Stereo

Raquel Urtasun

TTI Chicago

Feb 7, 2013

Today's Readings

- Chapter 11 of Szeliski's book

Let's look into stereo reconstruction

Stereo

Public Library, Stereoscopic Looking Room, Chicago, by Phillips, 1923

[Source: N. Snavely]

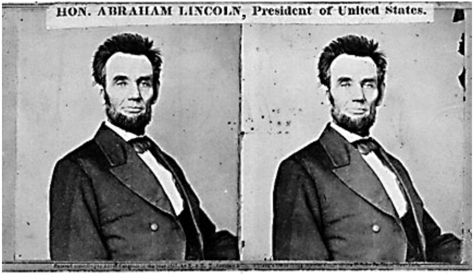
Stereo

- Stereo matching is the process of taking two or more images and **estimating a 3D model** of the scene by **finding matching pixels** in the images and converting their 2D positions into 3D depths
- We perceived depth based on the difference in appearance of the right and left eye.

Stereo

- Stereo matching is the process of taking two or more images and **estimating a 3D model** of the scene by **finding matching pixels** in the images and converting their 2D positions into 3D depths
- We perceived depth based on the difference in appearance of the right and left eye.

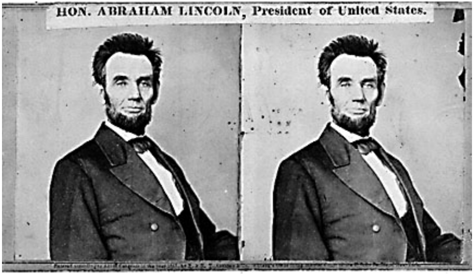
Stereo



Given two images from different viewpoints

- The depth is proportional to the inverse of the **disparity**
- How can we compute the depth of each point in the image?

Stereo

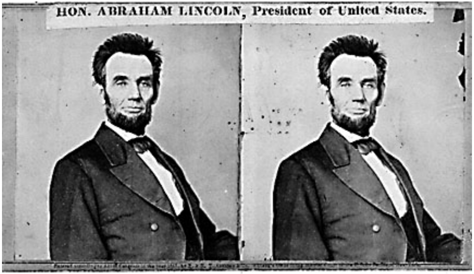


Given two images from different viewpoints

- The depth is proportional to the inverse of the **disparity**
- How can we compute the depth of each point in the image?
- Based on how much each pixel moves between the two images

[Source: N. Snavely]

Stereo

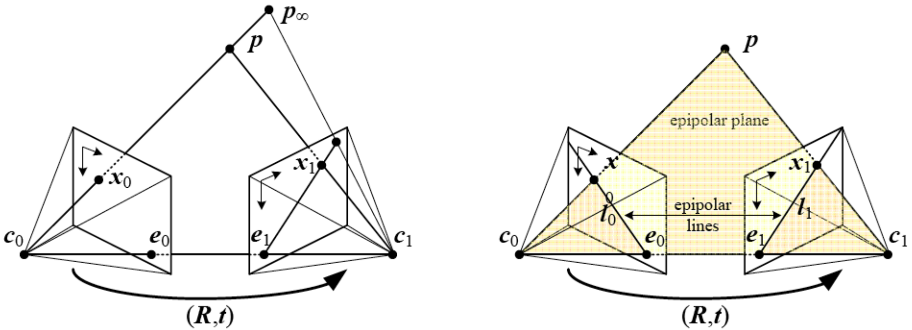


Given two images from different viewpoints

- The depth is proportional to the inverse of the **disparity**
- How can we compute the depth of each point in the image?
- Based on how much each pixel moves between the two images

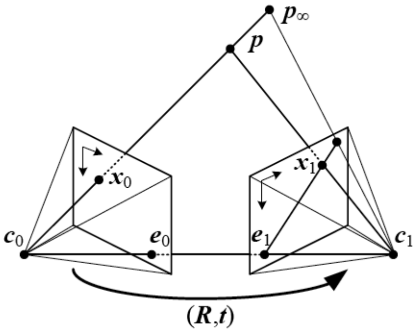
[Source: N. Snavely]

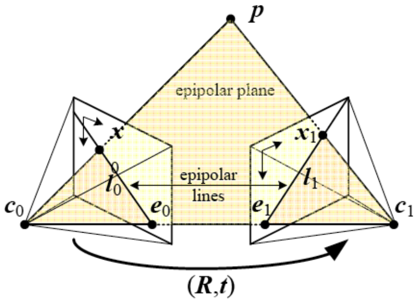
Epipolar Geometry



- Pixel in one image $\mathbf{x}_0$ projects to an **epipolar line** segment in the other image
- The segment is bounded at one end by the projection of the original viewing ray at infinity $\mathbf{p}_\infty$ and at the other end by the projection of the original camera center $\mathbf{c}_0$ into the second camera, which is known as the **epipole** $\mathbf{e}_1$.

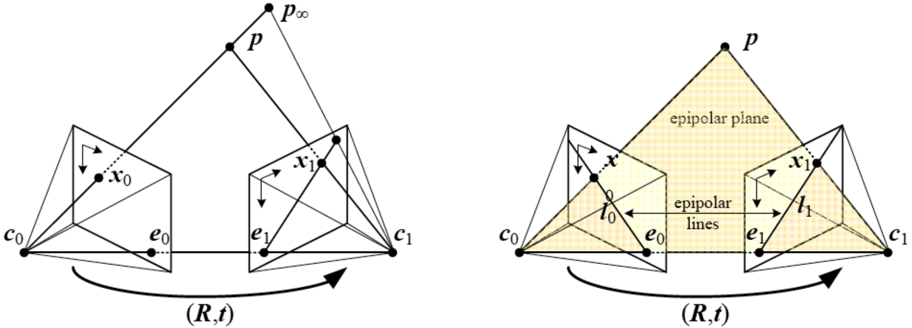
Epipolar Geometry





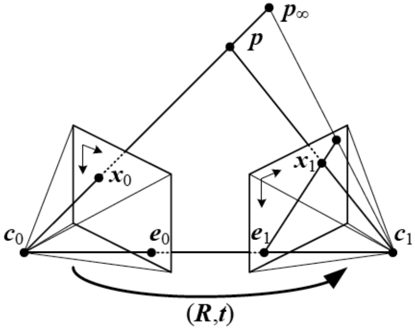
- Pixel in one image $\mathbf{x}_0$ projects to an **epipolar line** segment in the other image
- The segment is bounded at one end by the projection of the original viewing ray at infinity $\mathbf{p}_\infty$ and at the other end by the projection of the original camera center $\mathbf{c}_0$ into the second camera, which is known as the **epipole** $\mathbf{e}_1$.

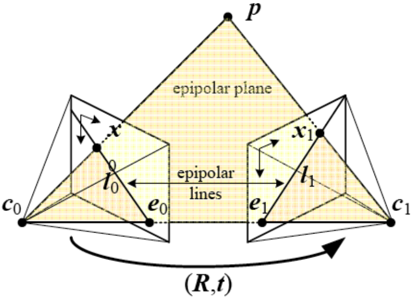
Epipolar Geometry



- If we project the epipolar line in the second image back into the first, we get another line (segment), this time bounded by the other corresponding **epipole** e_0
- Extending both line segments to infinity, we get a pair of corresponding epipolar lines, which are the intersection of the two image planes with the **epipolar plane** that passes through both camera centers $\mathbf{c}_0$ and $\mathbf{c}_1$ as well as the point of interest $\mathbf{p}$

