Announcement

Homework 4 has been posted online and in Blackboard.

Due: 2:20pm EST, Monday, April 3rd

Announcement

We will have Quiz #5 in class, March. 29

Open book and open notes

We will have 5 more quizzes. No late quizzes will be accepted without preapproval.

Reminder: Final Project Presentation Schedule

Presentation days:
April 19 and 24

A team project needs a combined presentation.

Each individual presentation has 12 + 3 minutes Q&A
Each team presentation has 18 + 3 minutes Q&A

Send me an email (tongy@cse.sc.edu) that includes:

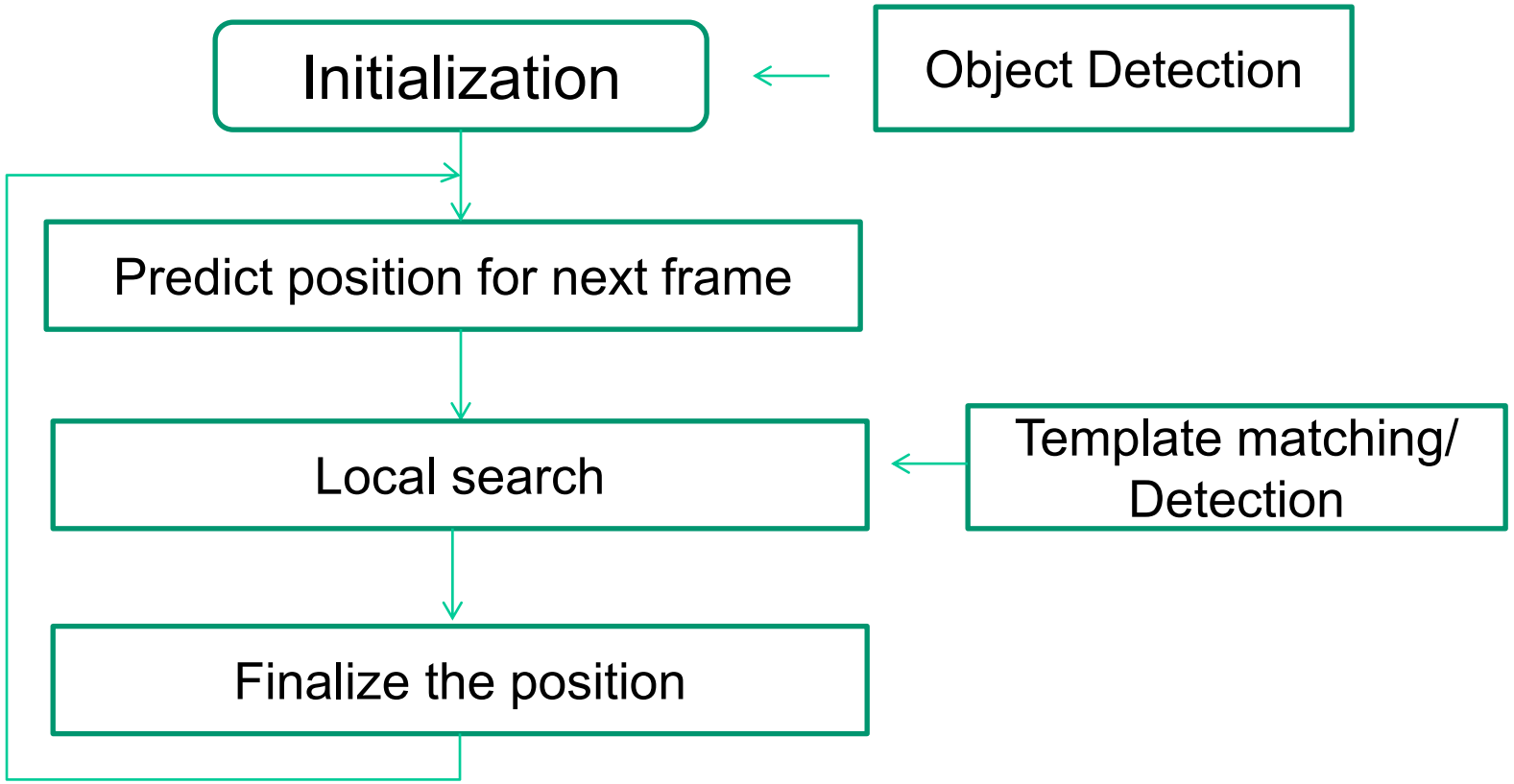
- Your name and the dates starting from the most preferred
- Earlier email has higher priority in choosing the date

Today

Early vision on multiple images

- Epipolar geometry

General Strategy for Object Tracking



Tracking – How to Obtain Measurements

- **Tracking by detection**
  - Specific type of objects
  - a very reliable detector (e.g., faces, pedestrian)
  - Sparse targets with distinctive properties
- **Tracking by template matching**

Without prior knowledge of the target – no detectors

- Moving foreground objects
- Tracking a target among multiple objects with similar appearance

Tracking by Template Matching

Template: a region (mostly a rectangular region) with known appearance

Template matching: searching a region that minimizes an SSD error (Sum of Squared Differences)

$$SSD = \sum_{i,j} (I_t\,(x+i, y+j) - T(i,j))^2$$

Tracking by Template Matching

Potential issues:

- **Template updating**
  - When and how
- **Template matching metrics**
  - SSD vs. histogram-based matching

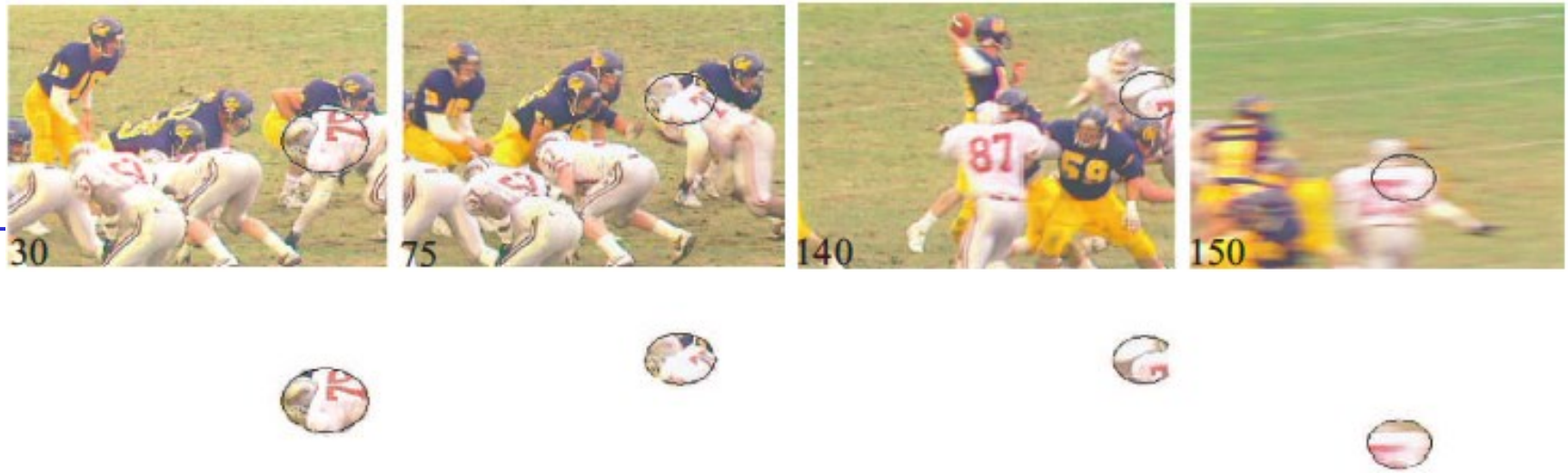


FIGURE 11.5: Four frames from a sequence depicting football players, with superimposed domains. The object to be tracked is the blob on top of player 78 (at the center right in frame 30). We have masked off these blobs (**below**) to emphasize just how strongly the pixels move around in the domain. Notice the motion blur in the final frame. These blobs can be matched to one another, and this is done by comparing histograms (in this case, color histograms), which are less affected by deformation than individual pixel values. *This figure was originally published as Figure 1 of "Kernel-Based Object Tracking" by D. Comaniciu, V. Ramesh, and P. Meer, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003,* © *IEEE 2003.*

