbf{t}]_\times K^{-1} \\ &= K'^{-T}RK^T[KR^T\mathbf{t}]_\times \end{aligned}$$

Example: Finding the Fundamental Matrix

- Note that the epipoles (defined as the image of the other camera center) are

$$\mathbf{e} = P \begin{bmatrix} -R^T\mathbf{t} \\ 1 \end{bmatrix} = KR^T\mathbf{t}, \qquad \mathbf{e}' = P' \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} = K'\mathbf{t}$$

thus we may write

$$\begin{aligned} F &= [\mathbf{e}']_\times K'RK^{-1} \\ &= K'^{-T}[\mathbf{t}]_\times RK^{-1} \\ &= K'^{-T}R[R^T\mathbf{t}]_\times K^{-1} \\ &= K'^{-T}RK^T[\mathbf{e}]_\times \end{aligned}$$

Correspondence Condition

- The fundamental matrix satisfies the **correspondence condition** that for any pair of corresponding points $\mathbf{x} \leftrightarrow \mathbf{x}'$ in the two images

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

- This is true, because if $\mathbf{x}$ and $\mathbf{x}'$ correspond, then $\mathbf{x}'$ lies on the epipolar line $\mathbf{l}' = F\mathbf{x}$ corresponding to $\mathbf{x}$ (i.e. $0 = \mathbf{x}'^T \mathbf{l}' = \mathbf{x}'^T F \mathbf{x}$

- Conversely, if the image points satisfy the relation $\mathbf{x}'^T F \mathbf{x} = 0$ then the rays defined by these points are coplanar (a necessary condition for the points to correspond)

Correspondence Condition

- The importance of the correspondence condition is that it gives a way of characterizing the fundamental matrix without reference to the camera matrices, i.e. only in terms of corresponding image points
- This enables F to be computed from image correspondences alone

Fundamental Matrix Properties

- **Transpose**: If F is the fundamental matrix of the pair of cameras (P, P'), then F^T is the fundamental matrix of the pair in opposite order: (P', P)
- **Epipolar Lines**: For any point $\mathbf{x}$ in the first image, the corresponding epipolar line is $\mathbf{l}' = F\mathbf{x}$; Similarly, $\mathbf{l} = F^T\mathbf{x}'$ represents the epipolar line corresponding to $\mathbf{x}'$ in the second image
- The **epipole**: For any point $\mathbf{x}$ (other than $\mathbf{e}$) the epipolar line $\mathbf{l}' = F\mathbf{x}$ contains the epipole $\mathbf{e}'$; Thus $\mathbf{e}'$ satisfies ${\mathbf{e}'}^T(F\mathbf{x}) = ({\mathbf{e}'}^T F)\mathbf{x} = 0$ for all $\mathbf{x}$ and it follows that $\mathbf{e}'$ and $\mathbf{e}$ are the left and right null vectors of F

Fundamental Matrix Properties

- F has seven degrees of freedom: a 3×3 homogeneous matrix has eight independent ratios (there are nine elements, and the common scaling is not significant); however, F also satisfies the constraint $det(F) = 0$ which removes one degreee of freedom
- F is a *correlation*, a projective map taking a point to a line, i.e. a point in the first image $\mathbf{x}$ defines a line in the second $\mathbf{l}' = F\mathbf{x}$ which is the epipolar line of $\mathbf{x}$

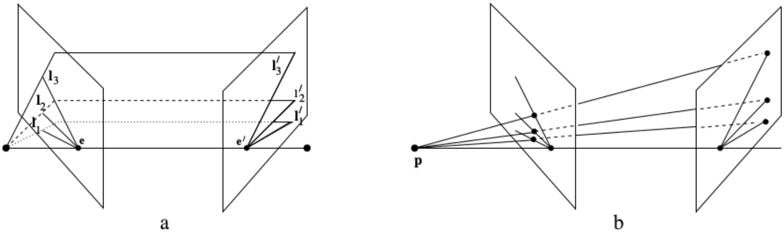
Summary of the Fundamental Matrix Properties