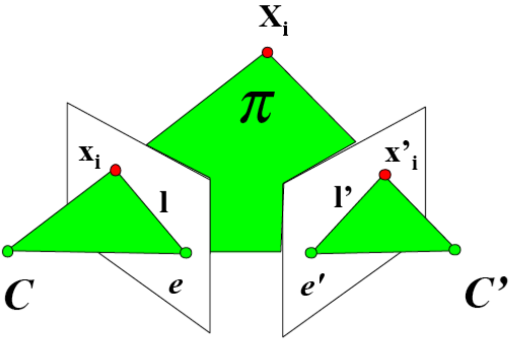
Epipolar Geometry





- If we project the epipolar line in the second image back into the first, we get another line (segment), this time bounded by the other corresponding **epipole** e_0
- Extending both line segments to infinity, we get a pair of corresponding epipolar lines, which are the intersection of the two image planes with the **epipolar plane** that passes through both camera centers $\mathbf{c}_0$ and $\mathbf{c}_1$ as well as the point of interest $\mathbf{p}$

Epipolar Plane



[Source: Ramani]

Rectification

- The epipolar geometry depends on the relative **pose** and **calibration** of the cameras
- This can be computed using the **fundamental matrix**

Rectification

- The epipolar geometry depends on the relative **pose** and **calibration** of the cameras
- This can be computed using the **fundamental matrix**
- Once this is computed, we can use the epipolar lines to restrict the search space to a 1D search

Rectification

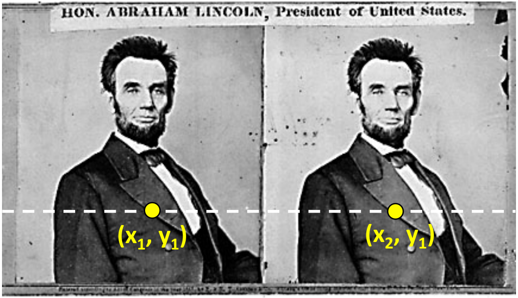
- The epipolar geometry depends on the relative **pose** and **calibration** of the cameras
- This can be computed using the **fundamental matrix**
- Once this is computed, we can use the epipolar lines to restrict the search space to a 1D search
- **Rectification**, the process of transforming the image so that the search is along horizontal line

Rectification

- The epipolar geometry depends on the relative **pose** and **calibration** of the cameras
- This can be computed using the **fundamental matrix**
- Once this is computed, we can use the epipolar lines to restrict the search space to a 1D search
- **Rectification**, the process of transforming the image so that the search is along horizontal line

Epipolar Geometry

epipolar
lines

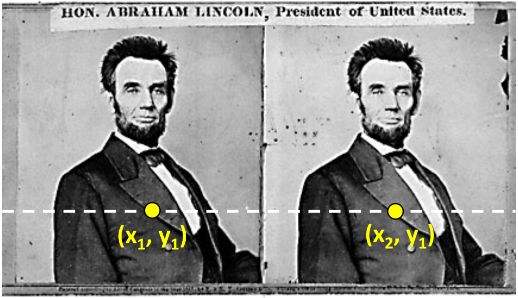


Two images captured by a purely horizontal translating camera
(*rectified* stereo pair)

- The disparity for pixel (x_1, y_1) is $(x_2 - x_1)$ if the images are rectified
- This is a one dimensional search for each pixel

Epipolar Geometry

epipolar lines

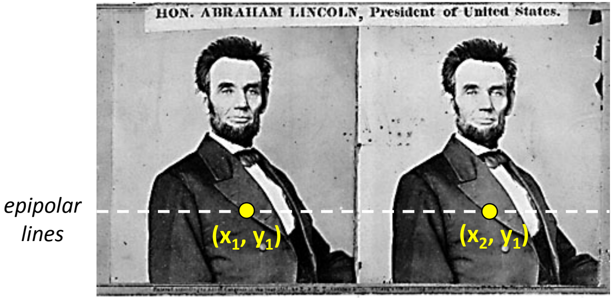


Two images captured by a purely horizontal translating camera
(*rectified* stereo pair)

- The disparity for pixel (x_1, y_1) is $(x_2 - x_1)$ if the images are rectified
- This is a one dimensional search for each pixel
- Very challenging to estimate the correspondences

[Source: N. Snavely]

Epipolar Geometry

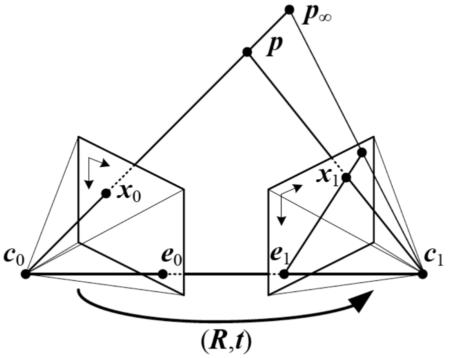


Two images captured by a purely horizontal translating camera
(*rectified* stereo pair)

- The disparity for pixel (x_1, y_1) is $(x_2 - x_1)$ if the images are rectified
- This is a one dimensional search for each pixel
- Very challenging to estimate the correspondences

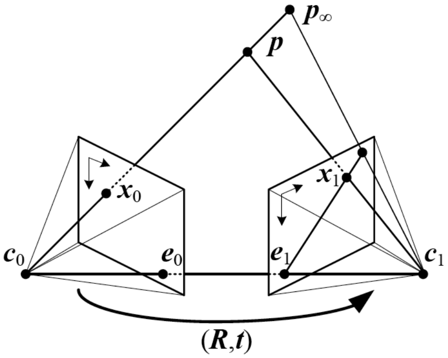
[Source: N. Snavely]

Fundamental Matrix



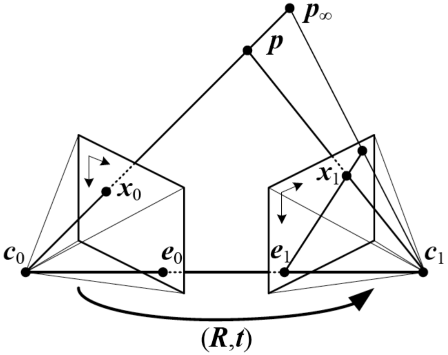
- Projective geometry depends only on the cameras internal parameters and relative pose of cameras (and not the 3D scene)
- Fundamental matrix **F** encapsulates this geometry

Fundamental Matrix



- Projective geometry depends only on the cameras internal parameters and relative pose of cameras (and not the 3D scene)
- Fundamental matrix **F** encapsulates this geometry

Computation of Fundamental Matrix

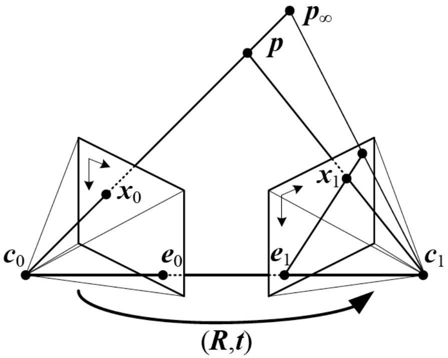


- We will show that for any pair of corresponding points in both images

$$\mathbf{x}_0^T \mathbf{F} \mathbf{x}_1 = 0$$

- **F** can be computed from correspondences between image points alone
- No knowledge of camera internal parameters required

Computation of Fundamental Matrix

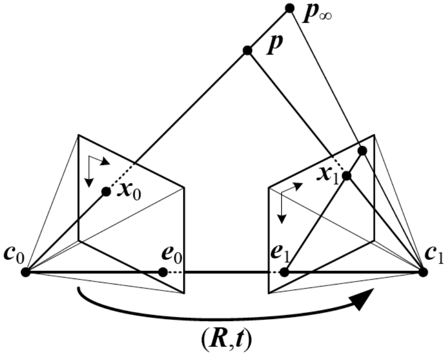


- We will show that for any pair of corresponding points in both images

$$\mathbf{x}_0^T \mathbf{F} \mathbf{x}_1 = 0$$

- **F** can be computed from correspondences between image points alone
- No knowledge of camera internal parameters required
- No knowledge of relative pose required