Forsyth and Ponce, "Computer Vision – A Modern Approach 2e"

Reading Assignments

- **Chapter 11 (Tracking) of Forsyth & Ponce**

- **CVonline: Motion and Time Sequence Analysis**

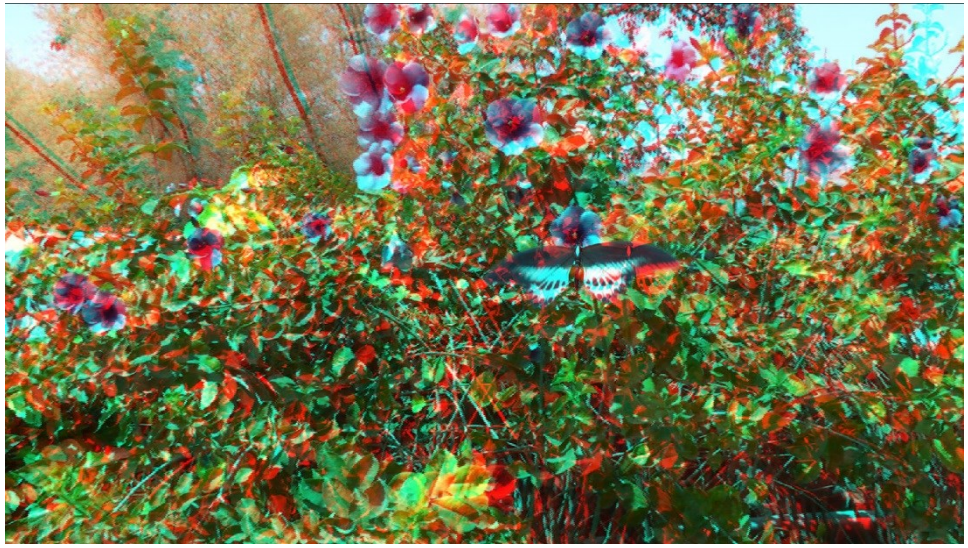
https://sites.google.com/site/cvonlinewiki/home/motion-and-time-sequence-analysis-related-concepts

Stereopsis

Stereo camera
California Museum of Photography

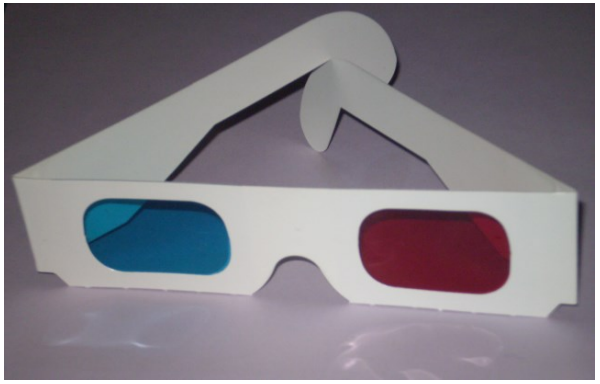
“Mark Twain at Pool Table", no date, UCR Museum of Photography

More Stereo Images



http://realvision.ae/blog/2009/09/stereoscopic-3d-in-classrooms/

https://www.dpreview.com/



By Snaily - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=1612671

Recover 3D from Images

Goal: recover the 3D property of an object including

- 3D coordinates
- Depth
- 3D shape

Two typical ways:

- **Reconstruct from image cues**
  - Single image
  - Multiple images

Human vision
Kinect

- **Obtain depth information directly using special devices**
  - Laser scanner
  - Radar
  - Ultrasound

Shape from X -- Image Cues

Shading

http://theindependentartschool.com/art-blog/2011/02/07/drawing-and-shading-a-sphere/

Texture

Visual cliff
http://brisray.com/optill/ovision1.htm

Focus/defocus

Motion

http://www.cs.ucla.edu/~soatto/research/projects.html

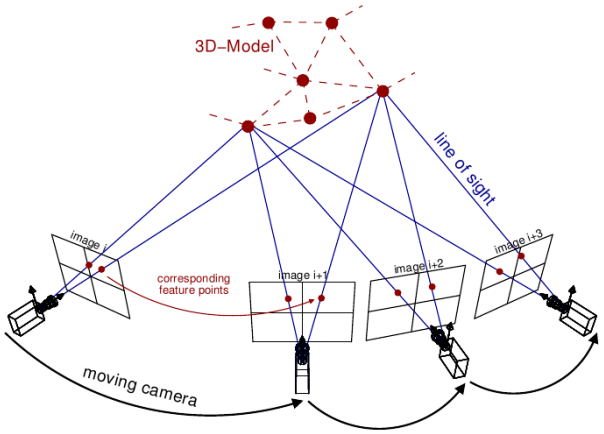
Many others

- Geometry
- Shadow
- Specularity
- etc.

Shape from Multiple Images: Stereo

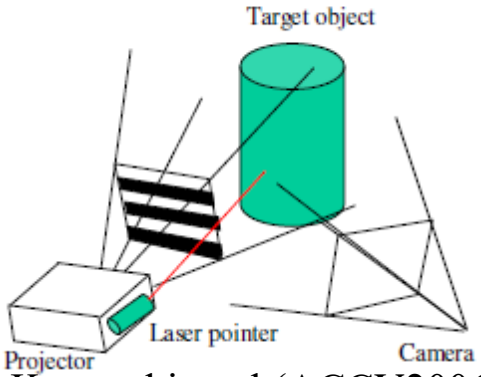
Reconstruct a 3D scene/object from images observed from multiple view points

- **Passive stereo**
  - Multiple (at least two) cameras



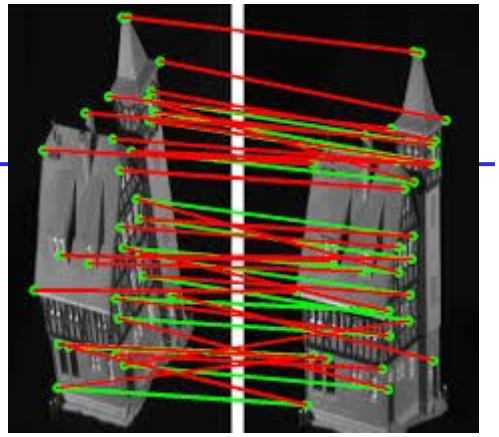
http://www.tnt.uni-hannover.de/project/motionestimation/

- **Active stereo**
  - Structured light
  - Usually consist of a camera and a projector
    - Kinect



Kawasaki et al (ACCV2006)

