
DTU

Lazaros Nalpantidis

Stereo Vision

some slides borrowed or adapted from:

- *Noah Snavely*
- *Aaron Bobick*
- *Antonio Torralba*

Outline

- What is Stereo Vision?

- Stereo/Epipolar Geometry
- Rectified Stereo Case
- Depth from Stereo Matches
- Correspondence Problem
  - **Dense** vs Sparse Correspondence
  - **Local** vs Global Correspondence
  - (Dis)-Similarity Measures

- Summary

Outline

- What is Stereo Vision?

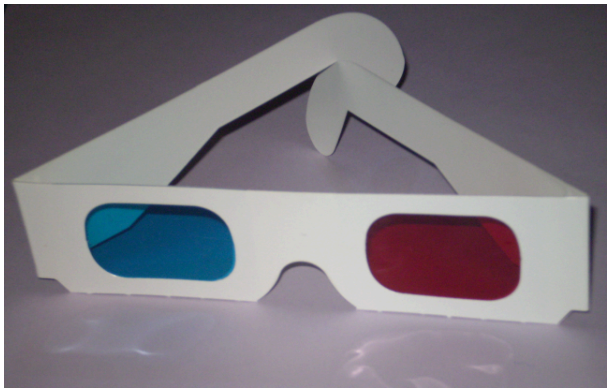
- Stereo/Epipolar Geometry
- Rectified Stereo Case
- Depth from Stereo Matches
- Correspondence Problem
  - **Dense** vs Sparse Correspondence
  - **Local** vs Global Correspondence
  - (Dis)-Similarity Measures

- Summary

What is Stereo Vision?

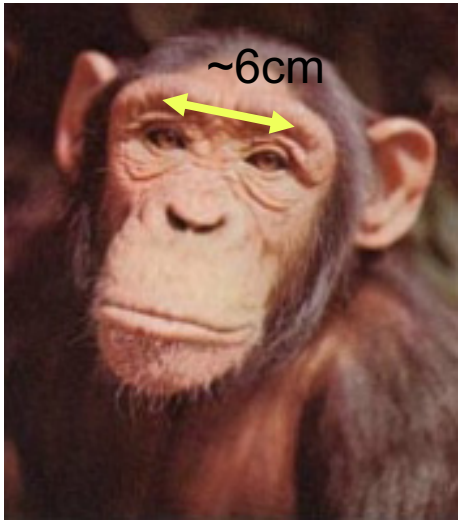
What is Stereo Vision?

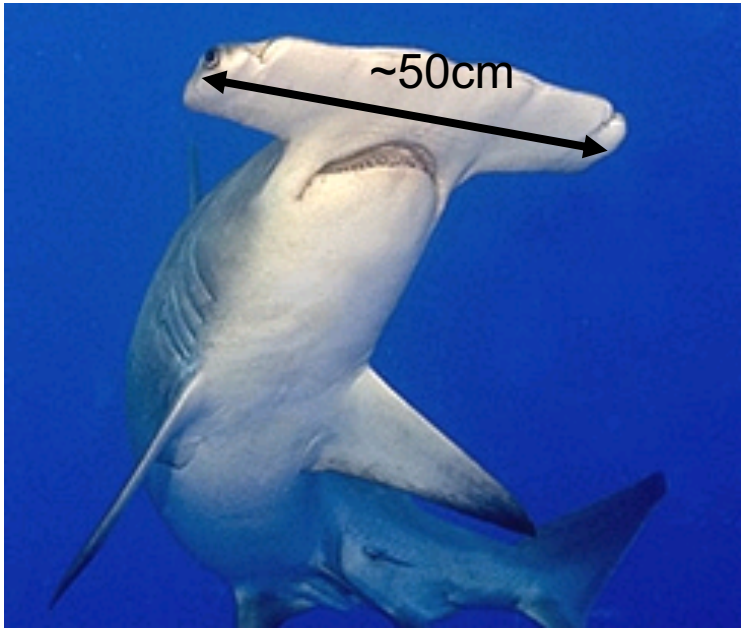
- 3D cinema
- 3D television
- …



Adapted from W. T. Freeman, P. Isola, A. Torralba

What is Stereo Vision?



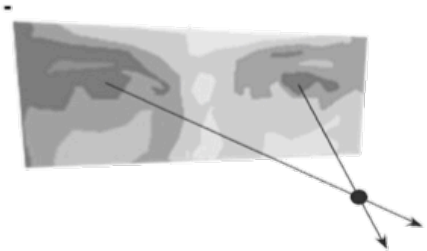


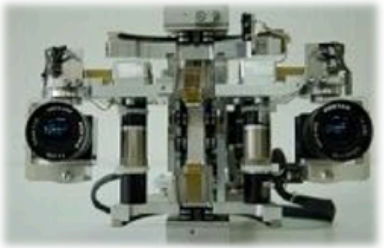
Adapted from W. T. Freeman, P. Isola, A. Torralba

What is Stereo Vision?

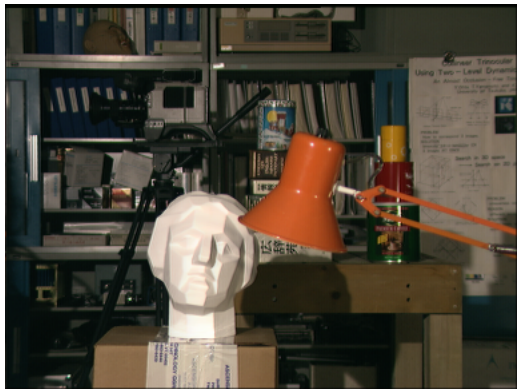
”… the mind perceives an object of three-dimensions by means of the two dissimilar pictures projected by it on the two retinae…”

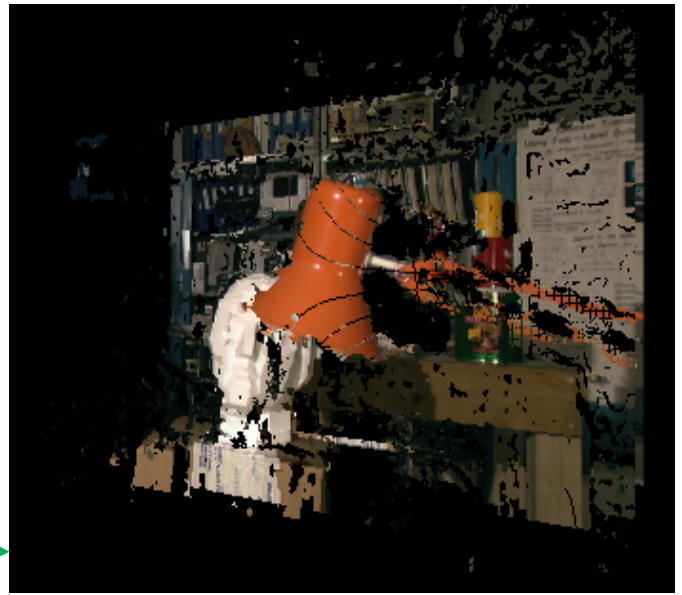
Sir Charles Wheatstone, 1838





Stereo Vision Computation





Outline

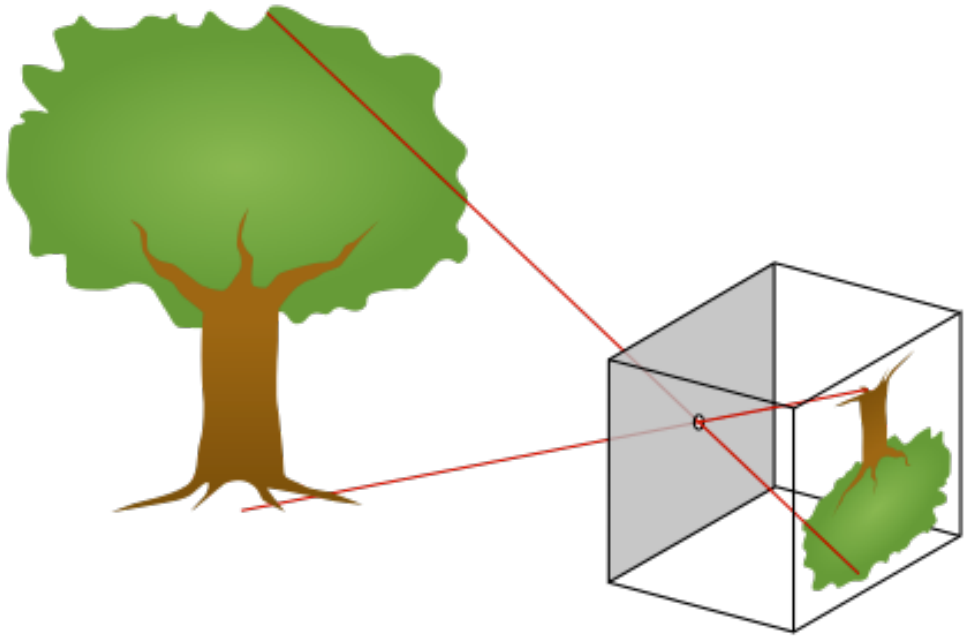
- What is Stereo Vision?

- Stereo/Epipolar Geometry
- Rectified Stereo Case
- Depth from Stereo Matches
- Correspondence Problem
  - **Dense** vs Sparse Correspondence
  - **Local** vs Global Correspondence
  - (Dis)-Similarity Measures