Computation of Fundamental Matrix

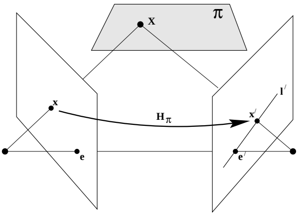


- We will show that for any pair of corresponding points in both images

$$\mathbf{x}_0^T \mathbf{F} \mathbf{x}_1 = 0$$

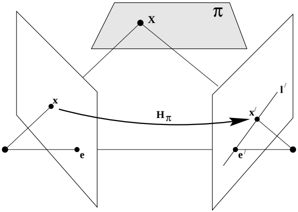
- **F** can be computed from correspondences between image points alone
- No knowledge of camera internal parameters required
- No knowledge of relative pose required

Fundamental Matrix and Projective Geometry



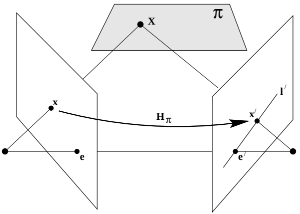
- Take **x** in camera **P** and find scene point **X** on ray of **x** in camera **P**
- Find the image **x**′ of **X** in camera **P**′

Fundamental Matrix and Projective Geometry



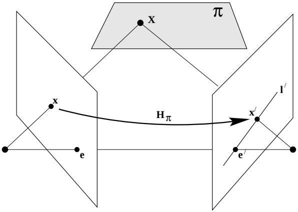
- Take $\mathbf{x}$ in camera $\mathbf{P}$ and find scene point $\mathbf{X}$ on ray of $\mathbf{x}$ in camera $\mathbf{P}$
- Find the image $\mathbf{x}'$ of $\mathbf{X}$ in camera $\mathbf{P}'$
- Find epipole $\mathbf{e}'$ as image of $\mathbf{C}$ in camera $\mathbf{P}'$, $\mathbf{e}' = \mathbf{P}'\mathbf{C}$

Fundamental Matrix and Projective Geometry



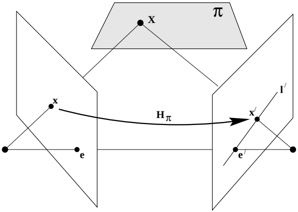
- Take $\mathbf{x}$ in camera $\mathbf{P}$ and find scene point $\mathbf{X}$ on ray of $\mathbf{x}$ in camera $\mathbf{P}$
- Find the image $\mathbf{x}'$ of $\mathbf{X}$ in camera $\mathbf{P}'$
- Find epipole $\mathbf{e}'$ as image of $\mathbf{C}$ in camera $\mathbf{P}'$, $\mathbf{e}' = \mathbf{P}'\mathbf{C}$
- Find epipolar line $\mathbf{l}'$ from $\mathbf{e}'$ to $\mathbf{x}'$ in $\mathbf{P}'$ as function of $\mathbf{x}$: $\mathbf{l}' = \mathbf{e}' \times \mathbf{x}'$

Fundamental Matrix and Projective Geometry



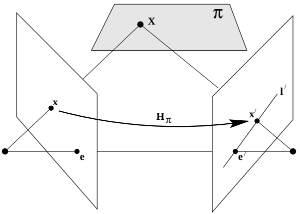
- Take $\mathbf{x}$ in camera $\mathbf{P}$ and find scene point $\mathbf{X}$ on ray of $\mathbf{x}$ in camera $\mathbf{P}$
- Find the image $\mathbf{x}'$ of $\mathbf{X}$ in camera $\mathbf{P}'$
- Find epipole $\mathbf{e}'$ as image of $\mathbf{C}$ in camera $\mathbf{P}'$, $\mathbf{e}' = \mathbf{P}'\mathbf{C}$
- Find epipolar line $\mathbf{l}'$ from $\mathbf{e}'$ to $\mathbf{x}'$ in $\mathbf{P}'$ as function of $\mathbf{x}$: $\mathbf{l}' = \mathbf{e}' \times \mathbf{x}'$
- $\mathbf{x}'$ and $\mathbf{x}$ are related via homography: $\mathbf{x}' = H_\pi \mathbf{x}$

Fundamental Matrix and Projective Geometry



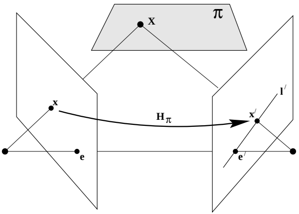
- Take $\mathbf{x}$ in camera $\mathbf{P}$ and find scene point $\mathbf{X}$ on ray of $\mathbf{x}$ in camera $\mathbf{P}$
- Find the image $\mathbf{x}'$ of $\mathbf{X}$ in camera $\mathbf{P}'$
- Find epipole $\mathbf{e}'$ as image of $\mathbf{C}$ in camera $\mathbf{P}'$, $\mathbf{e}' = \mathbf{P}'\mathbf{C}$
- Find epipolar line $\mathbf{l}'$ from $\mathbf{e}'$ to $\mathbf{x}'$ in $\mathbf{P}'$ as function of $\mathbf{x}$: $\mathbf{l}' = \mathbf{e}' \times \mathbf{x}'$
- $\mathbf{x}'$ and $\mathbf{x}$ are related via homography: $\mathbf{x}' = H_\pi \mathbf{x}$
- $\mathbf{l}' = \mathbf{e}' \times \mathbf{x}' = \mathbf{e}' \times H_\pi \mathbf{x} = \mathbf{F}\mathbf{x}$

Fundamental Matrix and Projective Geometry



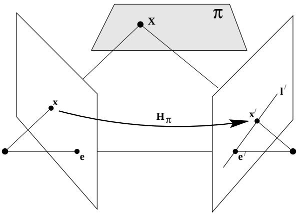
- Take $\mathbf{x}$ in camera $\mathbf{P}$ and find scene point $\mathbf{X}$ on ray of $\mathbf{x}$ in camera $\mathbf{P}$
- Find the image $\mathbf{x}'$ of $\mathbf{X}$ in camera $\mathbf{P}'$
- Find epipole $\mathbf{e}'$ as image of $\mathbf{C}$ in camera $\mathbf{P}'$, $\mathbf{e}' = \mathbf{P}'\mathbf{C}$
- Find epipolar line $\mathbf{l}'$ from $\mathbf{e}'$ to $\mathbf{x}'$ in $\mathbf{P}'$ as function of $\mathbf{x}$: $\mathbf{l}' = \mathbf{e}' \times \mathbf{x}'$
- $\mathbf{x}'$ and $\mathbf{x}$ are related via homography: $\mathbf{x}' = H_\pi \mathbf{x}$
- $\mathbf{l}' = \mathbf{e}' \times \mathbf{x}' = \mathbf{e}' \times H_\pi \mathbf{x} = \mathbf{F}\mathbf{x}$
- The fundamental matrix $\mathbf{F}$ is defined $\mathbf{l}' = \mathbf{F}\mathbf{x}$

Fundamental Matrix and Projective Geometry



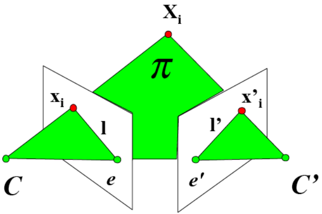
- Take **x** in camera **P** and find scene point **X** on ray of **x** in camera **P**
- Find the image $\mathbf{x}'$ of **X** in camera $\mathbf{P}'$
- Find epipole $\mathbf{e}'$ as image of **C** in camera $\mathbf{P}'$, $\mathbf{e}' = \mathbf{P}'\mathbf{C}$
- Find epipolar line $\mathbf{l}'$ from $\mathbf{e}'$ to $\mathbf{x}'$ in $\mathbf{P}'$ as function of **x**: $\mathbf{l}' = \mathbf{e}' \times \mathbf{x}'$
- $\mathbf{x}'$ and **x** are related via homography: $\mathbf{x}' = H_\pi \mathbf{x}$
- $\mathbf{l}' = \mathbf{e}' \times \mathbf{x}' = \mathbf{e}' \times H_\pi \mathbf{x} = \mathbf{F}\mathbf{x}$
- The fundamental matrix **F** is defined $\mathbf{l}' = \mathbf{F}\mathbf{x}$
- $\mathbf{x}'$ belongs to $\mathbf{l}'$, so $\mathbf{x}'^T \mathbf{l}' = 0$, so $\mathbf{x}'^T \mathbf{F}\mathbf{x} = 0$