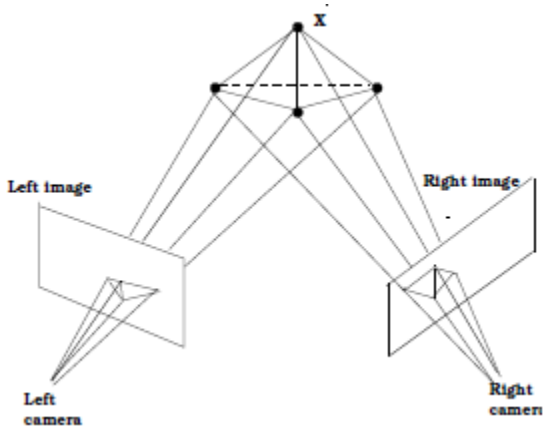
Passive Stereo



pages.cs.wisc.edu

General strategy:

- **Camera calibration**
- **Correspondence**
  - Detect a set of tokens (points, lines, and conic curves)
  - Build correspondence for the tokens in multiple (at least two) images
- **3D reconstruction from matched 2D points**
  - Image rectification
  - 3D triangulation from rectified images

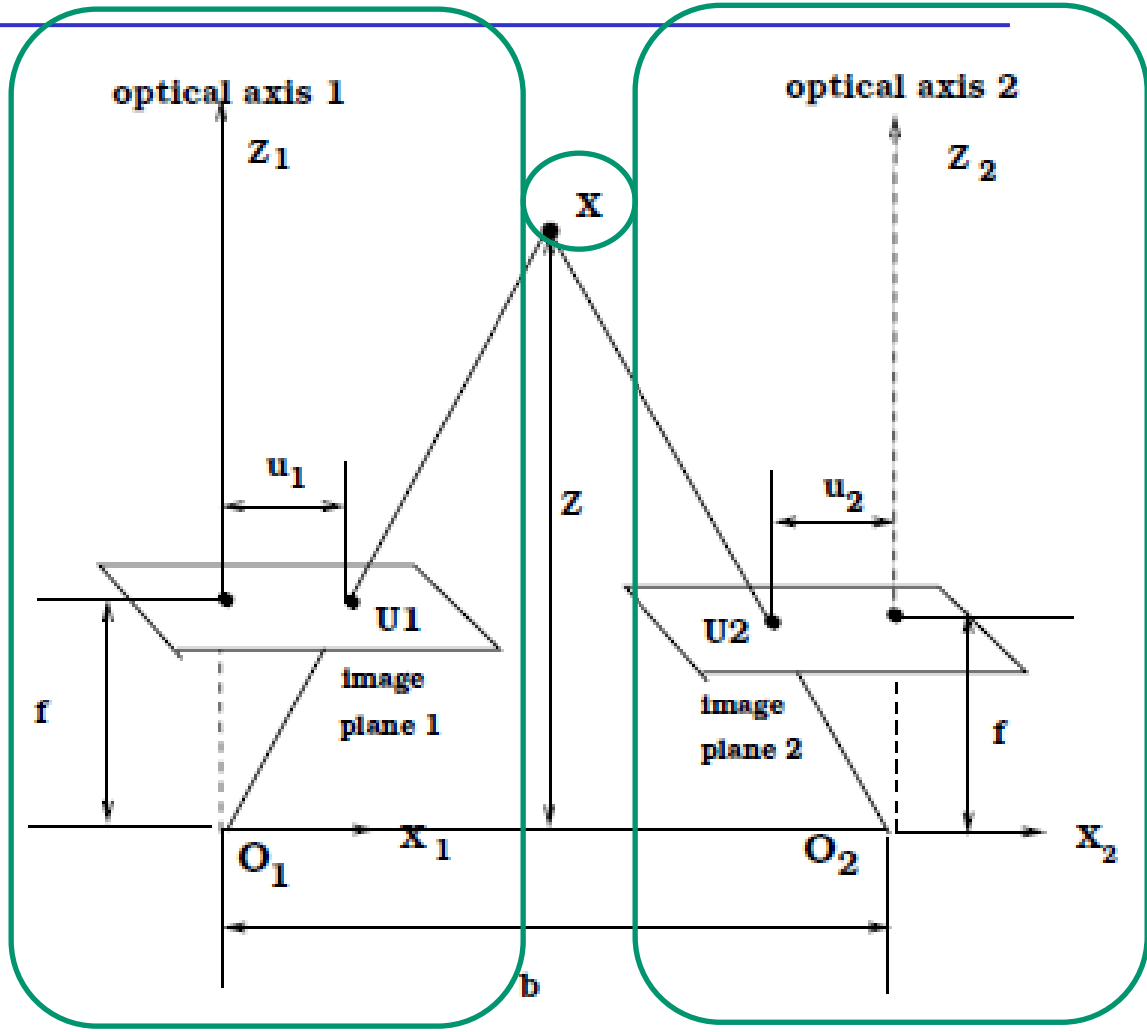


courtesy of Dr. Qiang Ji

A Simple Case: 3D Triangulation – Rectified Geometry

Rectified geometry:

- two image planes are coplanar and parallel to the line O_1O_2



courtesy of Dr. Qiang Ji

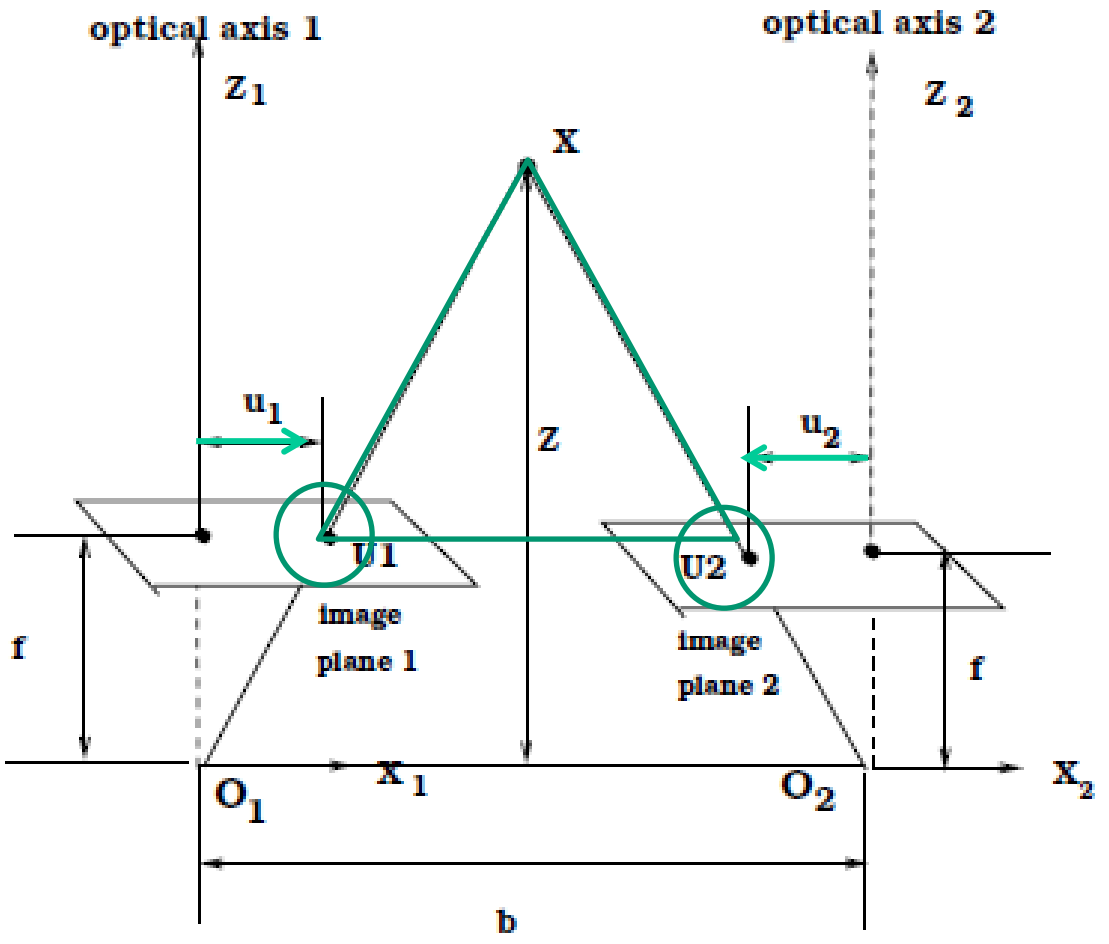
A Simple Case: 3D Triangulation – Rectified Geometry