- Summary

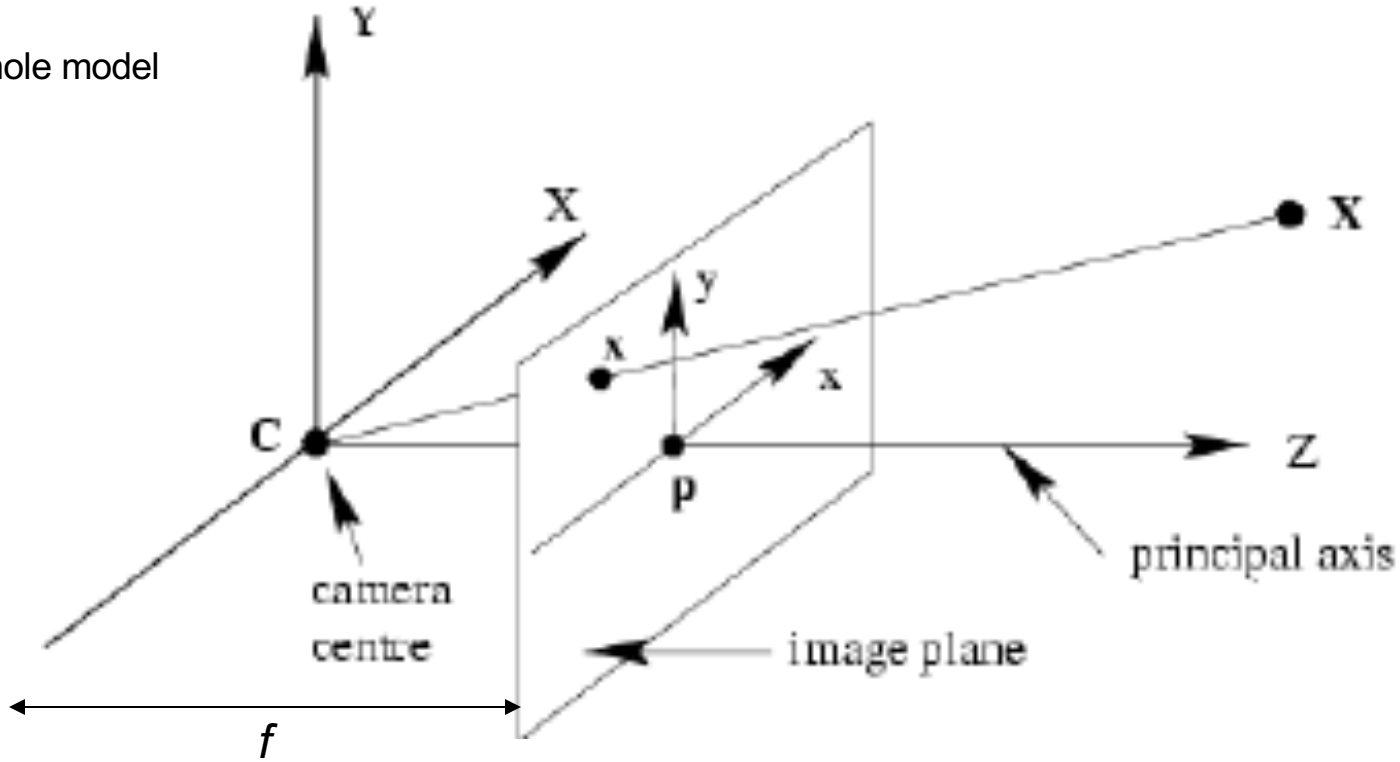
Camera Model

- Pinhole model

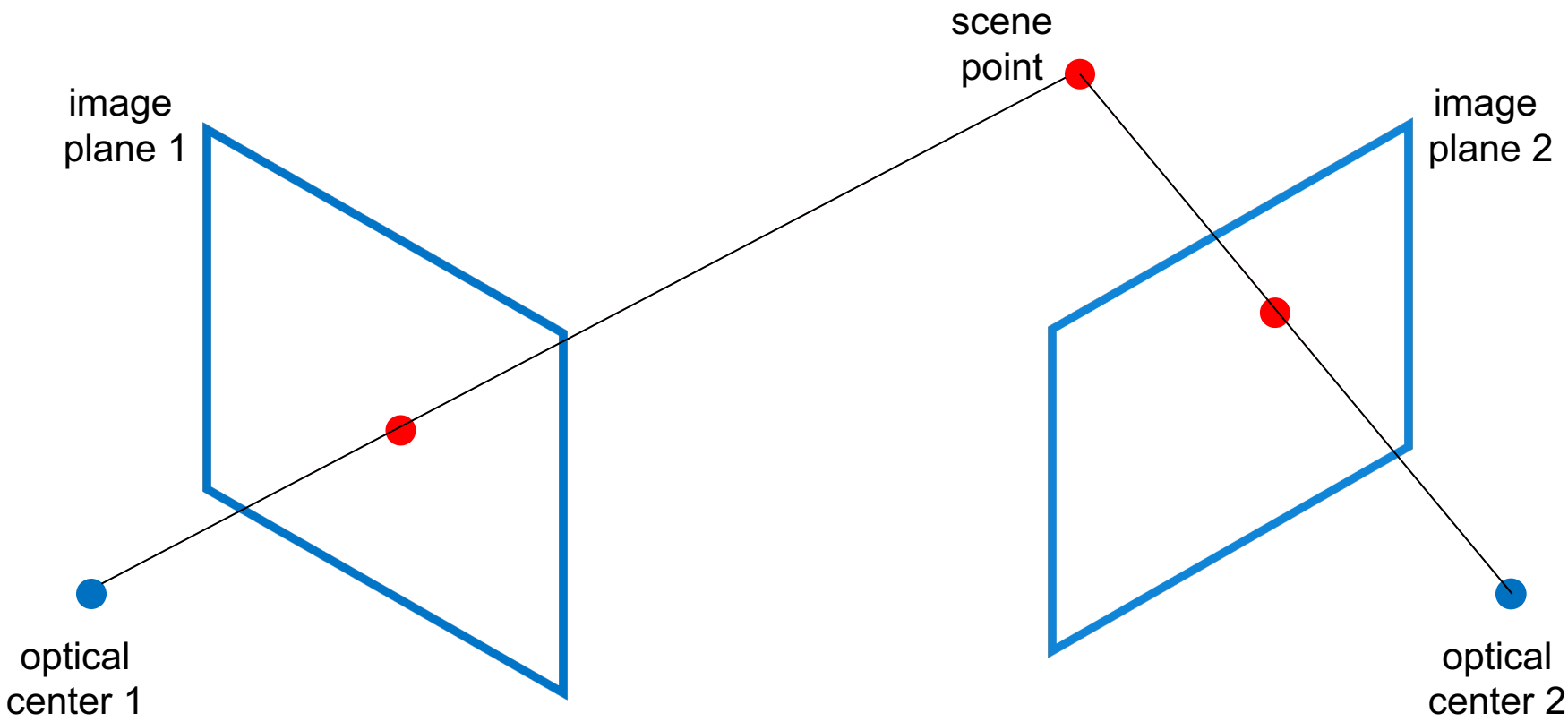


Camera Model

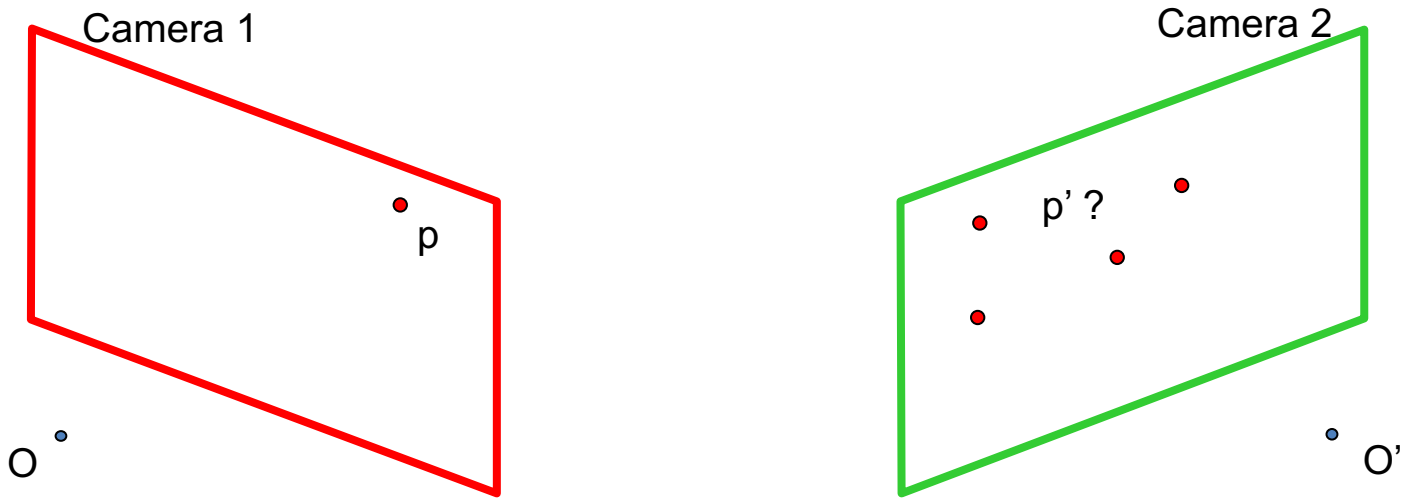
- Pinhole model



Epipolar Geometry

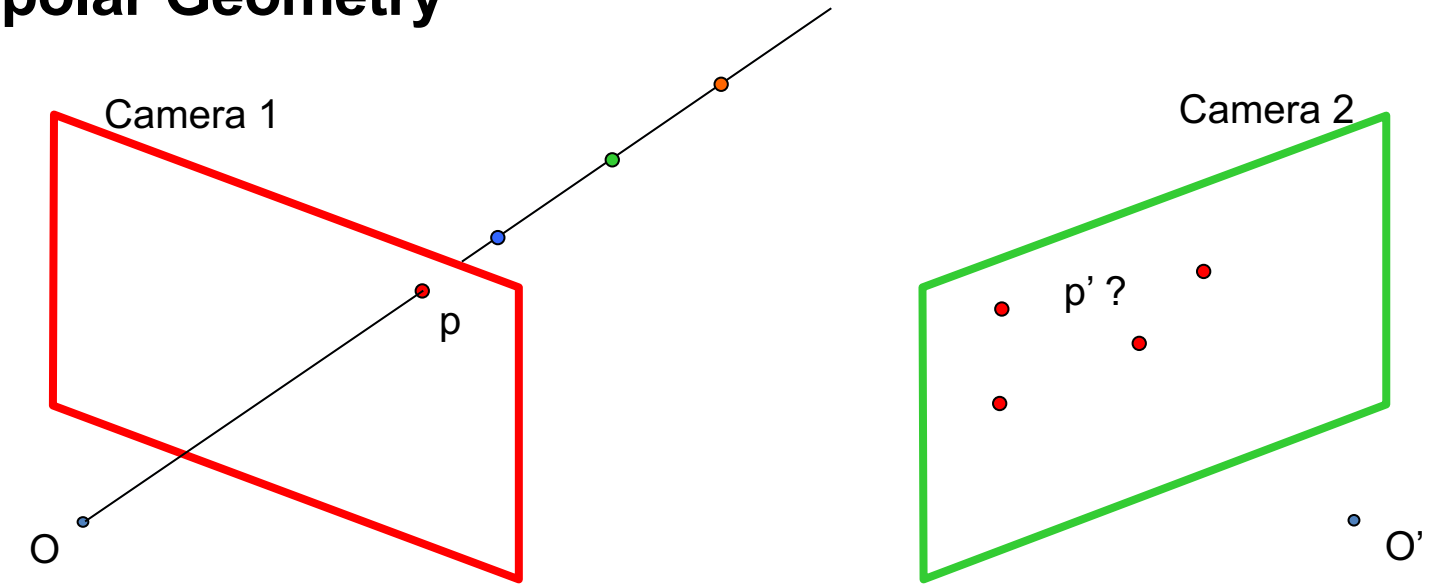


Epipolar Geometry



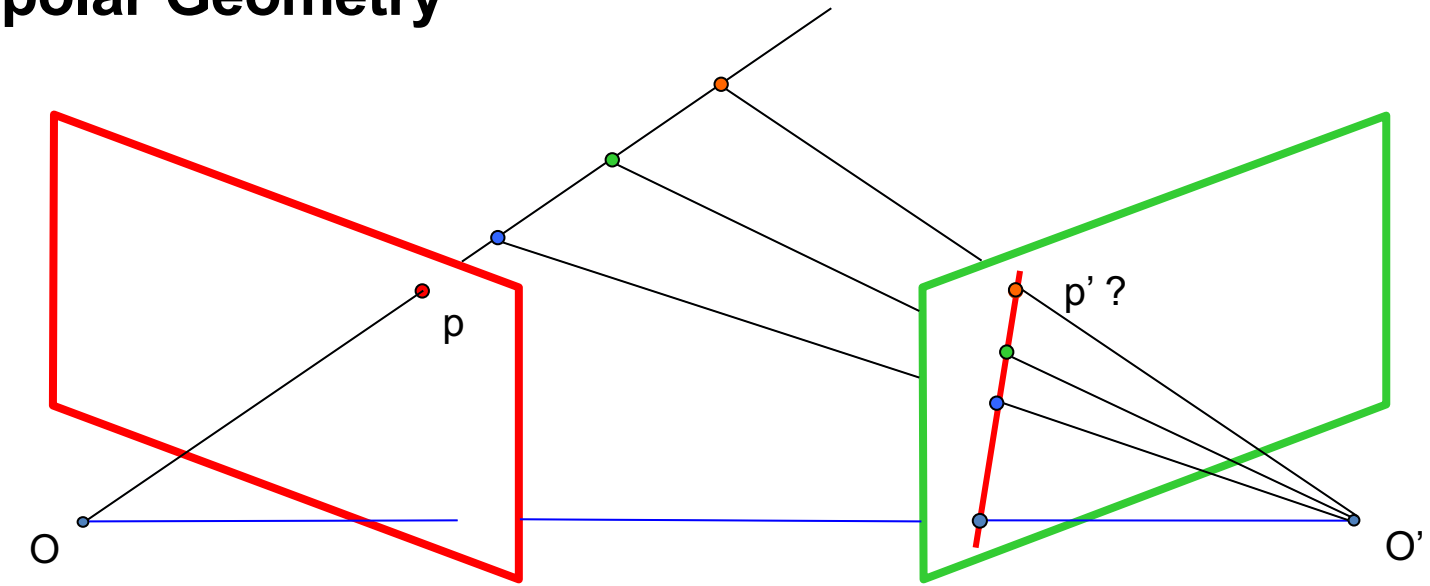
If we see a point in camera 1, are there any constraints on where we will find it on camera 2?

Epipolar Geometry



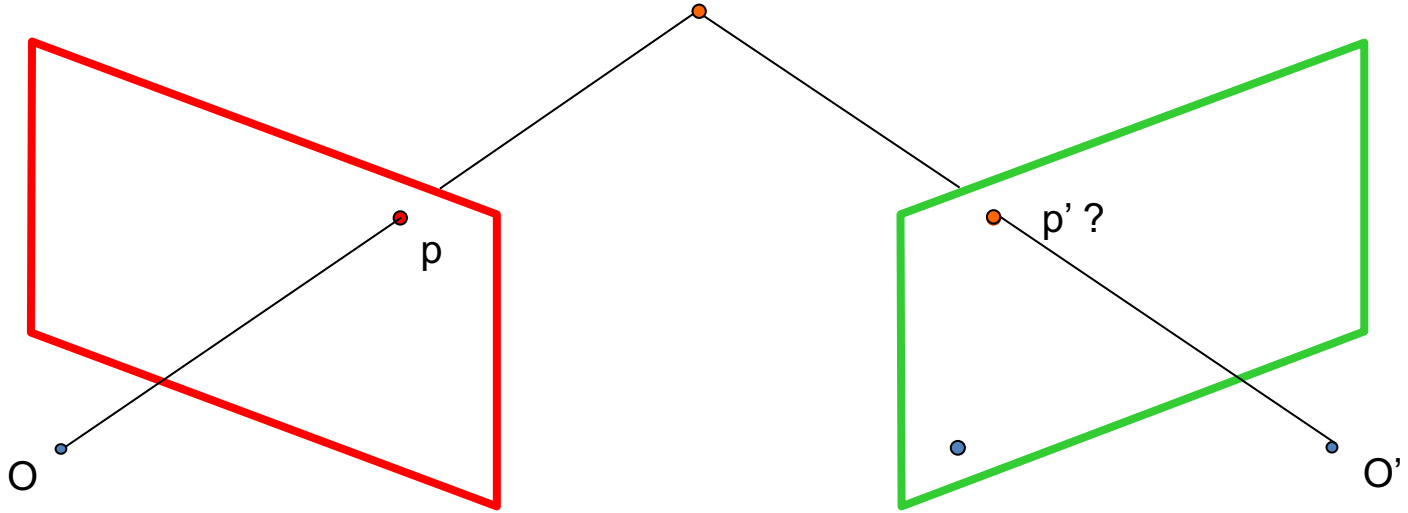
If we see a point in camera 1, are there any constraints on where we will find it on camera 2?