- $\mathtt{F}$ is a rank 2 homogeneous matrix with 7 degrees of freedom.
- **Point correspondence**: If $\mathbf{x}$ and $\mathbf{x}'$ are corresponding image points, then $\mathbf{x}'^{\mathsf{T}}\mathtt{F}\mathbf{x} = 0$.
- **Epipolar lines**:
  - $\mathbf{l}' = \mathtt{F}\mathbf{x}$ is the epipolar line corresponding to $\mathbf{x}$.
  - $\mathbf{l} = \mathtt{F}^{\mathsf{T}}\mathbf{x}'$ is the epipolar line corresponding to $\mathbf{x}'$.
- **Epipoles**:
  - $\mathtt{F}\mathbf{e} = \mathbf{0}$.
  - $\mathtt{F}^{\mathsf{T}}\mathbf{e}' = \mathbf{0}$.
- **Computation from camera matrices** $\mathtt{P}, \mathtt{P}'$:
  - General cameras, $\mathtt{F} = [\mathbf{e}']_\times \mathtt{P}'\mathtt{P}^+$, where $\mathtt{P}^+$ is the pseudo-inverse of $\mathtt{P}$, and $\mathbf{e}' = \mathtt{P}'\mathbf{C}$, with $\mathtt{P}\mathbf{C} = \mathbf{0}$.
  - Canonical cameras, $\mathtt{P} = [\mathtt{I} \mid \mathbf{0}]$, $\mathtt{P}' = [\mathtt{M} \mid \mathbf{m}]$, $\mathtt{F} = [\mathbf{e}']_\times \mathtt{M} = \mathtt{M}^{-\mathsf{T}}[\mathbf{e}]_\times$, where $\mathbf{e}' = \mathbf{m}$ and $\mathbf{e} = \mathtt{M}^{-1}\mathbf{m}$.
  - Cameras not at infinity $\mathtt{P} = \mathtt{K}[\mathtt{I} \mid \mathbf{0}]$, $\mathtt{P}' = \mathtt{K}'[\mathtt{R} \mid \mathbf{t}]$, $\mathtt{F} = \mathtt{K}'^{-\mathsf{T}}[\mathbf{t}]_\times \mathtt{R}\mathtt{K}^{-1} = [\mathtt{K}'\mathbf{t}]_\times \mathtt{K}'\mathtt{R}\mathtt{K}^{-1} = \mathtt{K}'^{-\mathsf{T}}\mathtt{R}\mathtt{K}^{\mathsf{T}}[\mathtt{K}\mathtt{R}^{\mathsf{T}}\mathbf{t}]_\times$.

Epipolar Line Homography

- The set of epipolar lines in each of the images forms a pencil of lines passing through the epipole
- Specifically, suppose $\mathbf{l}$ and $\mathbf{l}'$ are corresponding epipolar lines, and $\mathbf{k}$ is any line not passing through the epipole $\mathbf{e}$, then $\mathbf{l}$ and $\mathbf{l}'$ are related by $\mathbf{l}' = F[\mathbf{k}]_\times \mathbf{l}$ (symmetrically $\mathbf{l} = F^T[\mathbf{k}']_\times \mathbf{l}'$)

Epipolar Line Homography



- (a) The correspondence between epipolar lines $\mathbf{l}_i \leftrightarrow \mathbf{l'}_i$, is defined by the pencil of planes with axis the baseline
- (b) The corresponding lines are related by a perspective with center at point $\mathbf{p}$ on the baseline

Determining the Camera Matrices Between Views

- We have examined the properties of F and of image relations for a point correspondence $\mathbf{x} \leftrightarrow \mathbf{x}'$
- We now look at one of the most significant properities of F, that the matrix may be used to determine the camera matrices of the two views

Projective Invariance and Canonical Cameras

- The map $\mathbf{l}' = F\mathbf{x}$ and the correspondence condition ${\mathbf{x}'}^T F\mathbf{x} = 0$ are *projective* relationships
- In other words, the image relationships are projectively invariant: under a projective transformation of the image coordinates $\hat{\mathbf{x}} = H\mathbf{x}, \hat{\mathbf{x}'} = H'\mathbf{x}'$, there is a corresponding map $\hat{\mathbf{l}'} = \hat{F}\hat{\mathbf{x}}$ with $\hat{F} = H'^{-T}FH^{-1}$ the corresponding rank 2 fundamental matrix