Rectified geometry:

Hence, triangle XU_1U_2 is similar to triangle XO_1O_2

$$\frac{Z-f}{Z} = \frac{b-u_1+u_2}{b}$$

$$Z = \frac{fb}{u_1-u_2}$$



courtesy of Dr. Qiang Ji

3D Triangulation – Rectified Geometry

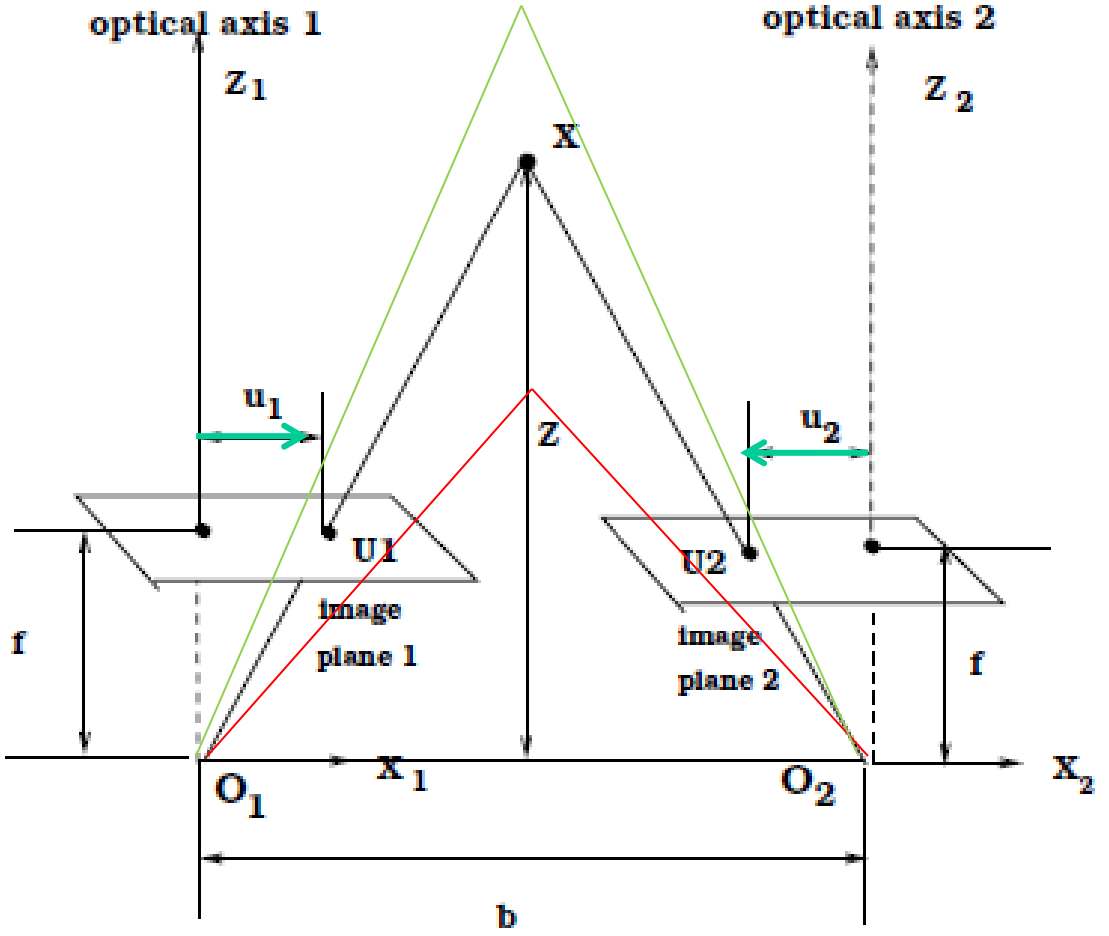
$$Z = \frac{fb}{u_1 - u_2}$$

b: base distance

$u_1 - u_2$: disparity

$$u_1 - u_2 = \frac{fb}{Z}$$

Disparity is inversely proportional to the depth.



Example of Disparity Map

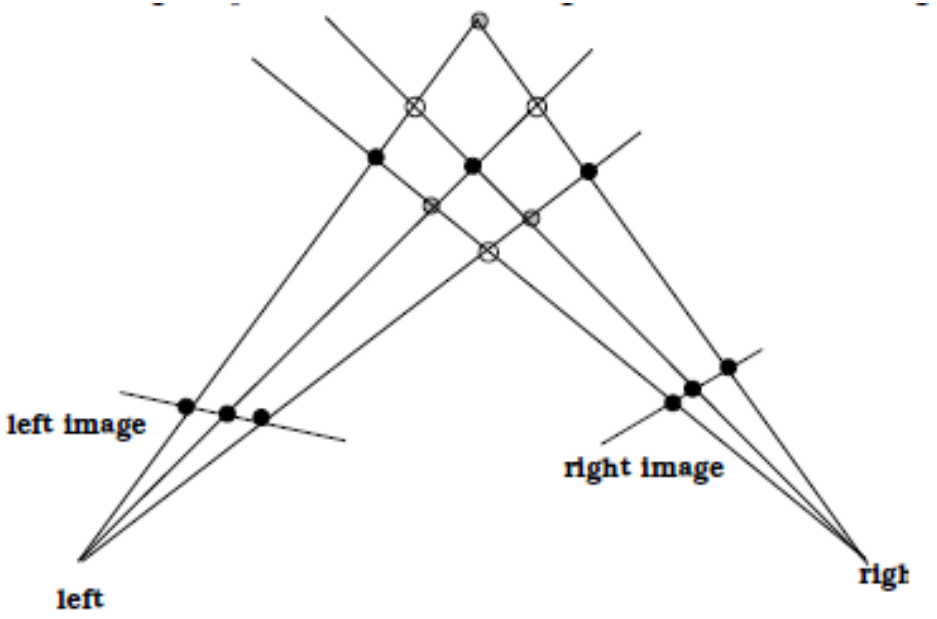
Groundtruth disparity map

Head and lamp from University of Tsukuba Stereo Dataset

http://cvlab-home.blogspot.com/2012/05/h2fecha-2581457116665894170-displaynone.html

Establish Correspondence

Problem: there are multiple images taken for the same object. We need to determine which point in one image is corresponding to the point in the others.