Epipolar Geometry



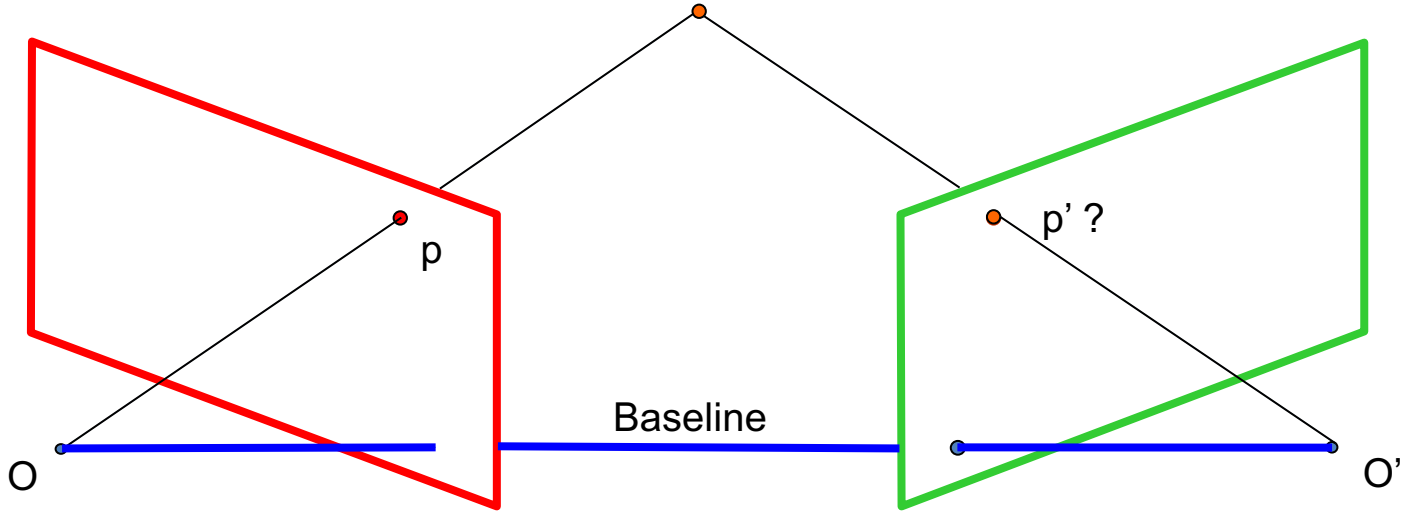
Adapted from W. T. Freeman, P. Isola, A. Torralba

Epipolar Geometry



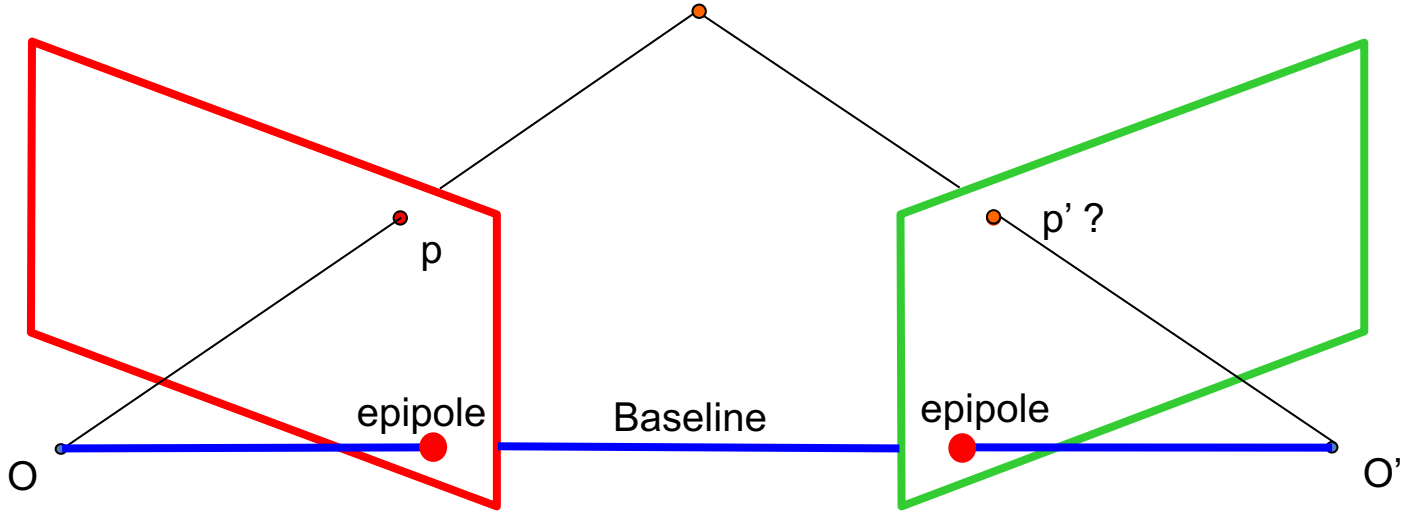
Adapted from W. T. Freeman, P. Isola, A. Torralba

Epipolar Geometry



Baseline: the line connecting the two camera centers

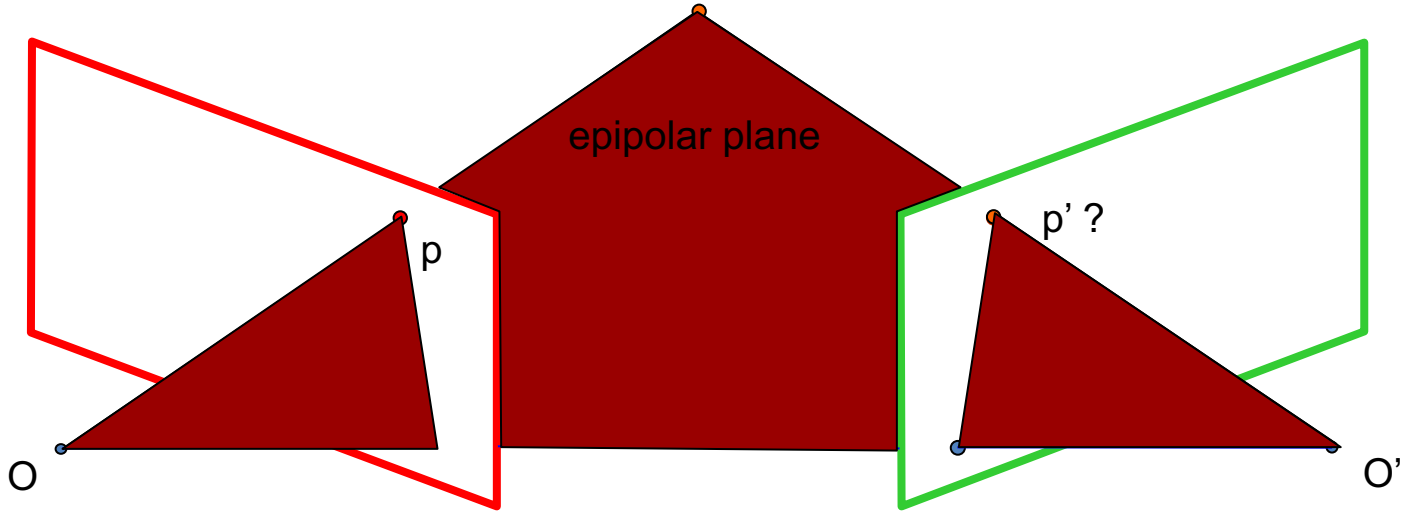
Epipolar Geometry



Baseline: the line connecting the two camera centers

Epipole: point of intersection of *baseline* with the image plane

Epipolar Geometry

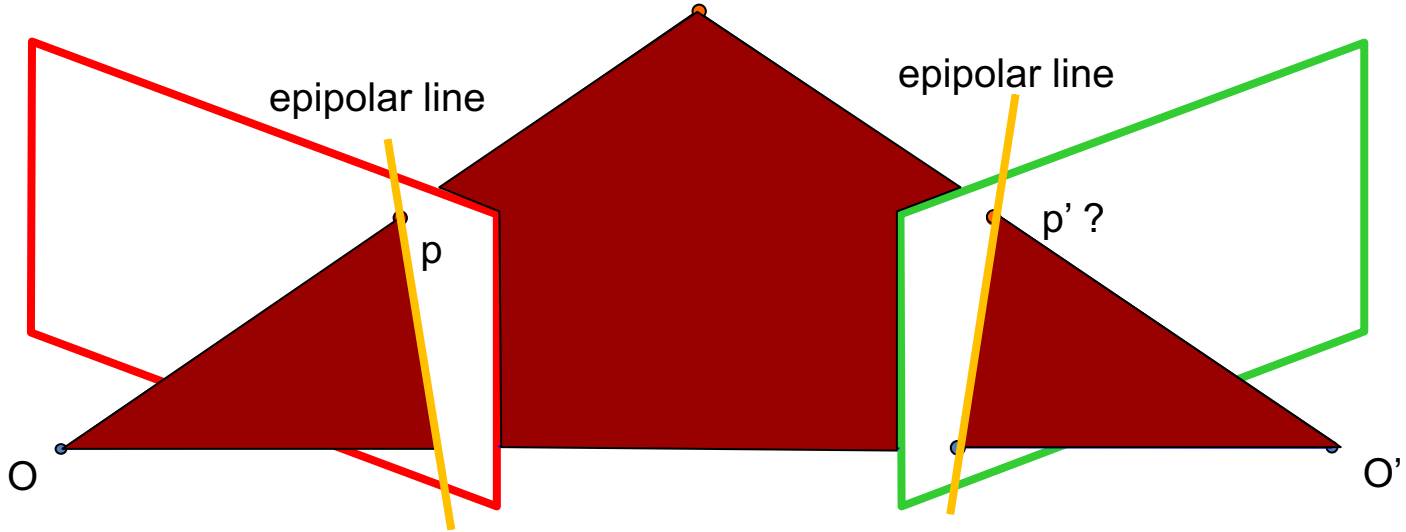


Baseline: the line connecting the two camera centers

Epipole: point of intersection of *baseline* with the image plane

Epipolar plane: the plane that contains the two camera centers and a 3D point in the world

Epipolar Geometry



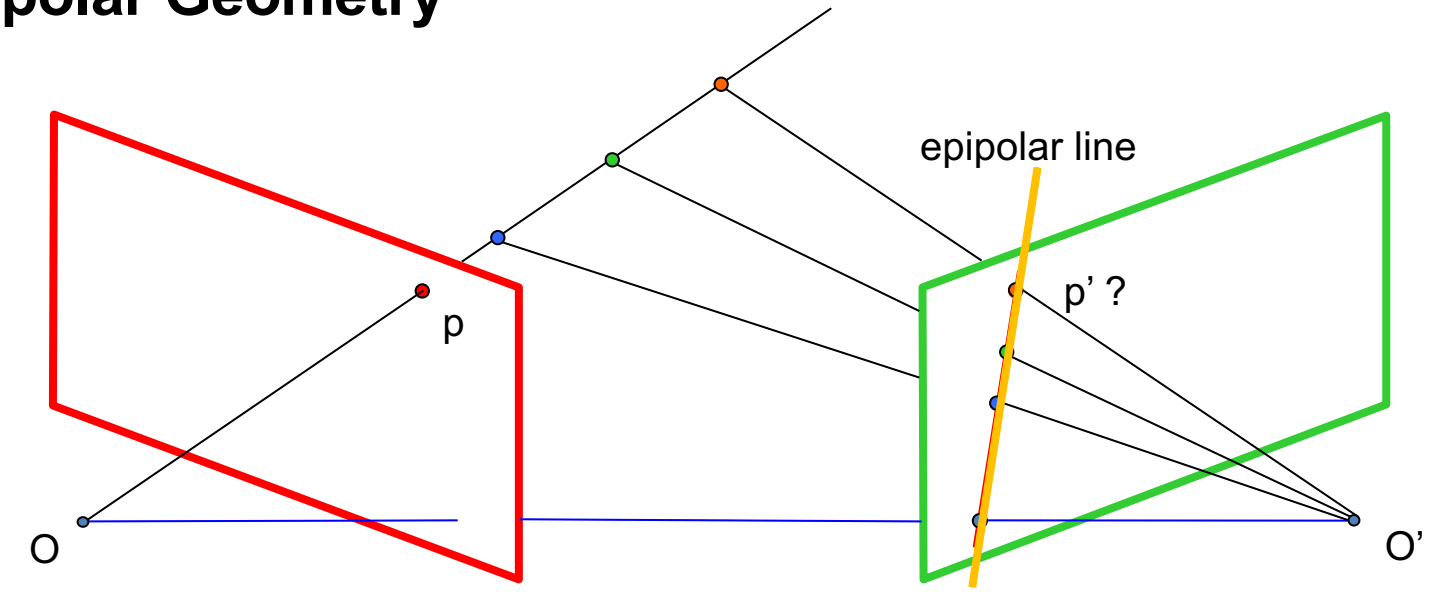
Baseline: the line connecting the two camera centers

Epipole: point of intersection of *baseline* with the image plane

Epipolar plane: the plane that contains the two camera centers and a 3D point in the world