Fundamental Matrix and Projective Geometry



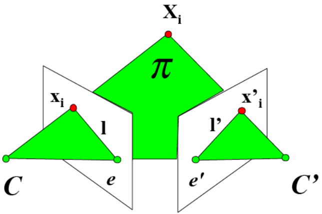
- Take $\mathbf{x}$ in camera $\mathbf{P}$ and find scene point $\mathbf{X}$ on ray of $\mathbf{x}$ in camera $\mathbf{P}$
- Find the image $\mathbf{x}'$ of $\mathbf{X}$ in camera $\mathbf{P}'$
- Find epipole $\mathbf{e}'$ as image of $\mathbf{C}$ in camera $\mathbf{P}'$, $\mathbf{e}' = \mathbf{P}'\mathbf{C}$
- Find epipolar line $\mathbf{l}'$ from $\mathbf{e}'$ to $\mathbf{x}'$ in $\mathbf{P}'$ as function of $\mathbf{x}$: $\mathbf{l}' = \mathbf{e}' \times \mathbf{x}'$
- $\mathbf{x}'$ and $\mathbf{x}$ are related via homography: $\mathbf{x}' = H_\pi \mathbf{x}$
- $\mathbf{l}' = \mathbf{e}' \times \mathbf{x}' = \mathbf{e}' \times H_\pi \mathbf{x} = \mathbf{F}\mathbf{x}$
- The fundamental matrix $\mathbf{F}$ is defined $\mathbf{l}' = \mathbf{F}\mathbf{x}$
- $\mathbf{x}'$ belongs to $\mathbf{l}'$, so $\mathbf{x}'^T \mathbf{l}' = 0$, so $\mathbf{x}'^T \mathbf{F}\mathbf{x} = 0$

Finding the Fundamental Matrix from known Projections



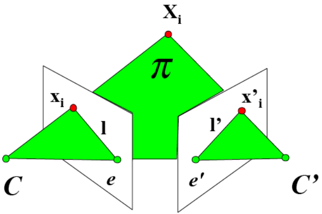
- Take $\mathbf{x}$ in camera $\mathbf{P}$ and find one scene point on ray from $\mathbf{C}$ to $\mathbf{x}$
- Point $\mathbf{X} = \mathbf{P}^{+}\mathbf{x}$ satisfies $\mathbf{x} = \mathbf{P}\mathbf{X}$ with $\mathbf{P}^{+} = \mathbf{P}^{T}(\mathbf{P}\mathbf{P}^{T})^{-1}$ so $\mathbf{P}\mathbf{X} = \mathbf{P}\mathbf{P}^{T}(\mathbf{P}\mathbf{P}^{T})^{-1}\mathbf{x} = \mathbf{x}$

Finding the Fundamental Matrix from known Projections



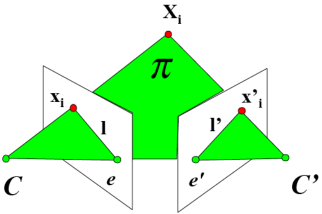
- Take $\mathbf{x}$ in camera $\mathbf{P}$ and find one scene point on ray from $\mathbf{C}$ to $\mathbf{x}$
- Point $\mathbf{X} = \mathbf{P}^+\mathbf{x}$ satisfies $\mathbf{x} = \mathbf{PX}$ with $\mathbf{P}^+ = \mathbf{P}^T(\mathbf{PP}^T)^{-1}$ so $\mathbf{PX} = \mathbf{PP}^T(\mathbf{PP}^T)^{-1}\mathbf{x} = \mathbf{x}$
- Image of this point in camera $\mathbf{P}'$ is $\mathbf{x}' = \mathbf{P}'\mathbf{X} = \mathbf{P}'\mathbf{P}^+\mathbf{x}$

Finding the Fundamental Matrix from known Projections



- Take $\mathbf{x}$ in camera $\mathbf{P}$ and find one scene point on ray from $\mathbf{C}$ to $\mathbf{x}$
- Point $\mathbf{X} = \mathbf{P}^+\mathbf{x}$ satisfies $\mathbf{x} = \mathbf{PX}$ with $\mathbf{P}^+ = \mathbf{P}^T(\mathbf{PP}^T)^{-1}$ so $\mathbf{PX} = \mathbf{PP}^T(\mathbf{PP}^T)^{-1}\mathbf{x} = \mathbf{x}$
- Image of this point in camera $\mathbf{P}'$ is $\mathbf{x}' = \mathbf{P}'\mathbf{X} = \mathbf{P}'\mathbf{P}^+\mathbf{x}$
- Image of $\mathbf{C}$ in camera $\mathbf{P}'$ is epipole $\mathbf{e}' = \mathbf{P}'\mathbf{C}$

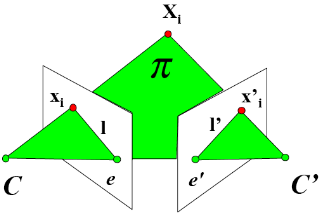
Finding the Fundamental Matrix from known Projections



- Take $\mathbf{x}$ in camera $\mathbf{P}$ and find one scene point on ray from $\mathbf{C}$ to $\mathbf{x}$
- Point $\mathbf{X} = \mathbf{P}^{+}\mathbf{x}$ satisfies $\mathbf{x} = \mathbf{P}\mathbf{X}$ with $\mathbf{P}^{+} = \mathbf{P}^{T}(\mathbf{P}\mathbf{P}^{T})^{-1}$ so $\mathbf{P}\mathbf{X} = \mathbf{P}\mathbf{P}^{T}(\mathbf{P}\mathbf{P}^{T})^{-1}\mathbf{x} = \mathbf{x}$
- Image of this point in camera $\mathbf{P}'$ is $\mathbf{x}' = \mathbf{P}'\mathbf{X} = \mathbf{P}'\mathbf{P}^{+}\mathbf{x}$
- Image of $\mathbf{C}$ in camera $\mathbf{P}'$ is epipole $\mathbf{e}' = \mathbf{P}'\mathbf{C}$
- Epipolar line of $\mathbf{x}$ in $\mathbf{P}'$ is

$$\mathbf{l}' = (\mathbf{e}') \times (\mathbf{P}'\mathbf{P}^{+}\mathbf{x})$$

Finding the Fundamental Matrix from known Projections



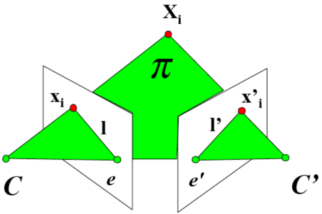
- Take $\mathbf{x}$ in camera $\mathbf{P}$ and find one scene point on ray from $\mathbf{C}$ to $\mathbf{x}$
- Point $\mathbf{X} = \mathbf{P}^+\mathbf{x}$ satisfies $\mathbf{x} = \mathbf{P}\mathbf{X}$ with $\mathbf{P}^+ = \mathbf{P}^T(\mathbf{P}\mathbf{P}^T)^{-1}$ so $\mathbf{P}\mathbf{X} = \mathbf{P}\mathbf{P}^T(\mathbf{P}\mathbf{P}^T)^{-1}\mathbf{x} = \mathbf{x}$
- Image of this point in camera $\mathbf{P}'$ is $\mathbf{x}' = \mathbf{P}'\mathbf{X} = \mathbf{P}'\mathbf{P}^+\mathbf{x}$
- Image of $\mathbf{C}$ in camera $\mathbf{P}'$ is epipole $\mathbf{e}' = \mathbf{P}'\mathbf{C}$
- Epipolar line of $\mathbf{x}$ in $\mathbf{P}'$ is