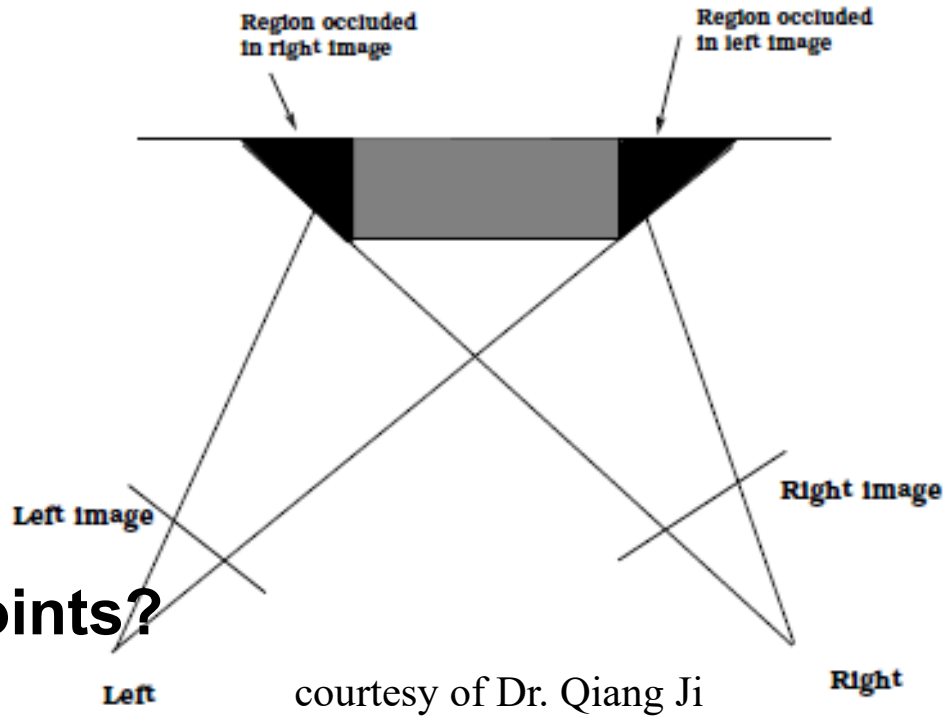


courtesy of Dr. Qiang Ji

Establish Correspondence

Challenges:

- Information associated with each image is not sufficient
- Image points in one image may not have corresponding point in another image due to
  - Field of view doesn't overl
  - Occlusion
  - Miss detection



courtesy of Dr. Qiang Ji

How to find corresponding points?

Establish Pointwise Correspondence using Epipolar Constraint

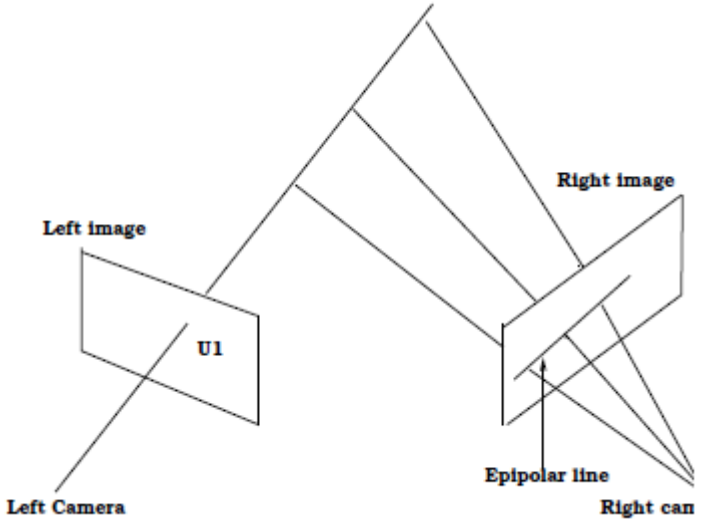
Epipolar constraint: For a 2D image point U_l in the left image, its corresponding point in the right image must lie on the conjugate epipolar line.

Epiploar line is the projection of the 3D line, which passes through U_l and the left camera center, on the right image.

Advantages:

• Simplify the searching for the corresponding points from 2D to 1D

• improve the efficiency significantly

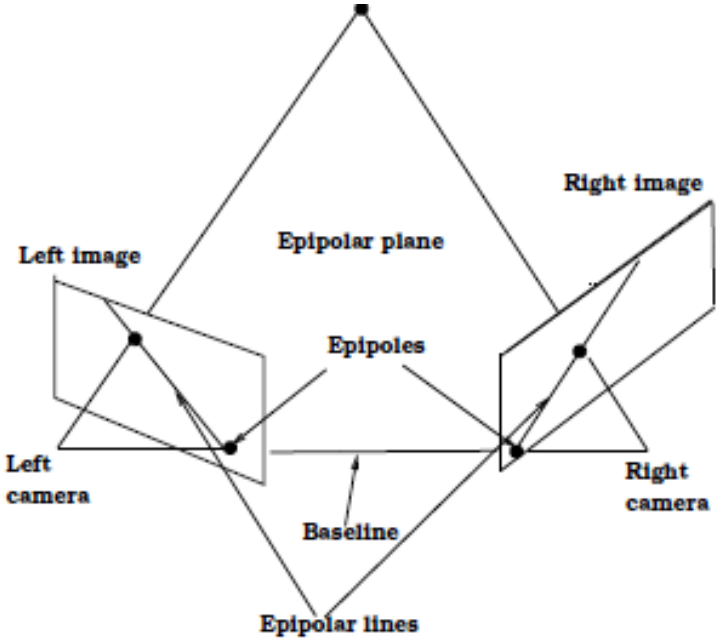


courtesy of Dr. Qiang Ji

Epipolar Geometry

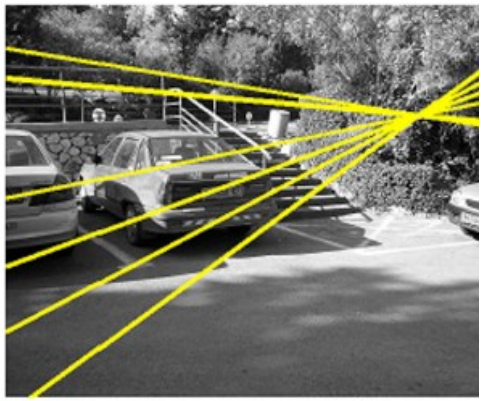
All epipolar lines go through an epipole on one image.

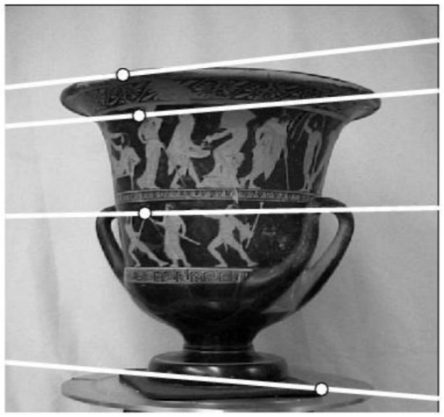
The epipole on the left (right) image is the projection of the optical center of the right (left) image on the left (right) image plane.