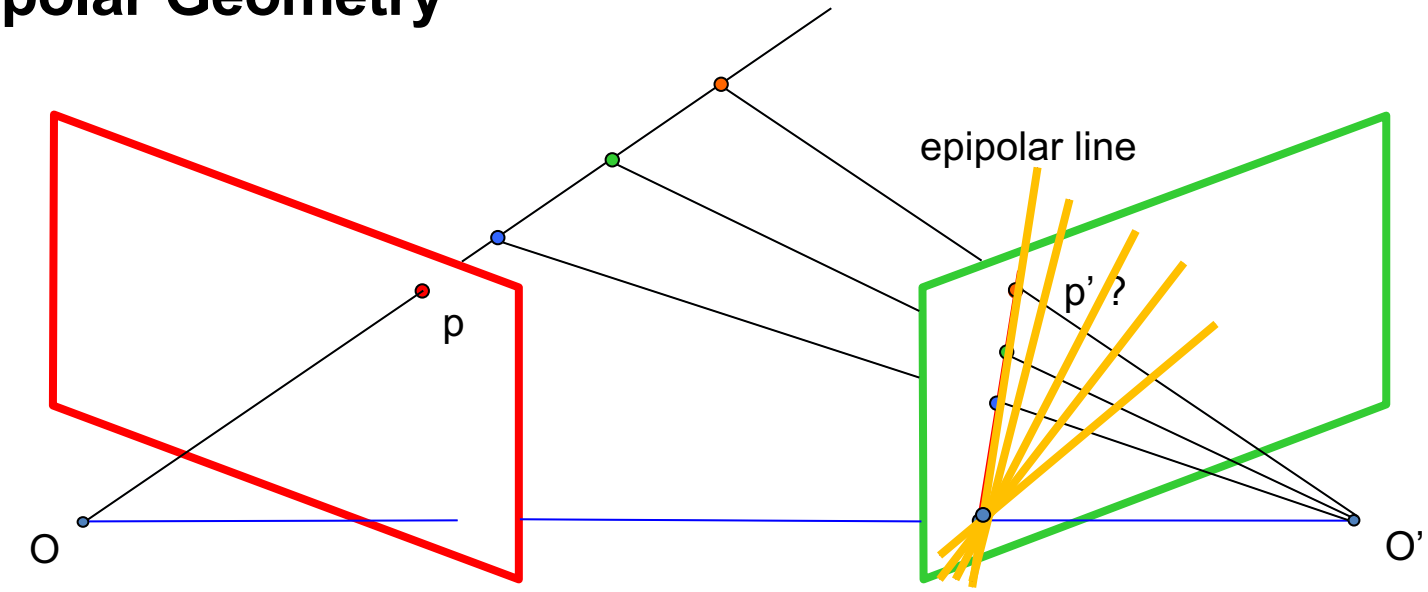
Epipolar line: intersection of the *epipolar plane* with each image plane

Epipolar Geometry



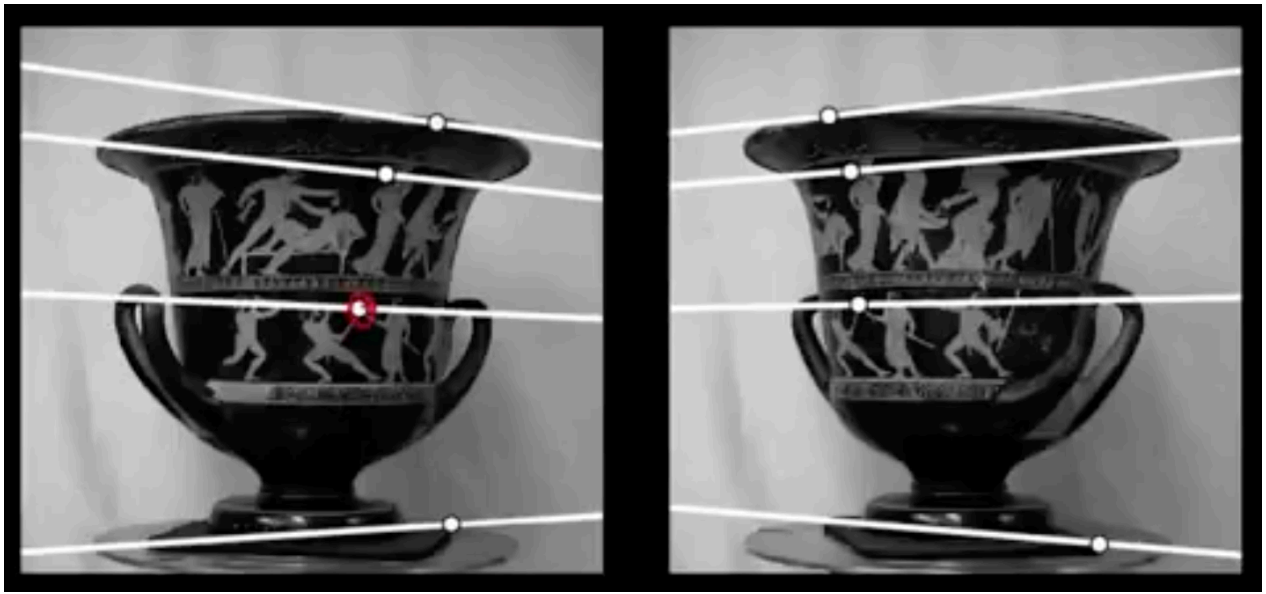
- We can search for matches across epipolar lines.
  - Search space for correspondences reduces to a 1D problem!

Epipolar Geometry



- We can search for matches across epipolar lines
  - Search space for correspondences reduces to a 1D problem!
- All epipolar lines intersect at the epipoles

Epipolar Geometry



- We can search for matches across epipolar lines
  - Search space for correspondences reduces to a 1D problem!
- All epipolar lines intersect at the epipoles
  - **Where are the epipoles in this case?**

Figure from Hartley & Zisserman

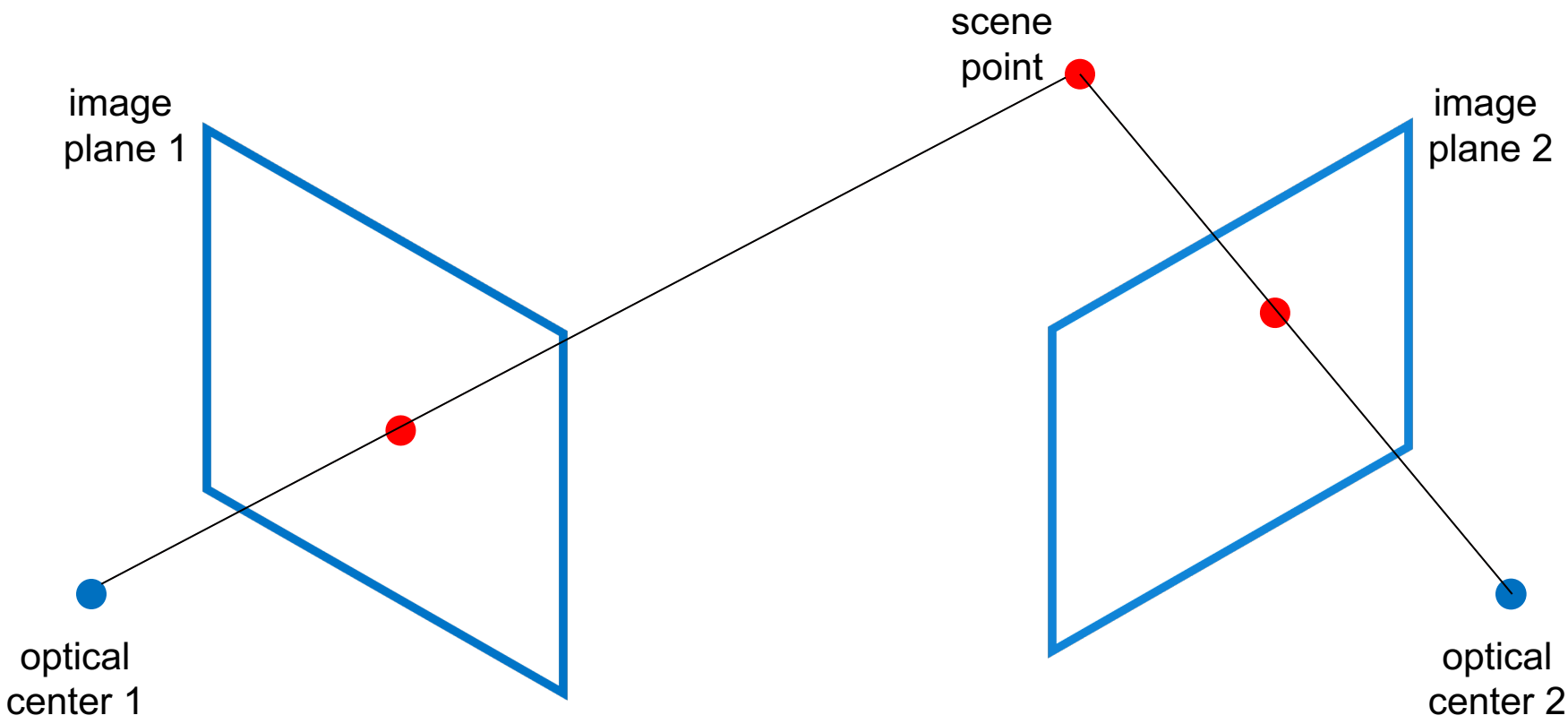
Outline

- What is Stereo Vision?

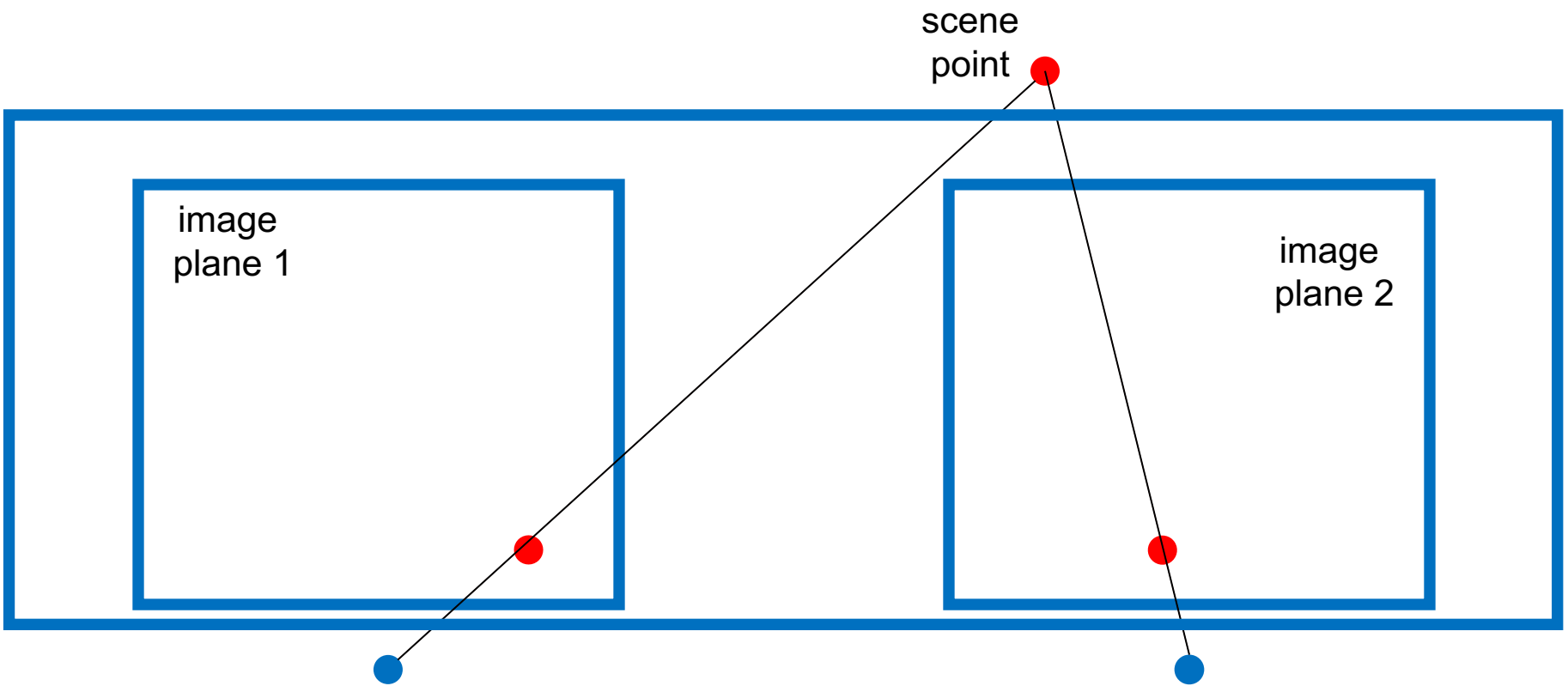
- Stereo/Epipolar Geometry
- Rectified Stereo Case
- Depth from Stereo Matches
- Correspondence Problem
  - **Dense** vs Sparse Correspondence
  - **Local** vs Global Correspondence
  - (Dis)-Similarity Measures