$$\mathbf{l}' = (\mathbf{e}') \times (\mathbf{P}'\mathbf{P}^+\mathbf{x})$$

- $\mathbf{l}' = \mathbf{F}\mathbf{x}$ defines the fundamental matrix

$$\mathbf{F} = (\mathbf{P}'\mathbf{C}) \times (\mathbf{P}'\mathbf{P}^+)$$

Finding the Fundamental Matrix from known Projections



- Take $\mathbf{x}$ in camera $\mathbf{P}$ and find one scene point on ray from $\mathbf{C}$ to $\mathbf{x}$
- Point $\mathbf{X} = \mathbf{P}^+\mathbf{x}$ satisfies $\mathbf{x} = \mathbf{PX}$ with $\mathbf{P}^+ = \mathbf{P}^T(\mathbf{PP}^T)^{-1}$ so $\mathbf{PX} = \mathbf{PP}^T(\mathbf{PP}^T)^{-1}\mathbf{x} = \mathbf{x}$
- Image of this point in camera $\mathbf{P}'$ is $\mathbf{x}' = \mathbf{P}'\mathbf{X} = \mathbf{P}'\mathbf{P}^+\mathbf{x}$
- Image of $\mathbf{C}$ in camera $\mathbf{P}'$ is epipole $\mathbf{e}' = \mathbf{P}'\mathbf{C}$
- Epipolar line of $\mathbf{x}$ in $\mathbf{P}'$ is

$$\mathbf{l}' = (\mathbf{e}') \times (\mathbf{P}'\mathbf{P}^+\mathbf{x})$$

- $\mathbf{l}' = \mathbf{Fx}$ defines the fundamental matrix

$$\mathbf{F} = (\mathbf{P}'\mathbf{C}) \times (\mathbf{P}'\mathbf{P}^+)$$

Properties of the fundamental matrix

- Matrix 3×3 since $\mathbf{x'}^T \mathbf{F} \mathbf{x} = 0$
- Let $\mathbf{F}$ be the fundamental matrix of camera pair $(\mathbf{P}, \mathbf{P}')$, the fundamental matrix of camera pair $(\mathbf{P}', \mathbf{P})$ is $\mathbf{F}' = F^T$

Properties of the fundamental matrix

- Matrix 3×3 since $\mathbf{x'}^T \mathbf{F} \mathbf{x} = 0$
- Let $\mathbf{F}$ be the fundamental matrix of camera pair $(\mathbf{P}, \mathbf{P'})$, the fundamental matrix of camera pair $(\mathbf{P'}, \mathbf{P})$ is $\mathbf{F'} = F^T$
- This is true since $\mathbf{x}^T \mathbf{F'} \mathbf{x'} = 0$ implies $\mathbf{x'}^T \mathbf{F'}^T \mathbf{x} = 0$, so $\mathbf{F'} = \mathbf{F}^T$

Properties of the fundamental matrix

- Matrix 3×3 since $\mathbf{x}'^T \mathbf{F} \mathbf{x} = 0$
- Let $\mathbf{F}$ be the fundamental matrix of camera pair $(\mathbf{P}, \mathbf{P}')$, the fundamental matrix of camera pair $(\mathbf{P}', \mathbf{P})$ is $\mathbf{F}' = F^T$
- This is true since $\mathbf{x}^T \mathbf{F}' \mathbf{x}' = 0$ implies $\mathbf{x}'^T \mathbf{F}'^T \mathbf{x} = 0$, so $\mathbf{F}' = \mathbf{F}^T$
- Epipolar line of $\mathbf{x}$ is $\mathbf{l}' = \mathbf{F}\mathbf{x}$

Properties of the fundamental matrix

- Matrix 3×3 since $\mathbf{x'}^T \mathbf{F} \mathbf{x} = 0$
- Let $\mathbf{F}$ be the fundamental matrix of camera pair $(\mathbf{P}, \mathbf{P}')$, the fundamental matrix of camera pair $(\mathbf{P}', \mathbf{P})$ is $\mathbf{F}' = F^T$
- This is true since $\mathbf{x}^T \mathbf{F}' \mathbf{x}' = 0$ implies $\mathbf{x'}^T \mathbf{F'}^T \mathbf{x} = 0$, so $\mathbf{F}' = \mathbf{F}^T$
- Epipolar line of $\mathbf{x}$ is $\mathbf{l}' = \mathbf{F}\mathbf{x}$
- Epipolar line of $\mathbf{x}'$ is $\mathbf{l} = \mathbf{F}^T \mathbf{x}'$

Properties of the fundamental matrix

- Matrix 3×3 since $\mathbf{x'}^T \mathbf{F} \mathbf{x} = 0$
- Let $\mathbf{F}$ be the fundamental matrix of camera pair $(\mathbf{P}, \mathbf{P'})$, the fundamental matrix of camera pair $(\mathbf{P'}, \mathbf{P})$ is $\mathbf{F'} = F^T$
- This is true since $\mathbf{x}^T \mathbf{F'} \mathbf{x'} = 0$ implies $\mathbf{x'}^T \mathbf{F'}^T \mathbf{x} = 0$, so $\mathbf{F'} = \mathbf{F}^T$
- Epipolar line of $\mathbf{x}$ is $\mathbf{l'} = \mathbf{F}\mathbf{x}$
- Epipolar line of $\mathbf{x'}$ is $\mathbf{l} = \mathbf{F}^T \mathbf{x'}$
- Epipole $\mathbf{e'}$ is left null space of $\mathbf{F}$, and $\mathbf{e}$ is right null space.

Properties of the fundamental matrix

- Matrix 3×3 since $\mathbf{x'}^T \mathbf{F} \mathbf{x} = 0$
- Let $\mathbf{F}$ be the fundamental matrix of camera pair $(\mathbf{P}, \mathbf{P'})$, the fundamental matrix of camera pair $(\mathbf{P'}, \mathbf{P})$ is $\mathbf{F'} = F^T$
- This is true since $\mathbf{x}^T \mathbf{F'} \mathbf{x'} = 0$ implies $\mathbf{x'}^T \mathbf{F'}^T \mathbf{x} = 0$, so $\mathbf{F'} = \mathbf{F}^T$
- Epipolar line of $\mathbf{x}$ is $\mathbf{l'} = \mathbf{F}\mathbf{x}$
- Epipolar line of $\mathbf{x'}$ is $\mathbf{l} = \mathbf{F}^T \mathbf{x'}$
- Epipole $\mathbf{e'}$ is left null space of $\mathbf{F}$, and $\mathbf{e}$ is right null space.
- All epipolar lines $\mathbf{l'}$ contains epipole e', so $\mathbf{e'}^T \mathbf{l'} = 0$, i.e. $\mathbf{e'}^T \mathbf{F} \mathbf{x} = 0$ for all $\mathbf{x}$, therefore $\mathbf{e'}^T \mathbf{F} = 0$

Properties of the fundamental matrix

- Matrix 3×3 since $\mathbf{x'}^T \mathbf{F} \mathbf{x} = 0$
- Let $\mathbf{F}$ be the fundamental matrix of camera pair $(\mathbf{P}, \mathbf{P'})$, the fundamental matrix of camera pair $(\mathbf{P'}, \mathbf{P})$ is $\mathbf{F'} = F^T$
- This is true since $\mathbf{x}^T \mathbf{F'} \mathbf{x'} = 0$ implies $\mathbf{x'}^T \mathbf{F'}^T \mathbf{x} = 0$, so $\mathbf{F'} = \mathbf{F}^T$
- Epipolar line of $\mathbf{x}$ is $\mathbf{l'} = \mathbf{F}\mathbf{x}$
- Epipolar line of $\mathbf{x'}$ is $\mathbf{l} = \mathbf{F}^T \mathbf{x'}$
- Epipole $\mathbf{e'}$ is left null space of $\mathbf{F}$, and $\mathbf{e}$ is right null space.
- All epipolar lines $\mathbf{l'}$ contains epipole e', so $\mathbf{e'}^T \mathbf{l'} = 0$, i.e. $\mathbf{e'}^T \mathbf{F}\mathbf{x} = 0$ for all $\mathbf{x}$, therefore $\mathbf{e'}^T \mathbf{F} = 0$
- $\mathbf{F}$ is of rank 2 because $\mathbf{F} = \mathbf{e'} \times (\mathbf{P'}\mathbf{P}^+)$ and $\mathbf{e'}\times$ is of rank 2