courtesy of Dr. Qiang Ji

Examples





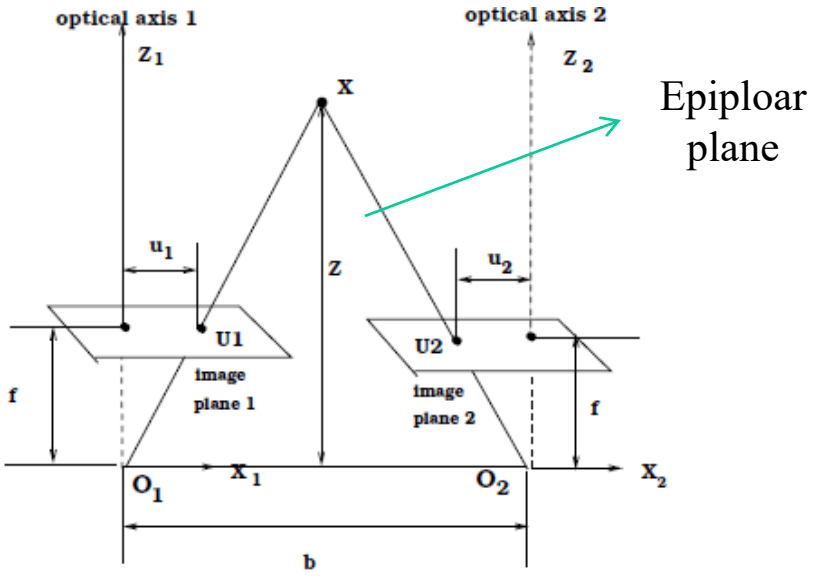
Epipolar Geometry

In rectified images,

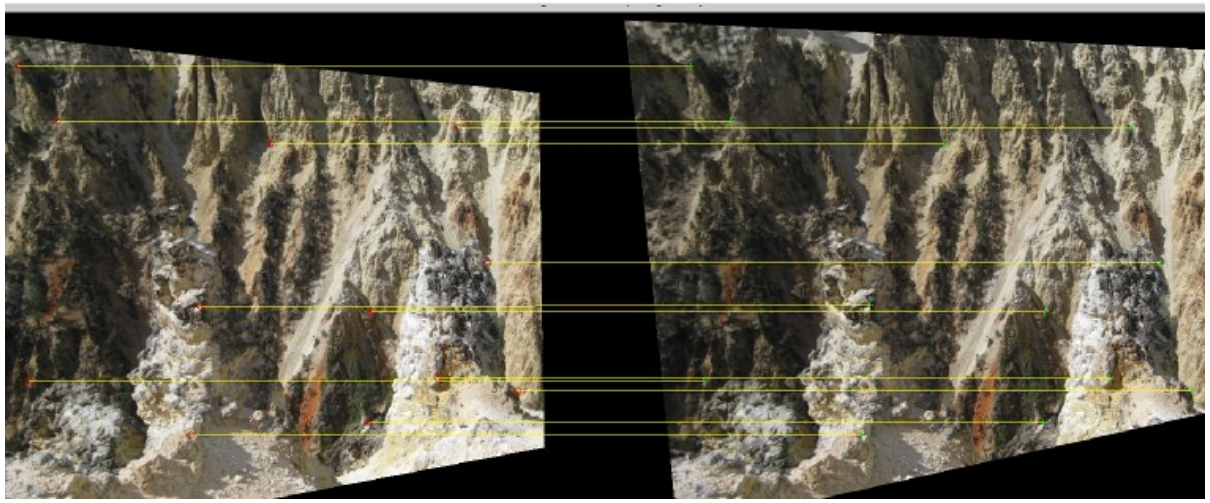
- the epipolar lines are parallel; ⇒ The epipole is at infinity
- conjugate epipolar lines are collinear and parallel to the base line.

Why?

- The conjugate epipolar lines are the intersections of the epipolar plane with the two image planes
- The two image planes are coplanar and parallel to the base line



Examples



Epipolar Geometry

P: a 3D point

$\boldsymbol{P_l}$**:** the coordinate of P on the left camera coordinate system

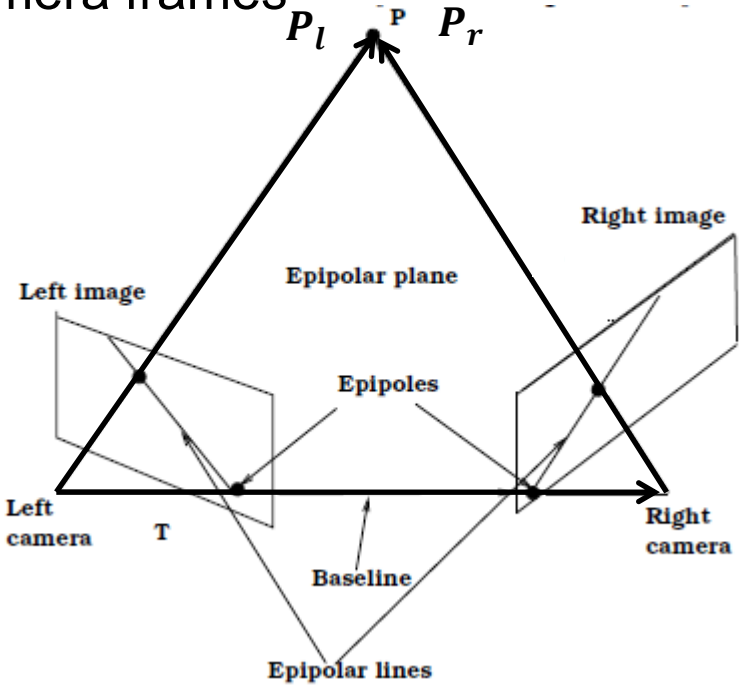
$\boldsymbol{P_r}$**:** the coordinate of P on the right camera coordinate system

T: the relative translation between two camera frames

$$\boldsymbol{O_r}^{\boldsymbol{l}} = \boldsymbol{O_r}^{\boldsymbol{r}} + \boldsymbol{T}$$

The right camera center measured in the left frame

The right camera center measured in the right frame



Epipolar Geometry

P: a 3D point

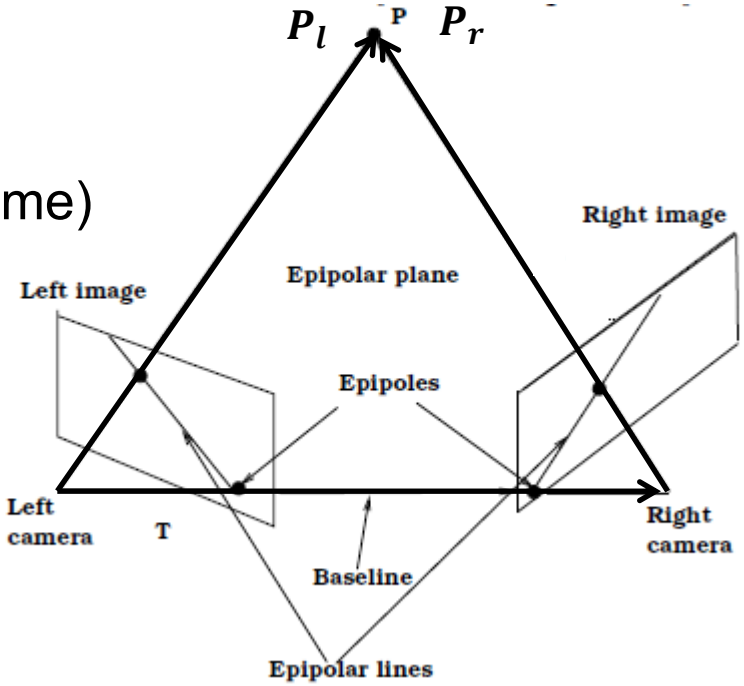
$\boldsymbol{P_l}$**:** the coordinate of P on the left camera coordinate system

$\boldsymbol{P_r}$**:** the coordinate of P on the right camera coordinate system

R: the relative rotation between two camera frames (from right frame to left frame)

$$\boldsymbol{P_l} = \mathbf{R}\boldsymbol{P_r} + \boldsymbol{T}$$

$$\boldsymbol{P_r} = \mathbf{R}^t(\boldsymbol{P_l} - \boldsymbol{T})$$



Epipolar Geometry

$\boldsymbol{P_l}$, $\boldsymbol{P_l - T}$ **and T** are coplanar → in the same epipolar plane