- Summary

Rectified Stereo – a simpler case

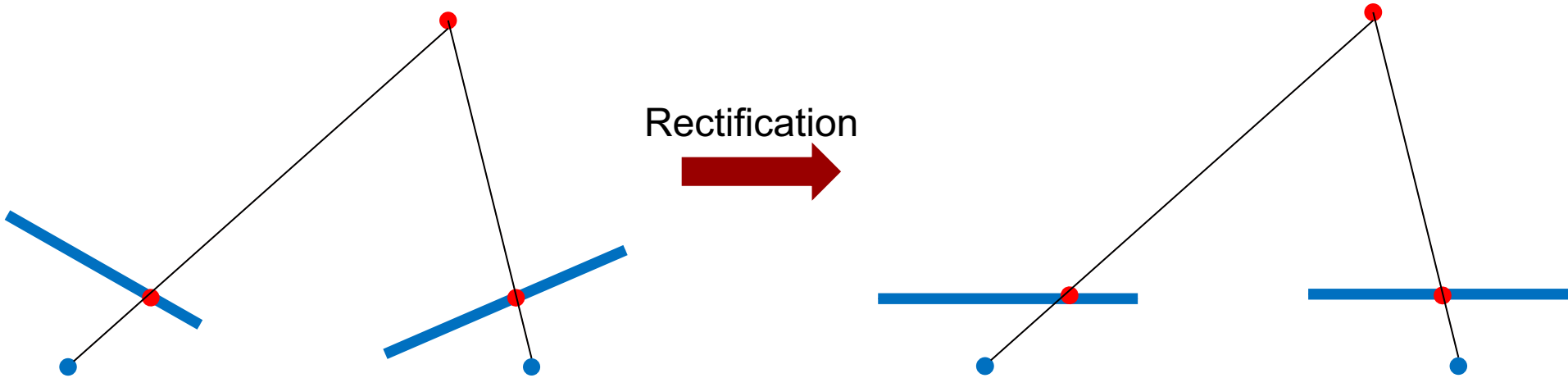


Rectified Stereo – a simpler case

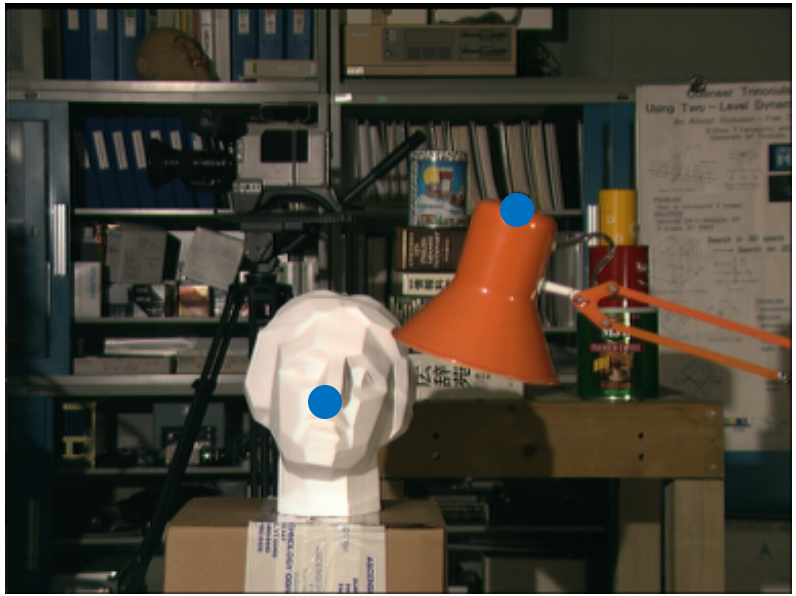


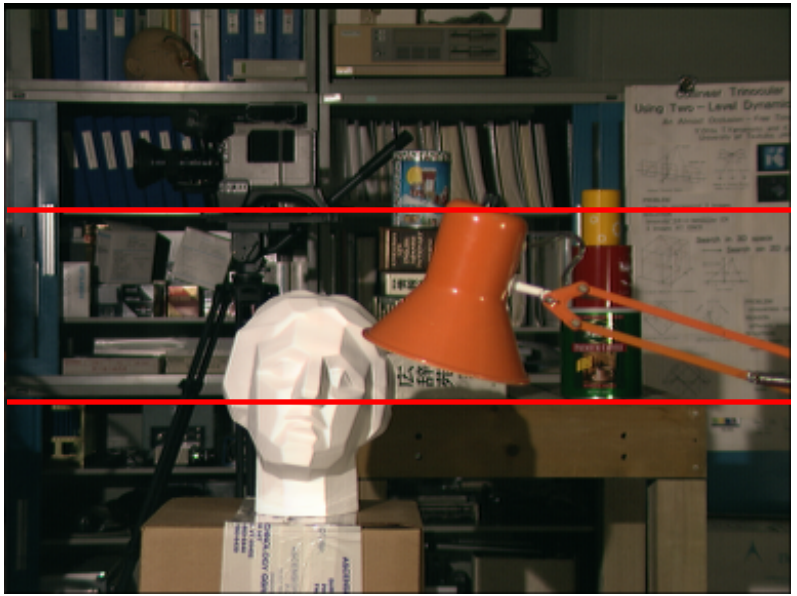
Rectified Stereo – a simpler case

- Rectification:
  - The initial images are reprojected on a common plane that is parallel to the baseline B joining the optical centers of the initial images.
  - Epipolar lines become parallel (and under certain conditions they become also horizontal)

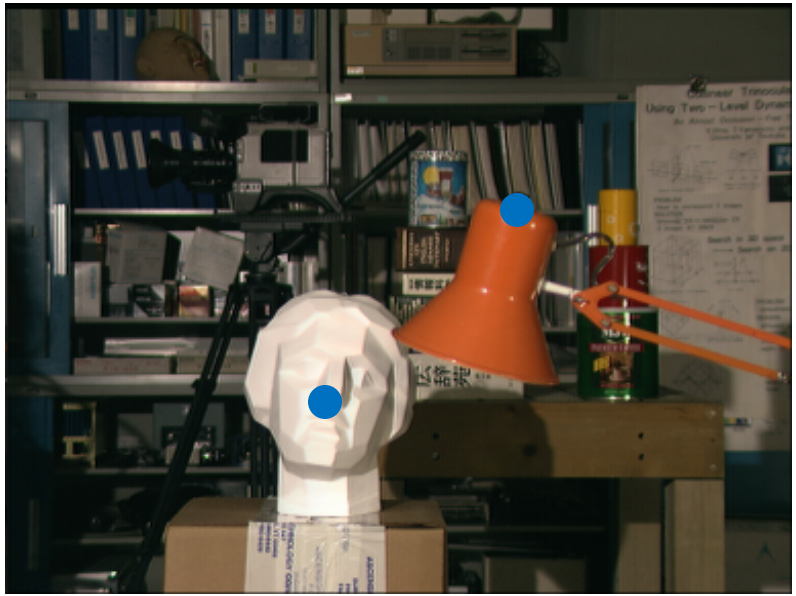


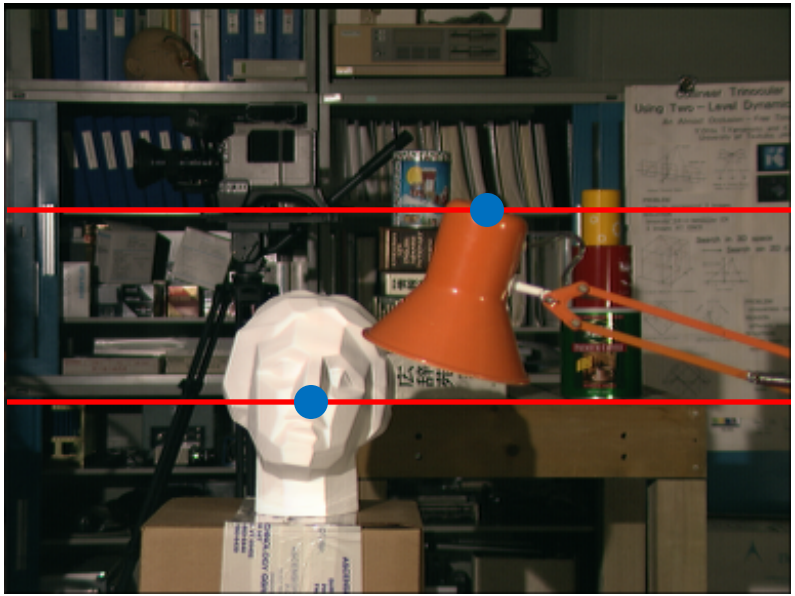
Rectified Stereo





Rectified Stereo





- “All epipolar lines intersect at the epipoles”
  - **Where are the epipoles in this case?**

Outline

- What is Stereo Vision?

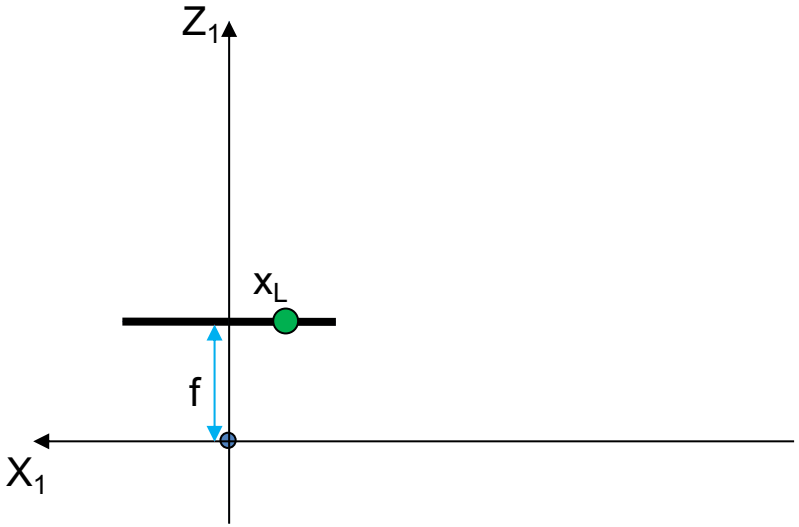
- Stereo/Epipolar Geometry
- Rectified Stereo Case
- Depth from Stereo Matches
- Correspondence Problem
  - **Dense** vs Sparse Correspondence
  - **Local** vs Global Correspondence
  - (Dis)-Similarity Measures

- Summary

Depth from Stereo Matches

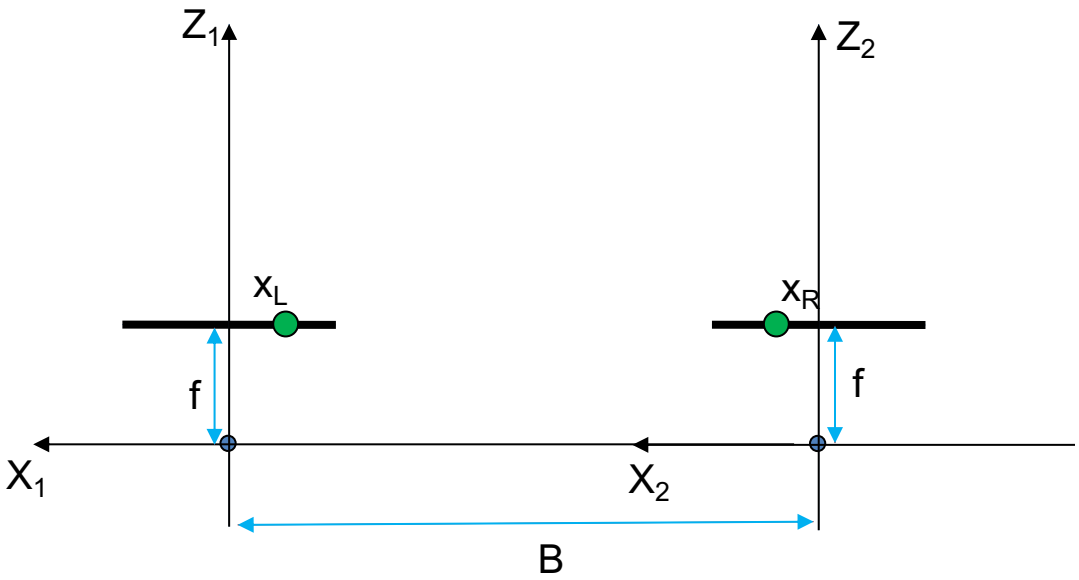
- Let us assume (for now!!) that:
  - we can check the points along the epipolar line and
  - we can find the point (on the right image) that is most similar to our reference point (in the left image), i.e. we can solve the correspondence problem!