Projective Invariance and Canonical Cameras

- Similarly, F only depends on projective properties of the cameras P, P'
- The camera matrix relates 3-space measurements to image measurements and so depends on both the image coordinate frame and the choice of world coordinate frame

Projective Invariance and Canonical Cameras

- F does not depend on the choice of world frame, for example a rotation of world coordinates changes P, P', but not F
- More precisely, if H is a 4×4 matrix representing a projective transformation of 3-space, then fundamental matrices corresponding to the pairs of camera matrices (P, P') and $(PH, P'H)$ are the same

Projective Invariance and Canonical Cameras

- Although a pair of camera matrices (P, P') uniquely determine a fundamental matrix F, the converse is not true
- The fundamental matrix determines the pair of camera matrices at best up to right-multiplication by a 3D projective transformation

Canonical Form of Camera Matrices

- Given this ambiguity, it is common to define a specific *canonical form* for the pair of camera matrices corresponding to a given fundamental matrix
- The first matrix is of the simple form $[I \mid \mathbf{0}]$, where I is the 3×3 identity matrix and $\mathbf{0}$ a null 3-vector
- Then, the fundamental matrix corresponding to a pair of camera matrices $P = [I \mid \mathbf{0}]$ and $P' = [M \mid \mathbf{m}]$ is equal to $[\mathbf{m}]_\times M$

Canonical Cameras Given F

- We want to derive a specific formula for a pair of cameras with canonical form given F
- To do this, we make use of the following characterization of F corresponding to the pair of camera matrices: *A nonzero matrix F is the fundamental matrix corresponding to a pair of camera matrices P and P' iff ${P'}^T FP$ is skew-symmetric*

  Proof: The condition that ${P'}^T FP$ is skew-symmetric is equivalent to $\mathbf{X}^T {P'}^T FP\mathbf{X} = 0$ for all $\mathbf{X}$. Setting $\mathbf{x}' = P'\mathbf{X}$ and $\mathbf{x} = P\mathbf{X}$, this is equivalent to ${\mathbf{x}'}^T F\mathbf{x} = 0$, which is the defining equation for the fundamental matrix.

Canonical Cameras Given F

- We can write a particular solution for the pairs of camera matrices in canonical form that correspond to a fundamental matrix as follows
- Let F be a fundamental matrix and S and skew-symmetric matrix
- Define the pair of camera matrices

$$P = [I \mid \mathbf{0}] \quad \text{and} \quad P' = [SF \mid \mathbf{e}']$$

where $\mathbf{e}'$ is the epipole such that ${\mathbf{e}'}^T F = \mathbf{0}$, and assume P' is a valid camera matrix (has rank 3), then F is the fundamental matrix corresponding to the pair (P, P')

Canonical Cameras Given F

- To demonstrate this, we can verify that

$$[SF \mid \mathbf{e}']^T F [I \mid \mathbf{0}] = \begin{bmatrix} F^T S^T F & \mathbf{0} \\ \mathbf{e}'^T F & 0 \end{bmatrix} = \begin{bmatrix} F^T S^T F & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix}$$

  which is skew-symmetric

- The skew-symmetric matrix S may be written in terms of its null vector as $S = [\mathbf{s}]_\times$

Canonical Cameras Given F

- A good choice for S is $S = [\mathbf{e}']_\times$, for in this case $\mathbf{e}'^T\mathbf{e}' \neq 0$
- Therefore, the camera matrices corresponding to a fundamental matrix F may be chosen as $P = [I \mid \mathbf{0}]$ and $P' = [[\mathbf{e}']_\times F \mid \mathbf{e}']$
- The general formula for a pair of canonic camera matrices corresponding to F is given by

$$P = [I \mid \mathbf{0}] \quad \text{and} \quad P' = [[\mathbf{e}']_\times F + \mathbf{e}'\mathbf{v}^T \mid \lambda\mathbf{e}']$$

where $\mathbf{v}$ is any 3-vector, and λ a non-zero scalar

Summary

- The fundamental matrix can be computed from correspondences of imaged scene points and does not require knowledge of the cameras' internal parameters or relative pose
- The camera matrices can be retrieved from the fundamental matrix up to a projective transformation of 3-space