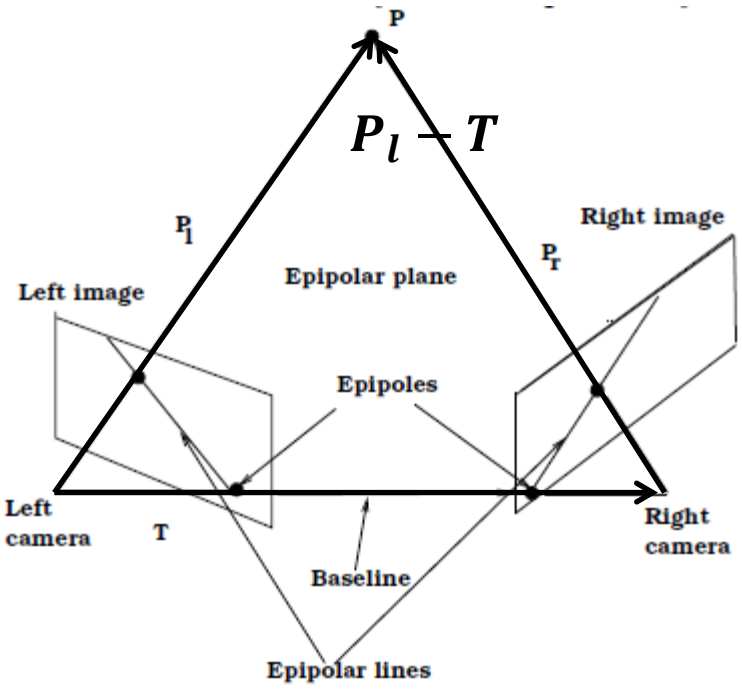
$$(\boldsymbol{T} \times \boldsymbol{P_l})^{\boldsymbol{t}} (\boldsymbol{P_l - T}) = \mathbf{0}$$

Normal of the plane

$$(\boldsymbol{T} \times \boldsymbol{P_l})^{\boldsymbol{t}} \mathbf{R} \boldsymbol{P_r} = \mathbf{0}$$

Let $\mathbf{S} = \begin{bmatrix} 0 & -T_z & T_y \\ T_z & 0 & -T_x \\ -T_y & T_x & 0 \end{bmatrix}$

$$\boldsymbol{T} \times \boldsymbol{P_l} = \mathbf{S} \boldsymbol{P_l}$$



Epipolar Geometry

$$(\mathbf{S}\boldsymbol{P}_l)^t \mathbf{R}\boldsymbol{P}_r = \mathbf{0} \leftrightarrow \boldsymbol{P}_l^t\, \mathbf{S}^t \mathbf{R}\boldsymbol{P}_r = 0$$

Let $\mathbf{E} = \mathbf{S}^t\mathbf{R}$

$$\boldsymbol{P}_l^t\, \mathbf{E}\boldsymbol{P}_r = \mathbf{0}$$

E is called the **essential matrix**

Since $\text{rank}(\mathbf{S}) = 2, \text{rank}(\mathbf{R}) = 3$ ⇒ $\text{rank}(\mathbf{E}) = 2$

The essential matrix

- establishes a link between the epipolar constraint and the relative position of the two camera systems
- has 5 degree of freedom

Fundamental Matrix

$\boldsymbol{U_l}$: the homogeneous coordinate of the projection of point P on the left image plane

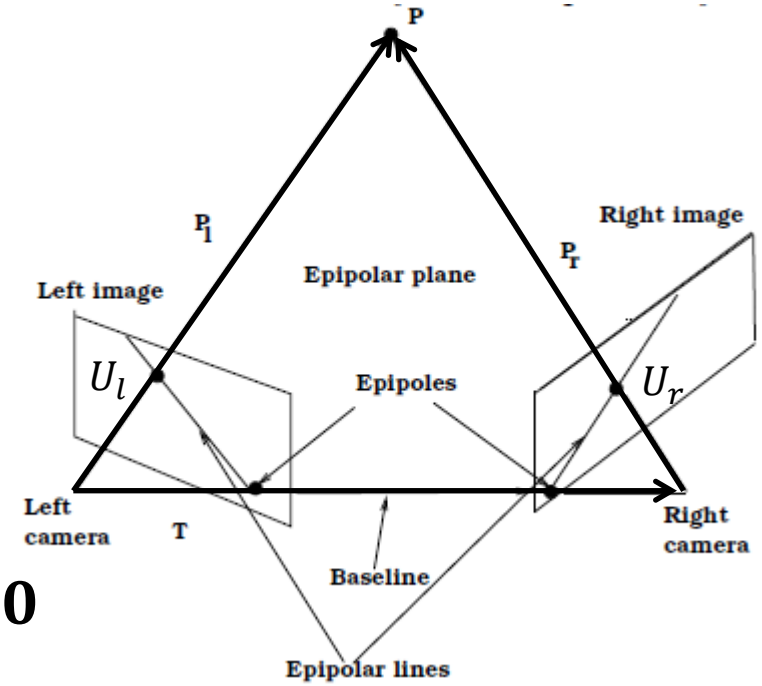
$\boldsymbol{U_r}$: the homogeneous coordinate of the projection of point P on the right image plane

$$\lambda_l U_l = \mathbf{W}_l P_l \qquad \lambda_r U_r = \mathbf{W}_r P_r$$

Intrinsic parameters for left and right cameras

Substitute to $P_l^t \mathbf{E} P_r = 0$

$$U_l^t \mathbf{W}_l^{-t} \mathbf{E} \mathbf{W}_r^{-1} U_r = 0$$



Fundamental Matrix

Fundamental matrix

$$\mathbf{F} = \mathbf{W_l}^{-\mathbf{t}}\mathbf{E}\mathbf{W_r}^{-\mathbf{1}}$$

$$U_l^t \mathbf{F} U_r = \mathbf{0}$$

A pair of corresponding image points are related by **F**!

A fundamental matrix encodes

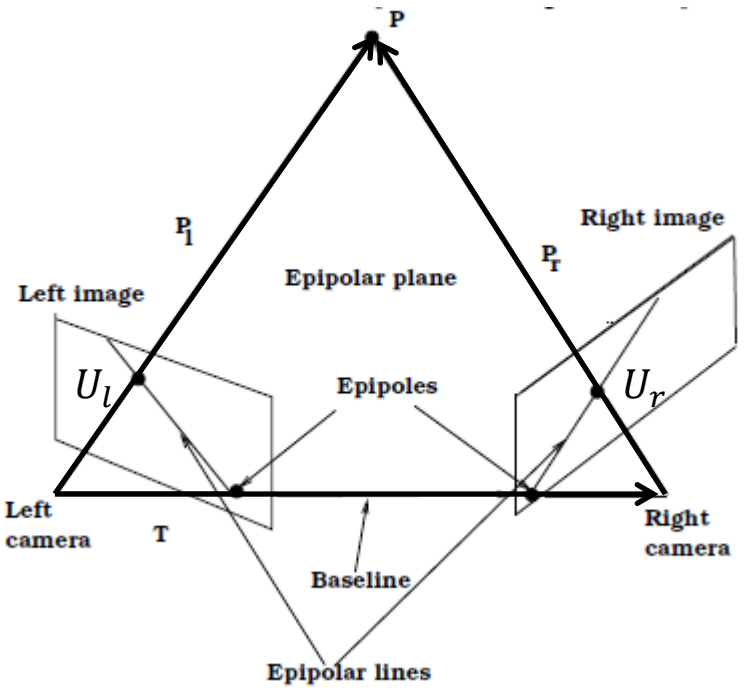
- the relative orientation and translation

between two camera systems

- the intrinsic parameters of the cameras

What is the rank of **F**?