Depth from Stereo Matches



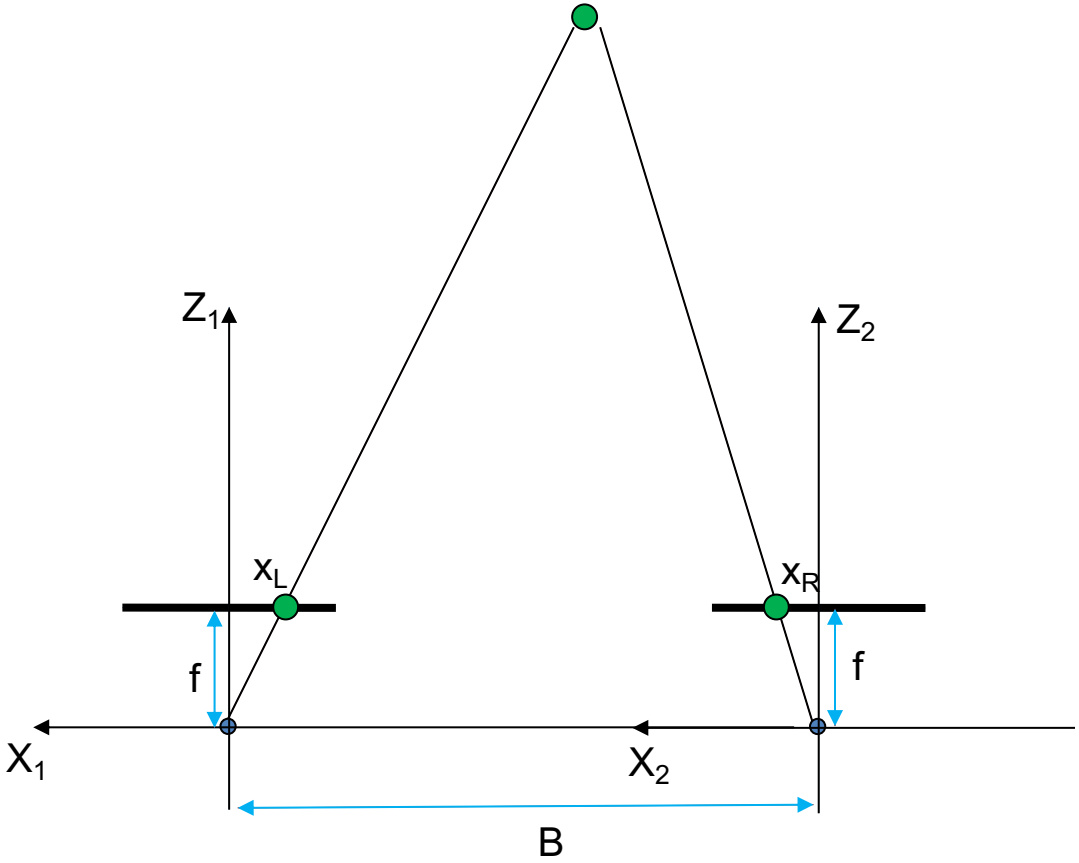
Adapted from W. T. Freeman, P. Isola, A. Torralba

Depth from Stereo Matches



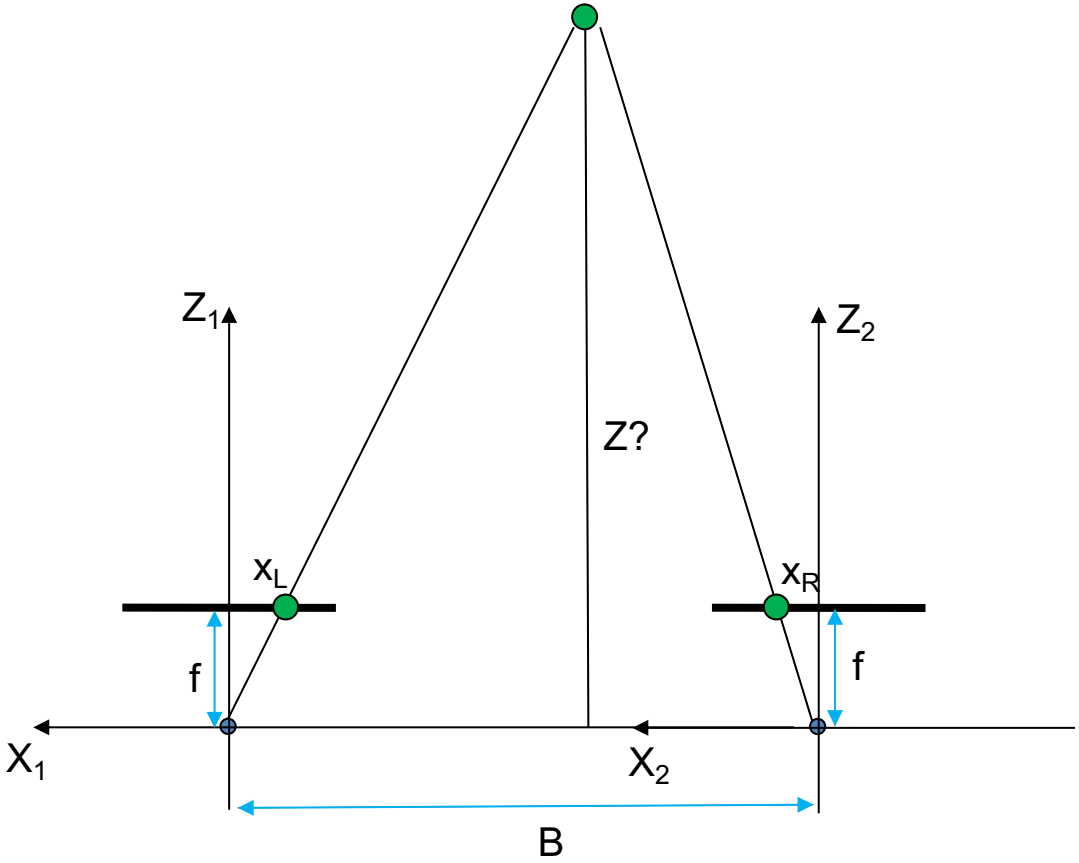
Adapted from W. T. Freeman, P. Isola, A. Torralba

Depth from Stereo Matches



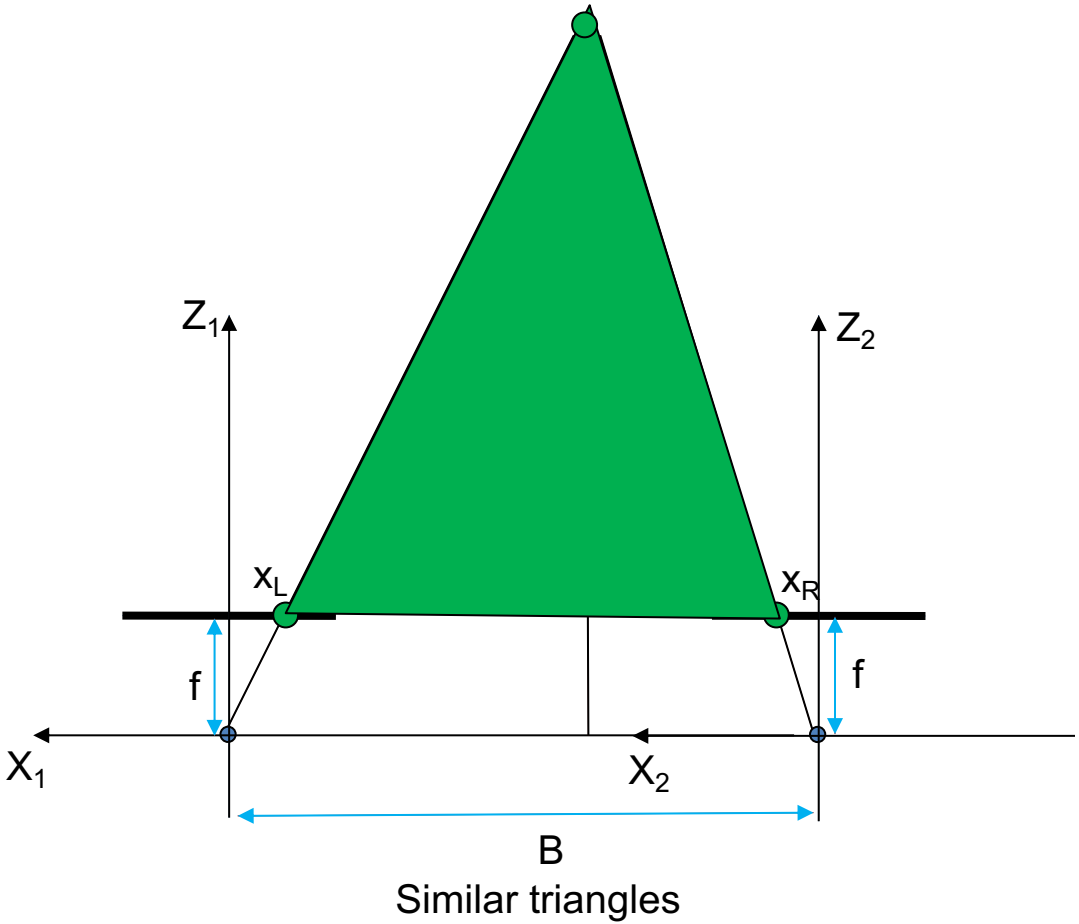
Adapted from W. T. Freeman, P. Isola, A. Torralba

Depth from Stereo Matches



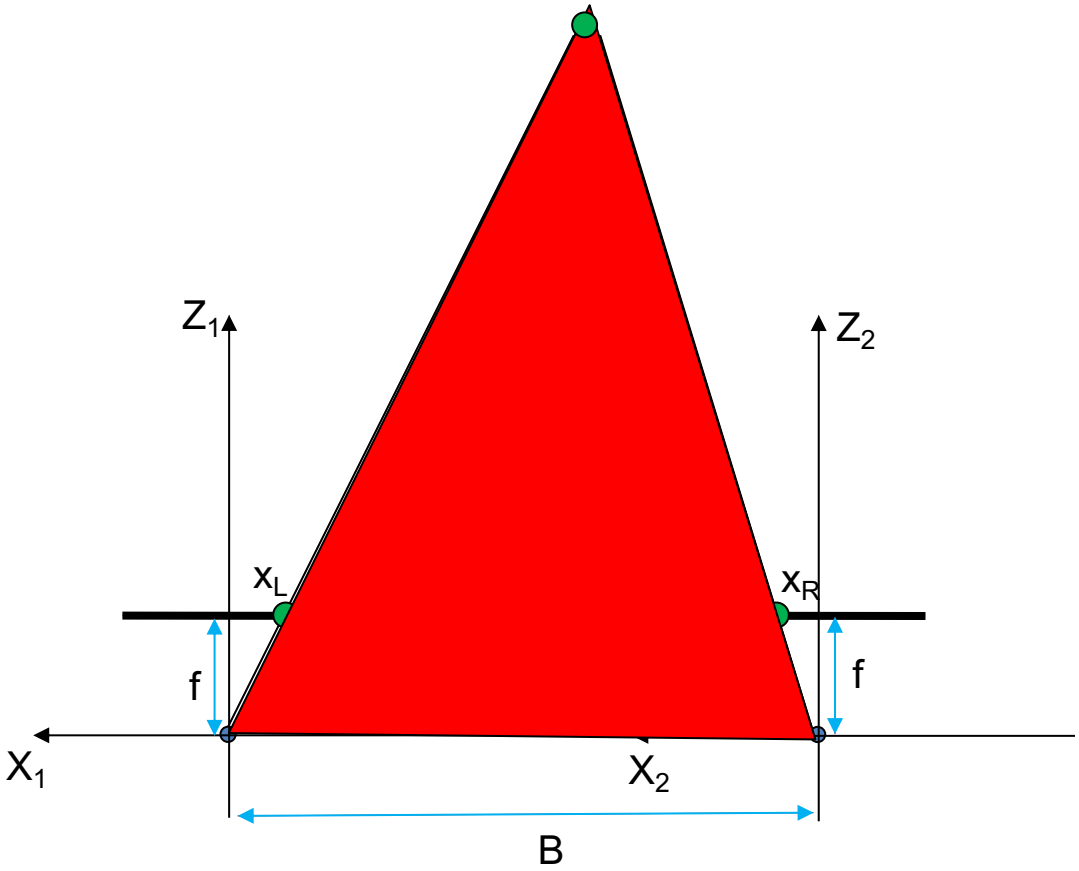
Adapted from W. T. Freeman, P. Isola, A. Torralba