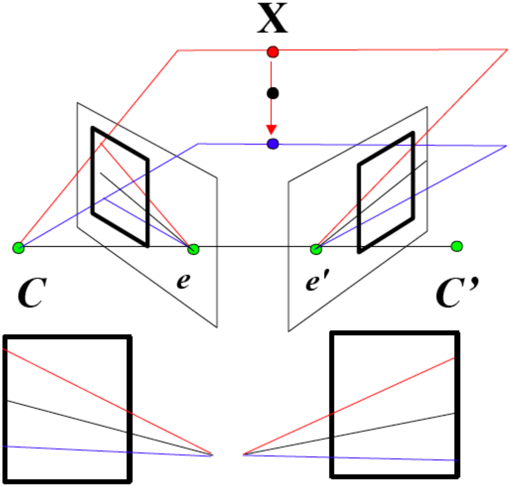
Properties of the fundamental matrix

- Matrix 3×3 since $\mathbf{x'}^T \mathbf{F} \mathbf{x} = 0$
- Let $\mathbf{F}$ be the fundamental matrix of camera pair $(\mathbf{P}, \mathbf{P'})$, the fundamental matrix of camera pair $(\mathbf{P'}, \mathbf{P})$ is $\mathbf{F'} = F^T$
- This is true since $\mathbf{x}^T \mathbf{F'} \mathbf{x'} = 0$ implies $\mathbf{x'}^T \mathbf{F'}^T \mathbf{x} = 0$, so $\mathbf{F'} = \mathbf{F}^T$
- Epipolar line of $\mathbf{x}$ is $\mathbf{l'} = \mathbf{F}\mathbf{x}$
- Epipolar line of $\mathbf{x'}$ is $\mathbf{l} = \mathbf{F}^T \mathbf{x'}$
- Epipole $\mathbf{e'}$ is left null space of $\mathbf{F}$, and $\mathbf{e}$ is right null space.
- All epipolar lines $\mathbf{l'}$ contains epipole e', so $\mathbf{e'}^T \mathbf{l'} = 0$, i.e. $\mathbf{e'}^T \mathbf{F}\mathbf{x} = 0$ for all $\mathbf{x}$, therefore $\mathbf{e'}^T \mathbf{F} = 0$
- $\mathbf{F}$ is of rank 2 because $\mathbf{F} = \mathbf{e'} \times (\mathbf{P'}\mathbf{P}^+)$ and $\mathbf{e'}\times$ is of rank 2
- $\mathbf{F}$ has 7 degrees of freedom, there are 9 elements, but scaling is not important and $det(\mathbf{F}) = 0$ removes one degree of freedom

Properties of the fundamental matrix

- Matrix 3×3 since $\mathbf{x}'^T \mathbf{F} \mathbf{x} = 0$
- Let $\mathbf{F}$ be the fundamental matrix of camera pair $(\mathbf{P}, \mathbf{P}')$, the fundamental matrix of camera pair $(\mathbf{P}', \mathbf{P})$ is $\mathbf{F}' = F^T$
- This is true since $\mathbf{x}^T \mathbf{F}' \mathbf{x}' = 0$ implies $\mathbf{x}'^T \mathbf{F}'^T \mathbf{x} = 0$, so $\mathbf{F}' = \mathbf{F}^T$
- Epipolar line of $\mathbf{x}$ is $\mathbf{l}' = \mathbf{F}\mathbf{x}$
- Epipolar line of $\mathbf{x}'$ is $\mathbf{l} = \mathbf{F}^T \mathbf{x}'$
- Epipole $\mathbf{e}'$ is left null space of $\mathbf{F}$, and $\mathbf{e}$ is right null space.
- All epipolar lines $\mathbf{l}'$ contains epipole e', so $\mathbf{e}'^T \mathbf{l}' = 0$, i.e. $\mathbf{e}'^T \mathbf{F} \mathbf{x} = 0$ for all $\mathbf{x}$, therefore $\mathbf{e}'^T \mathbf{F} = 0$
- $\mathbf{F}$ is of rank 2 because $\mathbf{F} = \mathbf{e}' \times (\mathbf{P}' \mathbf{P}^+)$ and $\mathbf{e}' \times$ is of rank 2
- $\mathbf{F}$ has 7 degrees of freedom, there are 9 elements, but scaling is not important and $det(\mathbf{F}) = 0$ removes one degree of freedom

Pensils of Epipolar Lines



[Source: Ramani]

Mapping between epipolar lines (Homography)

- Define $\mathbf{x}$ as intersection between line $\mathbf{l}$ and a line $\mathbf{k}$ (that doesn't pass through $\mathbf{e}$)

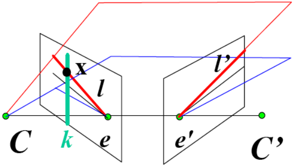
$$\begin{aligned} \mathbf{x} &= \mathbf{k} \times \mathbf{l} \\ \mathbf{l}' &= \mathbf{F}\mathbf{x} = \mathbf{F}(\mathbf{k} \times \mathbf{l}) \end{aligned}$$

- We can also write

$$\mathbf{l}' = \mathbf{F}\mathbf{x} = \mathbf{F}(\mathbf{e} \times \mathbf{l})$$

- and similarly

$$\mathbf{l} = \mathbf{F}^T\mathbf{x}' = \mathbf{F}^T(\mathbf{e}' \times \mathbf{l}')$$



[Source: Ramani]

Retrieving Camera Matrices from $\mathbf{F}$

- Select world coordinates as camera coordinates of first camera, select focal length = 1, and count pixels from the principal point. Then $\mathbf{P} = [\mathbf{I}_3, 0]$
- Then $\mathbf{P} = [\mathbf{SF}|\mathbf{e}']$ with $\mathbf{S}$ any skew-symmetric matrix is a solution
- How do we prove this?

$$\mathbf{x'}^T\mathbf{Fx} = \mathbf{X}^T\mathbf{P'}^T\mathbf{FPX}$$

- The middle part is skew symmetric

$$\mathbf{P'}^T\mathbf{FP} = [\mathbf{SF}|\mathbf{e}']^T F[\mathbf{I}_3|0]$$

The middle part is skew symmetric

$$\mathbf{P'}^T\mathbf{F}\mathbf{P} = [\mathbf{S}\mathbf{F}|\mathbf{e'}]^T\mathbf{F}[\mathbf{I}_3|0] = \begin{bmatrix} \mathbf{F}^T\mathbf{S}^T\mathbf{F} & 0_3 \\ \mathbf{e'}^T\mathbf{F} & 0 \end{bmatrix}$$

- $\mathbf{e'}^T\mathbf{F} = 0$ because $\mathbf{e'}$ is left null space of $\mathbf{F}$
- $\mathbf{F}^T\mathbf{S}^T\mathbf{F}$ is skew-symmetric for any $\mathbf{F}$ and any skew-symmetric $\mathbf{S}$

$$\mathbf{P'}^T\mathbf{F}\mathbf{P} = [\mathbf{S}\mathbf{F}|\mathbf{e'}]^T\mathbf{F}[\mathbf{I}_3|0] = \begin{bmatrix} \mathbf{F}^T\mathbf{S}^T\mathbf{F} & 0_3 \\ 0 & 0 \end{bmatrix}$$

Retrieving Camera Matrices from $\mathbf{F}$

- Select world coordinates as camera coordinates of first camera, select focal length $= 1$, and count pixels from the principal point. Then $\mathbf{P} = [\mathbf{I}_3, 0]$
- Then $\mathbf{P} = [\mathbf{SF}|\mathbf{e}']$ with $\mathbf{S}$ any skew-symmetric matrix is a solution
- How do we proof this?