rank(**F**)=2 because of **E**



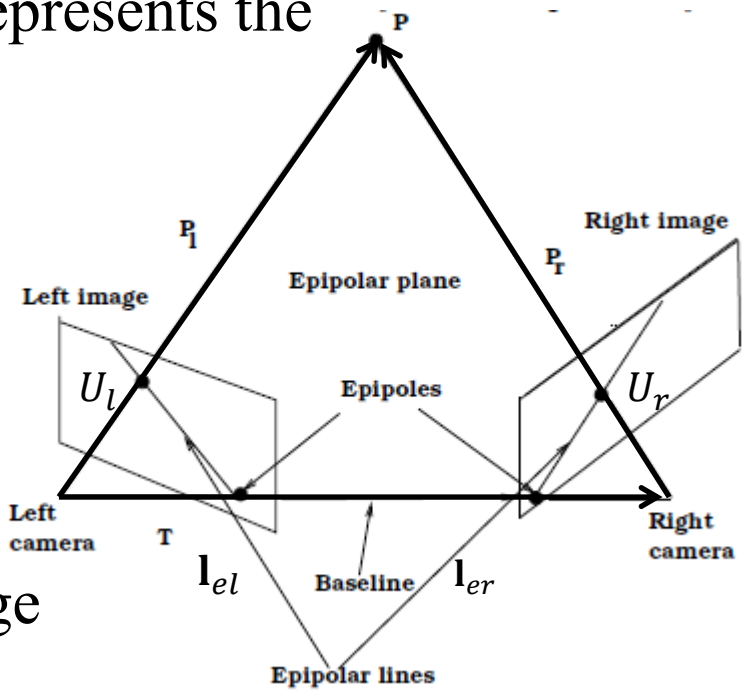
Information from Fundamental Matrix

$$U_l^t \mathbf{F} U_r = \mathbf{0}$$

Epipolar constraint: $U_l^t(\mathbf{F}U_r) = \mathbf{0}$ and $(\mathbf{F}^t U_l)^t \; U_r = \mathbf{0}$

Recall line function: $\mathbf{l} \cdot \mathbf{p} = \mathbf{0}$, where $\mathbf{l}$ represents the line parameters

$\mathbf{l}_{el} = \mathbf{F}U_r$



U_l is on a line determined by $\mathbf{l}_{el}$

- Epipolar line on the left image
- Corresponding to U_r on the right image

Information from Fundamental Matrix

$$U_l^t \mathbf{F} U_r = 0$$

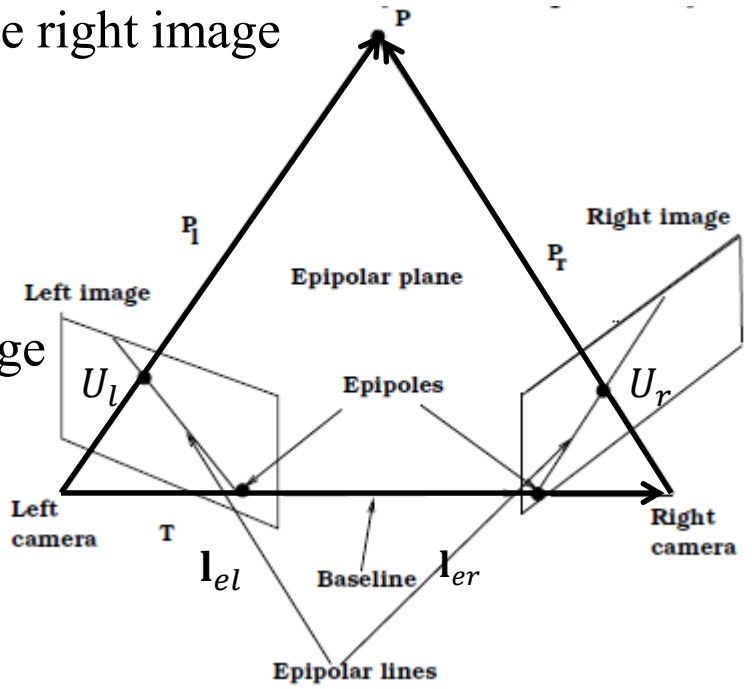
Epipolar constraint: $U_l^t(\mathbf{F}U_r) = 0$ and $(\mathbf{F}^t U_l)^t \; U_r = 0$

$\mathbf{l}_{el} = \mathbf{F}U_r$ Epipolar line on the left image corresponding to U_r on the right image

Similarly

$\mathbf{l}_{er} = \mathbf{F}^t U_l$ Epipolar line on the right image corresponding to U_l on the left image

Given F and a point on one image, we can estimate the corresponding epipolar line



Information from Fundamental Matrix

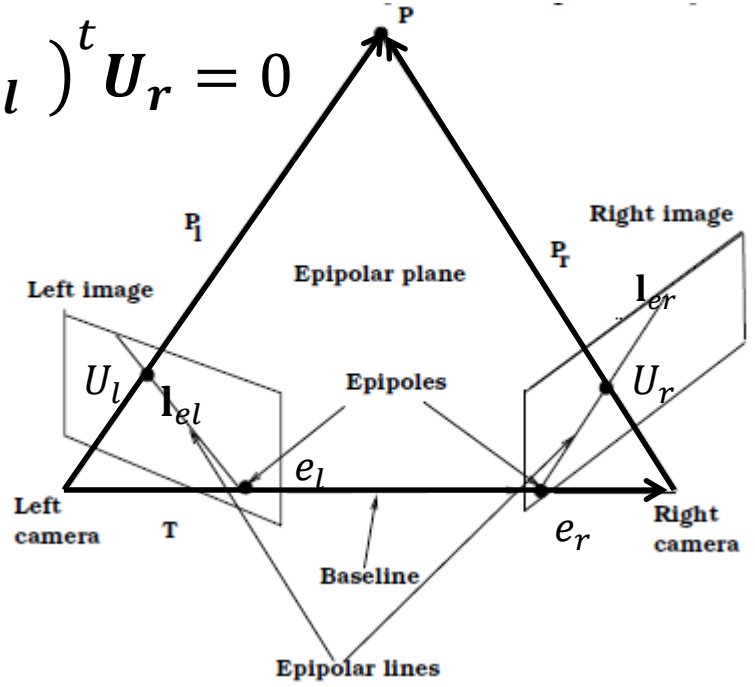
$$U_l^t \, \mathbf{F} U_r = \mathbf{0}$$

Since all epipolar line pass through the epipole,

$$e_r^t \mathbf{l}_{er} = \mathbf{0} \rightarrow \; e_r^t \mathbf{F}^t U_l = 0 \qquad \rightarrow \left(\mathbf{F} e_r\right)^t U_l = 0$$

$$e_l^t \mathbf{l}_{el} = \mathbf{0} \rightarrow \; e_l^t \mathbf{F} U_r = 0 \qquad \rightarrow \left(\mathbf{F}^t e_l\right)^t U_r = 0$$

$$\mathbf{F} e_r = 0$$

$$\mathbf{F}^t e_l = 0$$

Given F, the epipole of right (left) image is the null eigenvector of **F** ($\mathbf{F}^t$)