Depth from Stereo Matches



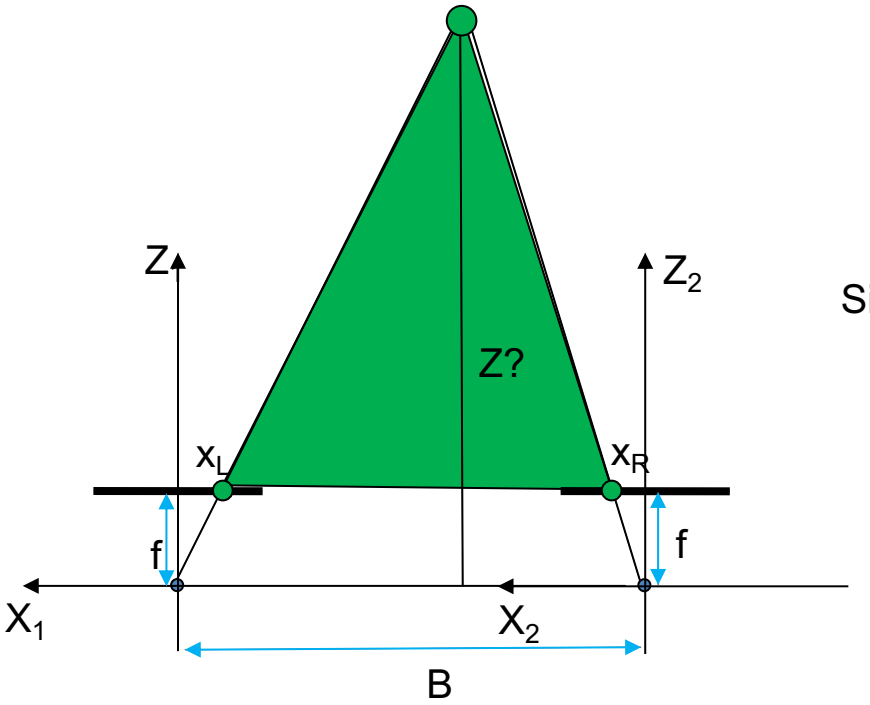
Similar triangles

Depth from Stereo Matches



Similar triangles

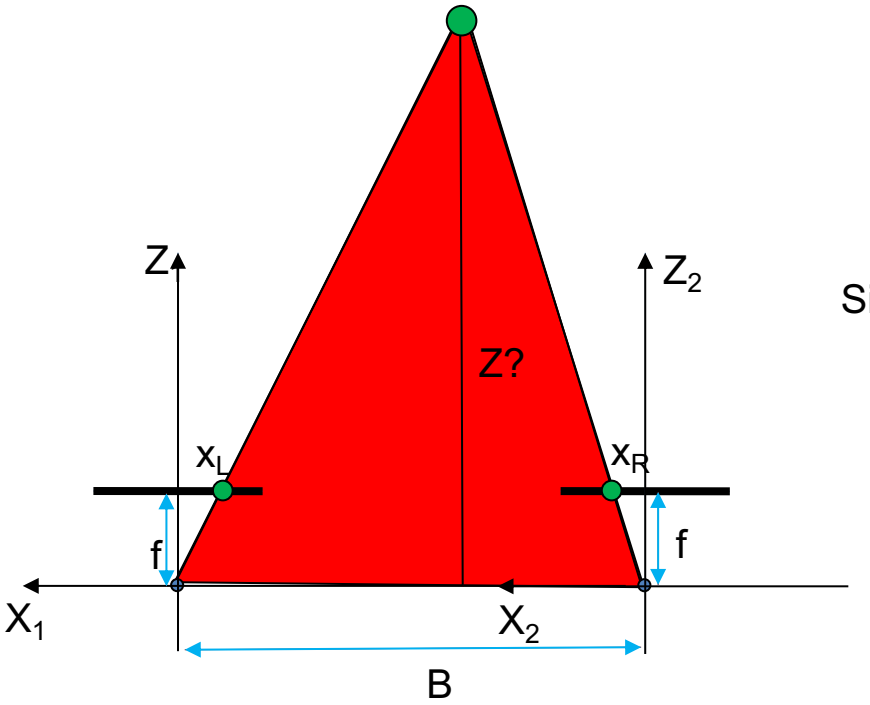
Depth from Stereo Matches



Similar triangles:

$$\frac{B+X_L-X_R}{Z-f} =$$

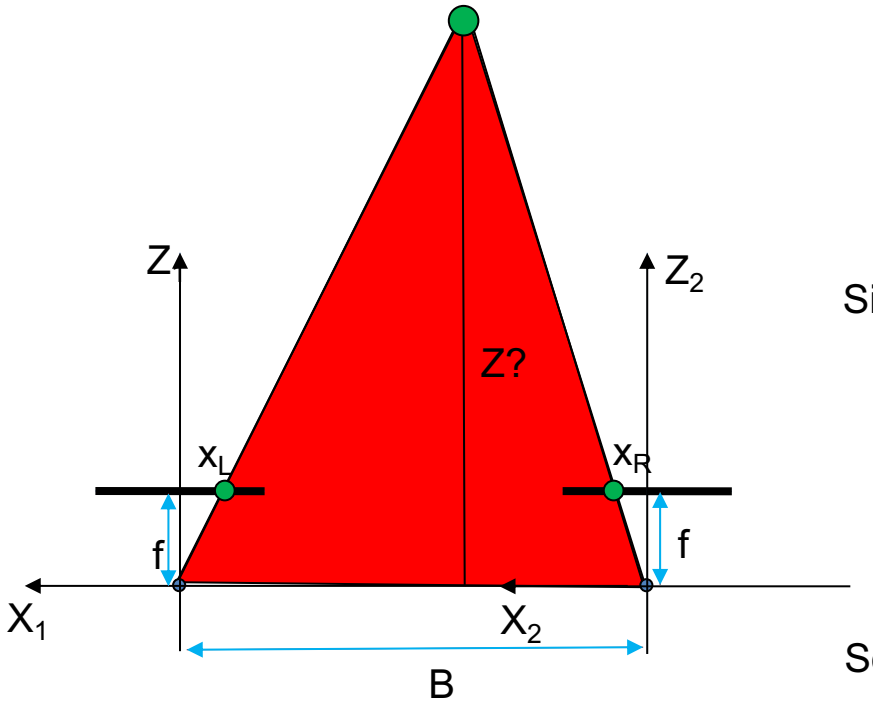
Depth from Stereo Matches



Similar triangles:

$$\frac{B+X_L-X_R}{Z-f} = \frac{B}{Z}$$

Depth from Stereo Matches



Similar triangles:

$$\frac{B+X_L-X_R}{Z-f} = \frac{B}{Z}$$

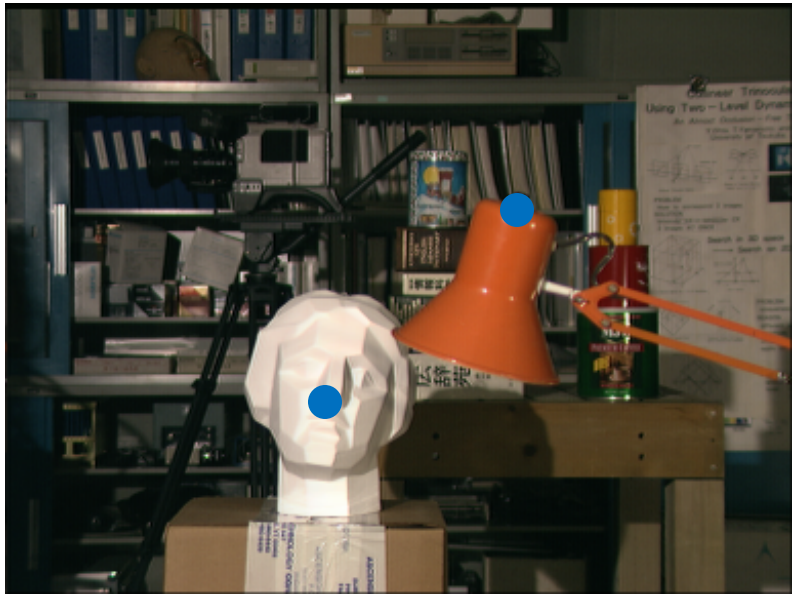
Solving for Z:

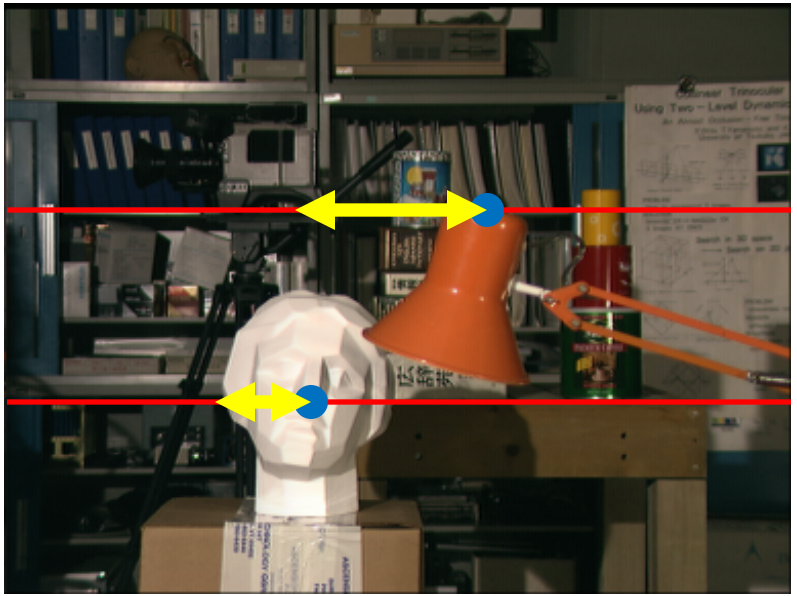
$$Z = f\,\frac{B}{X_L - X_R}$$

Disparity

Adapted from W. T. Freeman, P. Isola, A. Torralba

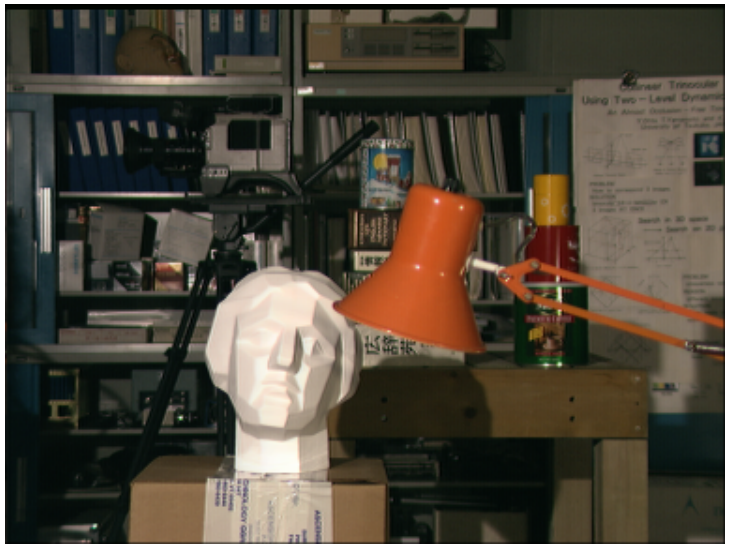
Depth from Stereo Matches





Depth from Stereo Matches

Reference Image



Disparity Map

$$Z = f \frac{B}{X_L - X_R}$$

Disparity

Outline

- What is Stereo Vision?

- Stereo/Epipolar Geometry
- Rectified Stereo Case
- Depth from Stereo Matches
- Correspondence Problem
  - **Dense** vs Sparse Correspondence
  - **Local** vs Global Correspondence
  - (Dis)-Similarity Measures

- Summary

Correspondence Problem

- We have assumed (up to now!!) that:
  - we can check the points along the epipolar line and
  - we can find the point (on the right image) that is **most similar** to our reference point (in the left image), i.e. we can **solve the correspondence problem**!

- *How can we indeed match corresponding pixels between the two stereo images?*

Correspondence Problem

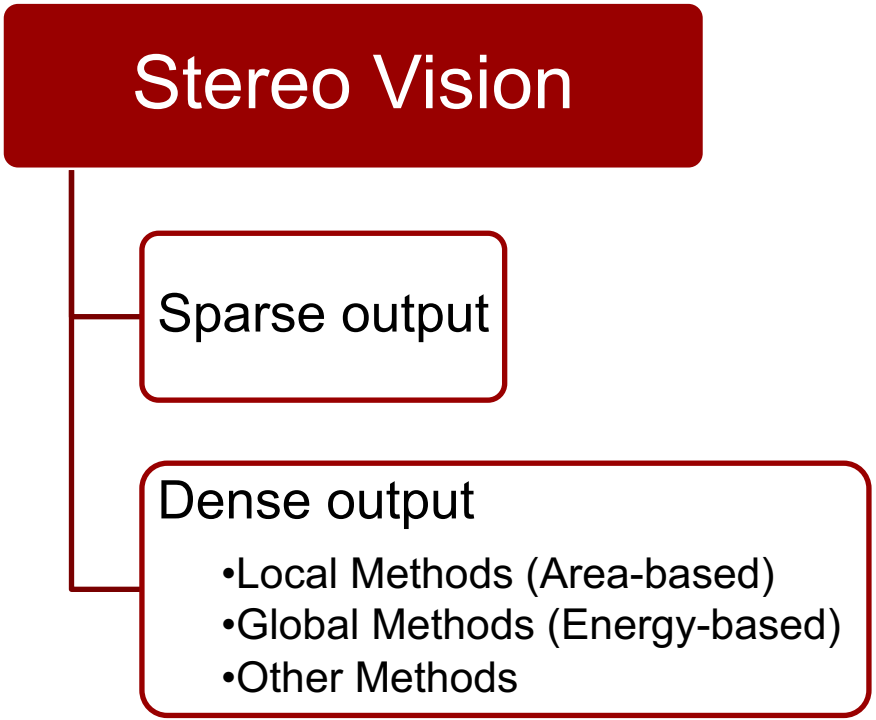
- Beyond the hard constraint of epipolar geometry, there are “soft” constraints to help identify corresponding points
  - Similarity
  - Uniqueness
  - Ordering
  - Disparity gradient is limited

Correspondence Problem

- To find matches in the image pair, we will assume
  - Most scene points visible from both views
  - Image regions for the matches are similar in appearance

Correspondence Problem

- It depends!
  - Do we need **dense** or **sparse** stereo matching?