$$\mathbf{x}'^T\mathbf{F}\mathbf{x} = \mathbf{X}^T\mathbf{P}'^T\mathbf{F}\mathbf{P}\mathbf{X}$$

- The middle part is skew symmetric

$$\mathbf{P}'^T\mathbf{F}\mathbf{P} = [\mathbf{SF}|\mathbf{e}']^T\mathbf{F}[\mathbf{I}_3|0]$$

- For any skew-symmetric matrix $\mathbf{S}'$ and any $\mathbf{X}$

$$\mathbf{X}^T\mathbf{S}'\mathbf{X} = 0$$

- A good choice is $\mathbf{S} = [\mathbf{e}']_\times$, therefore

$$\mathbf{P}' = [(\mathbf{e}' \times \mathbf{F}), \mathbf{e}']$$

Notation $[\mathbf{a}]_\times$

- $[\mathbf{a}']_\times$ is defined as:

$$[\mathbf{a}]_\times = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

- $\mathbf{a} \times \mathbf{b} = [\mathbf{a}]_\times \mathbf{b} = \left(\mathbf{a}^T [\mathbf{b}]_\times\right)^T$
- $[\mathbf{a}]_\times \mathbf{a} = 0$

Essential Matrix

- Specialization of fundamental matrix for calibrated cameras and normalized coordinates

$$\mathbf{x} = \mathbf{P}\mathbf{X}$$

- Normalized coordinates $\mathbf{x}_0 = \mathbf{K}^{-1}\mathbf{x}$
- Consider pair of normalized cameras

$$\mathbf{P} = [\mathbf{I}|0], \quad \mathbf{P}' = [\mathbf{R}|\mathbf{T}]$$

- Then we compute

$$\mathbf{F} = [\mathbf{P}'\mathbf{C}]_\times \mathbf{P}'\mathbf{P}^+$$

with

$$[\mathbf{P}'\mathbf{C}] = [\mathbf{R}|\mathbf{T}] \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = [\mathbf{T}]$$

$$\mathbf{P}^+ = \begin{bmatrix} \mathbf{I}_3 \\ 0_3^T \end{bmatrix} \qquad \mathbf{P}'\mathbf{P}^+ = \mathbf{R} \qquad \rightarrow \mathbf{F} = \mathbf{T} \times \mathbf{R} = \mathbf{E}$$

Essential Matrix and Fundamental Matrix

- The defining equation for essential matrix is

$$\mathbf{x}'_o \mathbf{E} \mathbf{x}_0 = 0$$

  with $\mathbf{x}_0 = \mathbf{K}^{-1}\mathbf{x}$ and $\mathbf{x}'_0 = \mathbf{K}'^{-1}\mathbf{x}'$

- Therefore

$$\mathbf{x}'^T \mathbf{K}'^{-T} \mathbf{E} \mathbf{K}^{-1} \mathbf{x} = 0$$

- Since $\mathbf{x}'\mathbf{F}\mathbf{x} = 0$ then

$$\mathbf{E} = \mathbf{K}'^T \mathbf{F} \mathbf{K}$$

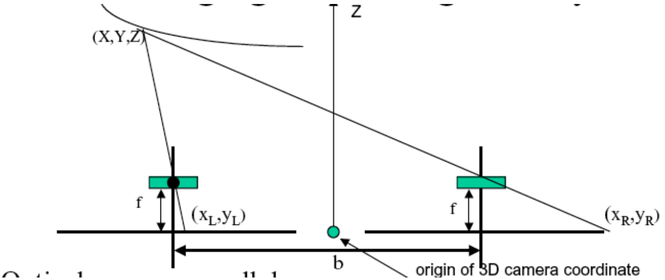
Computing Fundamental Matrix from correspondences

- For any pair of correspondence points you have an equation

$$\mathbf{x}_i' \mathbf{F} \mathbf{x}_i = 0$$

  with $(x, y, 1)$ and $(x', y', 1)$
- solve the linear system for N matches
- How? do it in the board
- How many points do you need?
- There are 8 unknowns, use the **8-point algorithm**
- How do we make this robust?

Stereo with ideal geometry



- Optical axes are parallel and separated by **baseline** b
- Line connecting lens centers is perpendicular to the optical axis, and the x axis is parallel to that line
- 3D coordinate system is a cyclopean system centered between the cameras

[Source: Ramani]

Stereo imaging

- The coordinates of a point are (X, Y, Z) in the cyclopean coordinate system
- The coordinates of the point in the left camera coordinate system are

$$(X_L, Y_L, Z_L) = (X - b/2, Y, Z)$$

  and in the right camera coordinate system are

$$(X_R, Y_R, Z_R) = (X + b/2, Y, Z)$$

- The x image coordinates of the projection in both cameras are

$$x_L = (X + \frac{b}{2})\frac{f}{Z} \qquad x_R = (X - \frac{b}{2})\frac{f}{Z}$$

- Subtracting the second equation from the first, and solving for Z we obtain:

$$Z = \frac{b \cdot f}{x_L - x_R} = \frac{b \cdot f}{d}$$

  with d the disparity

- Subtracting the second equation from the first, and solving for Z we obtain:

$$Z = \frac{b \cdot f}{x_L - x_R} = \frac{b \cdot f}{d}$$

  with d the disparity

- We can also solve for X and Y

$$X = \frac{b(x_L + x_R)}{2(x_L - x_R)} = \frac{b(x_L + x_R)}{2d} \qquad Y = \frac{b \cdot y}{x_L - x_R} = \frac{b \cdot y}{d}$$

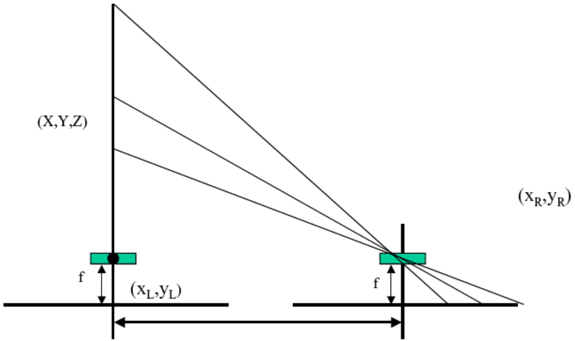
- d is call the disparity and is always negative

Properties of Disparity

- Distance is inversely proportional to absolute value of the disparity
  - Disparity of 0 corresponds to points that are infinitely far away from the cameras
  - Disparity typically in integers (some methods use subpixel accuracy)
  - Thus a disparity measurement in the image just constrains distance to lie in a given range
- Disparity is directly proportional to b
  - the larger b, the further we can accurately range
  - but as b increases, the images decrease in common field of view

[Source: Ramani]

Range vs Disparity



[Source: Ramani]

More on stereo

- A scene point, P, visible in both cameras gives rise to a pair of image points called a **conjugate pair**
- the conjugate of a point in the left (right) image must lie on the same image row (line) in the right (left) image because the two have the same y coordinate
- this line is called the **conjugate line**.
- for our simple image geometry, all conjugate lines are parallel to the x axis

[Source: Ramani]

More practical scenario

- Difficult in practice to
  - have the optical axes parallel
  - have the baseline perpendicular to the optical axes
- we might want to tilt the cameras towards one another to have more overlap in the images
- Calibration problem - finding the transformation between the two cameras

General Stereo Algorithm

- Assume relative orientation of cameras is known
- An image point (x_L, y_L) in the left coordinate system is the image of some point on a ray through the origin of the left camera coordinate system, thus

$$X_L = x_L s \qquad Y_L = y_L s \qquad Z_L = fs$$

- In the right image system, the coordinates of points on this ray are:

$$\begin{aligned} X_R &= (r_{11}x_L + r_{12}y_L + r_{13}f)s + u \\ Y_R &= (r_{21}x_L + r_{22}y_L + r_{23}f)s + v \\ Z_R &= (r_{31}x_L + r_{32}y_L + r_{33}f)s + w \end{aligned}$$

- Why?
- These points project on the right camera onto

$$x_R = f\frac{X_R}{Z_R} \qquad y_R = f\frac{Y_R}{Z_R}$$

[Source: Ramani]

General stereo

- In the right image system, the coordinates of points on this ray are:

$$\begin{aligned} X_R &= (r_{11}x_L + r_{12}y_L + r_{13}f)s + u = as + u \\ Y_R &= (r_{21}x_L + r_{22}y_L + r_{23}f)s + v = bs + v \\ Z_R &= (r_{31}x_L + r_{32}y_L + r_{33}f)s + w = cs + w \end{aligned}$$

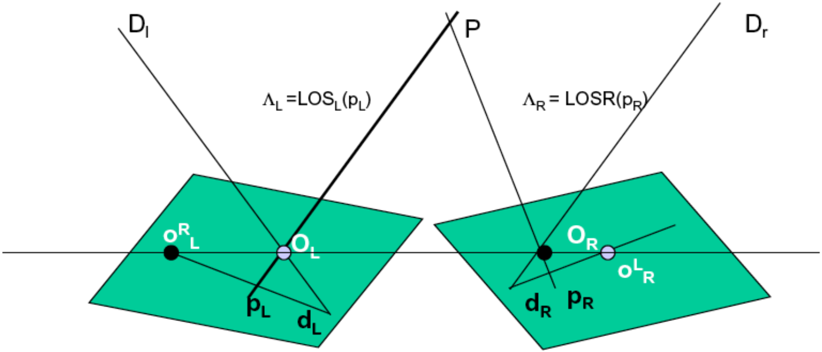
- Then

$$\frac{x_R}{f} = \frac{as + u}{cs + w} \qquad \frac{y_R}{f} = \frac{bs + u}{cs + w}$$

- This is a straight line connecting the point
  - $(u/w, v/w)$ which occurs for $s = 0$ and is the image of the left camera center in the right camera coordinate system to
  - $(a/c, b/c)$ which occurs as s approaches infinity,the vanishing point for the ray

[Source: Ramani]

General Stereo Geometry



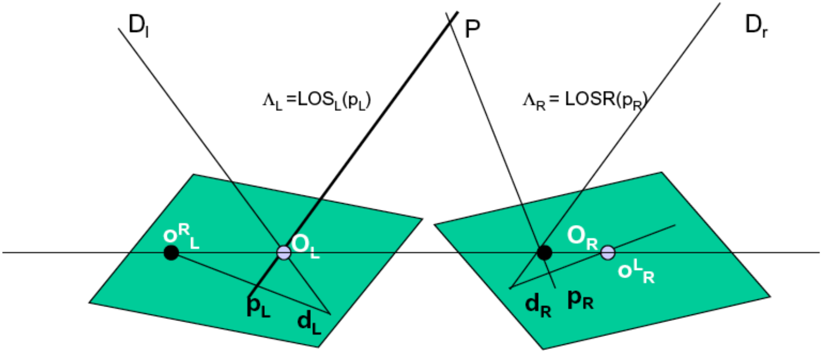
[Source: Ramani]

General Stereo

- Point P lies somewhere on the ray (line) from p_L through O_L
- from the left image alone, we do not know where on this ray P lies
- perspective projection of a line is a line
- The first point on the line that might correspond to P is O_L, any point closer to the left image than O_L could not be seen
- the perspective projection of O_L in the right camera is the point o_R^L
- The last point on line that might correspond to P is the point infinitely far away along the ray
- its image is the vanishing point of the ray in the right camera, d_R
- any other possible location for P will project to a point in R on the line joining o_R^L to d_R.

[Source: Ramani]

General Stereo Geometry

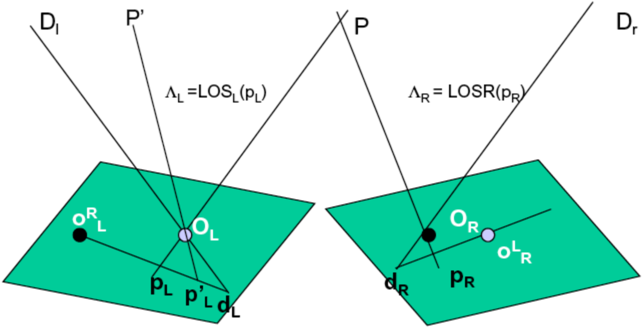


[Source: Ramani]

General Stereo

- Given any point, p_L, in the left image of a stereo pair, its conjugate point must appear on a line in the right image
- all of the conjugate lines for all of the points in the left image must pass through a common point in the right image
  - this is image of the left lens center in the right image
  - this point lies on the line of sight for every point in the left image
  - the conjugate lines must all contain (i.e., pass through) the image of this point
  - This point is called an **epipole**.
- The conjugate line for p_L must also pass through the vanishing point in the right image for the line of sight through p_L

General Stereo Geometry



[Source: Ramani]

More on geometry

- The points O_L, p_L, and o_L^R are three noncollinear points, so they form a plane,
- The intersection of this plane with the right image plane is the conjugate line of p_L, and this would be the image of any line on this plane
- Let p_L' be some other point on the line joining p_L and o_L^R,
- the line of sight through p_L' to P' lies on the plane since two points on that line (p_L and o_L^R lie on the plane
- Thus, the conjugate line for p_L' must be the same line as the conjugate line for p_L, or **for any other point** on the line containing p_L and o_L^R
- We use this epipolar lines for matching

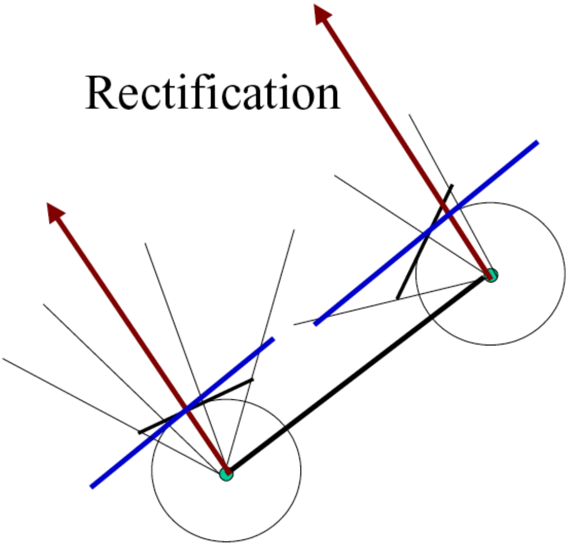
[Source: Ramani]

Rectification

- transforming a stereo pair taken under general conditions into the ideal configuration
- Involves a rotation of one image so that the optical axes of the two image coordinate systems are parallel
- Simplifies computational structure of stereo matching algorithm
- But requires interpolation to create rotated image and can create a large rectified image if the rotation angles are large.

[Source: Ramani]

Rectification



[Source: Ramani]

Stereo correspondence problem

- Given a point,p, in the left image, find its conjugate point in the right image
- What constraints simplify this problem?
  - Epipolar constraint - need only search for the conjugate point on the epipolar line
  - Disparity sign constraint - need only search the epipolar line to the right of the vanishing point in the right image of the ray through p in the left coordinate system
  - Continuity constraint - if we are looking at a continuous surface, images of points along a given epipolar line will be ordered the same way
  - Disparity gradient constraint - disparity changes slowly over most of the image (Exceptions occur at and near occluding boundaries)

[Source: Ramani]

Why is the correspondence problem hard

- Foreshortening effects
- A square match window in one image will be distorted in the other if disparity is not constant (complicates correlation)
- Variations in intensity between images due to: noise, specularities, shape-from-shading differences
- Occlusion: points visible in one image and not the other
- Coincidence of edge and epipolar line orientation

Applications of Stereo

- Photogrammetry: Creation of digital elevation models from high resolution aerial imagery
- Visual navigation: Obstacle detection
- Creating models for graphics applications: For objects difficult to design using CAD systems

[Source: Ramani]