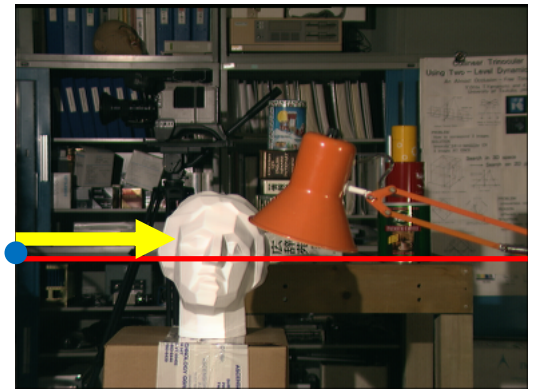
Sparse Stereo Correspondence

- Extract features (e.g. SIFT, SURF, Harris,…) and match them!

  – Pros?
  – Cons?

Dense Stereo Correspondence : Local Methods

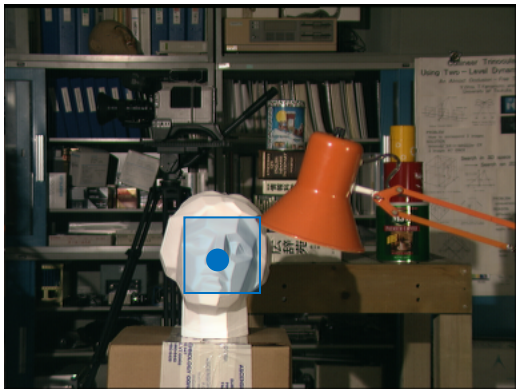
- Try to find correspondences for all the pixels of the reference image.
- For each epipolar line
  - For each pixel in the left image
    - Compare with every pixel on same epipolar line in right image
    - Choose the pixel that maximizes a similarity metric (or minimizes a dissimilarity metric!).

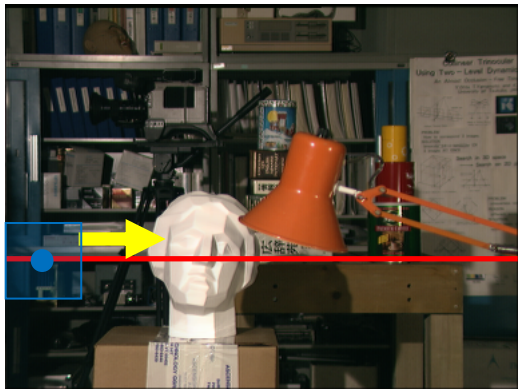


Dense Stereo Correspondence

- Try to find correspondences for all the pixels of the reference image.
- For each epipolar line
  - For each pixel in the left image
    - Compare with every pixel on same epipolar line in right image
    - Choose the pixel that maximizes a similarity metric (or minimizes a dissimilarity metric!).

- Improvement: don't match individual pixels, but rather match windows!





Stereo Correspondence Metrics

- Sum of Absolute Differences (SAD)

$$SAD(x, y, d) = \sum_{x,y \in W} |I_l(x, y) - I_r(x, y - d)|$$

- Sum of Squared Differences (SSD)

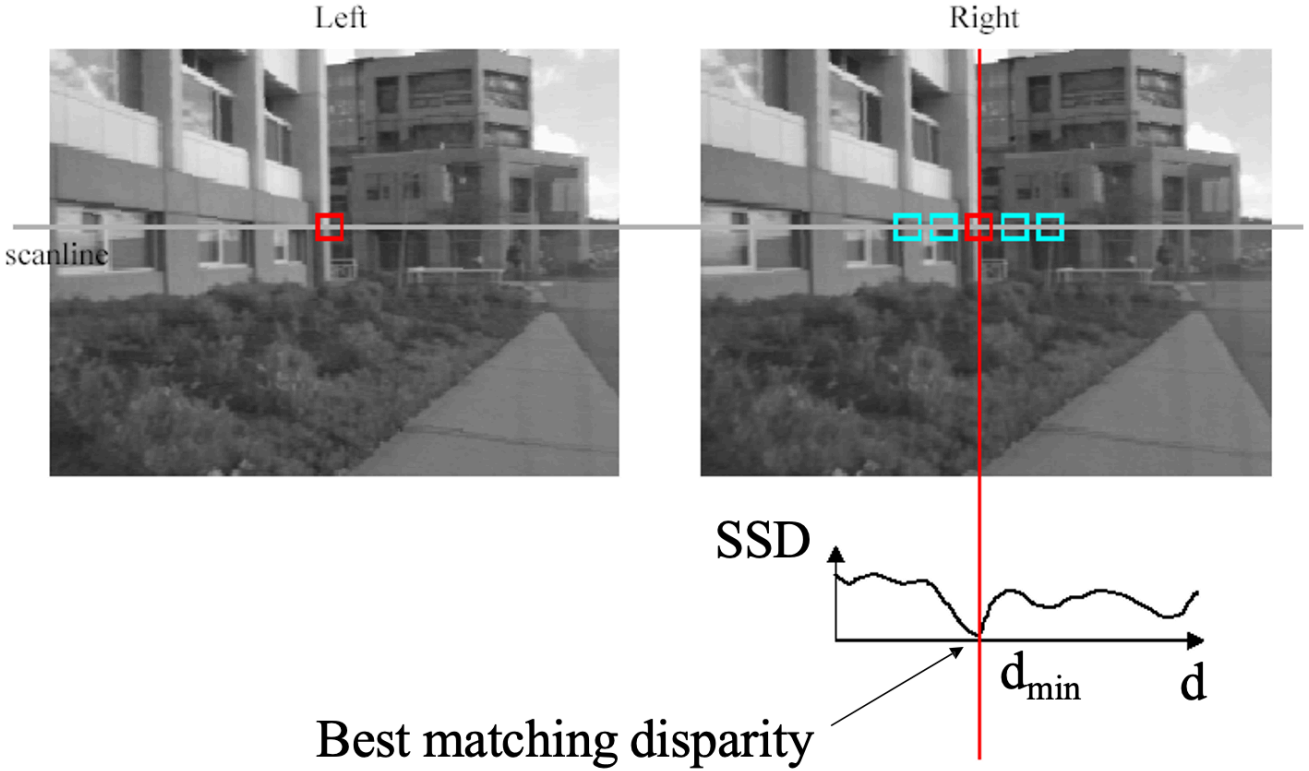
$$SSD(x, y, d) = \sum_{x,y \in W} (I_l(x, y) - I_r(x, y - d))^2$$

- Normalized Cross-Correlation

$$NCC(x, y, d) = \frac{\sum_{x,y \in W} I_l(x, y) \cdot I_r(x, y - d)}{\sqrt{\sum_{x,y \in W} I_l^2(x, y) \cdot \sum_{x,y \in W} I_r^2(x, y - d)}}$$

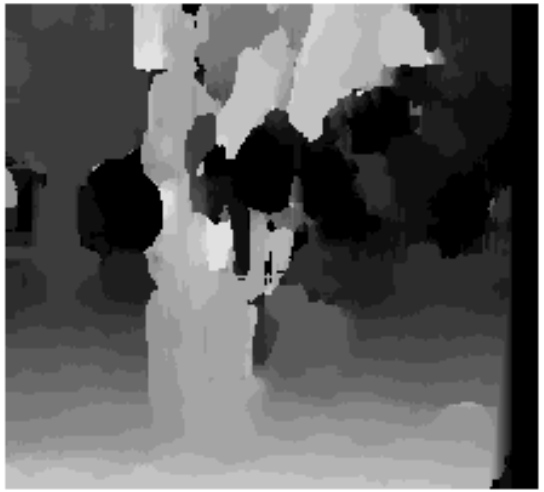
- …many many more!!!

Stereo Correspondence Metrics: SSD



Stereo Correspondence Metrics: SSD on various windows

W = 3



W = 20

- Small vs Big windows
- What are their Pros and Cons?

Figure from Li Zhang

Stereo Correspondence Metrics: Good/Bad areas

- In this stereo image pair:
  - what would be good areas to match?
  - where would you expect to face problems and why?

Figure from Hartley & Zisserman

Global Stereo Correspondence

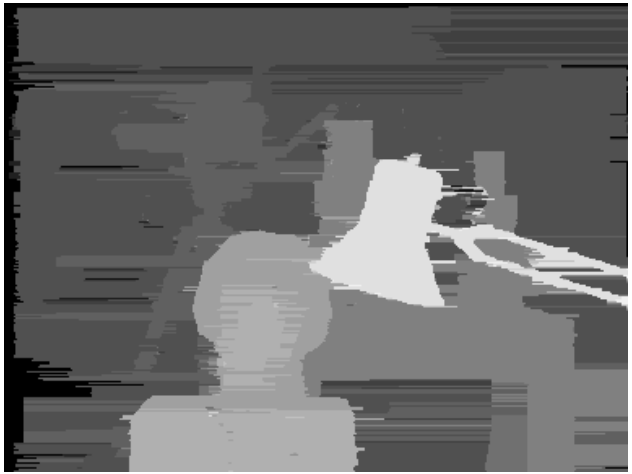
- Up to this point, the disparity of each pixel was determined only by the information of the pixel itself and its neighborhood.
  – Thus, those methods are called ”local” or “area-based” methods.

- Example: Result of a **local** SSD algorithm with W=21:

Global Stereo Correspondence

- Up to this point, the disparity of each pixel was determined only by the information of the pixel itself and its neighborhood.
  – Thus, those methods are called ”local” or “area-based” methods.

- Global methods find better solutions in expense of more computations
  – Optimize jointly the disparity values of all the pixels of each scanline (e.g. Dynamic Programming)



Global Stereo Correspondence

- Up to this point, the disparity of each pixel was determined only by the information of the pixel itself and its neighborhood.
  - Thus, those methods are called ”local” or “area-based” methods.

- Global methods find better solutions in expense of more computations
  - Optimize jointly the disparity values of all the pixels of each scanline (e.g. Dynamic Programming)
  - Optimize jointly the disparity values of all the pixels of the image (e.g. graph cuts)

Global Stereo Correspondence

- Up to this point, the disparity of each pixel was determined only by the information of the pixel itself and its neighborhood.
  - Thus, those methods are called ”local” or “area-based” methods.

- Global methods find better solutions in expense of more computations
  - Optimize jointly the disparity values of all the pixels of each scanline (e.g. Dynamic Programming)
  - Optimize jointly the disparity values of all the pixels of the image (e.g. graph cuts)

    - *In global algorithms, stereo correspondence is formulated as an energy function minimization problem, consisting of data and smoothness terms.*

Outline

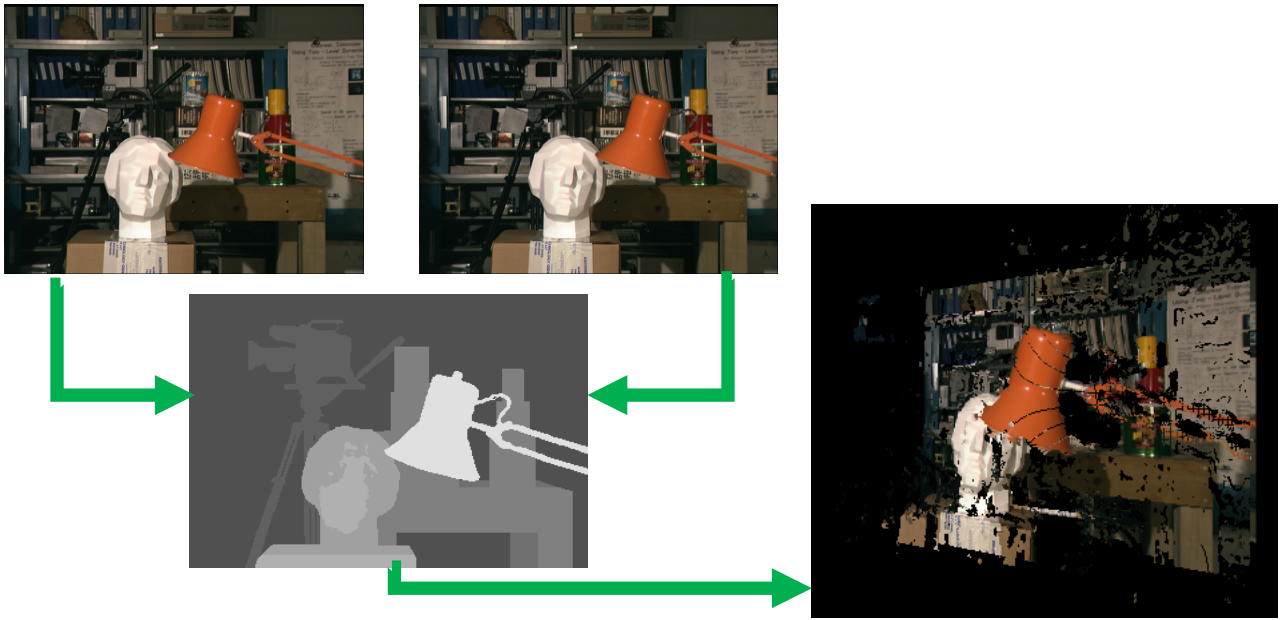
- What is Stereo Vision?

- Stereo/Epipolar Geometry
- Rectified Stereo Case
- Depth from Stereo Matches
- Correspondence Problem
  - **Dense** vs Sparse Correspondence
  - **Local** vs Global Correspondence
  - (Dis)-Similarity Measures

- Summary

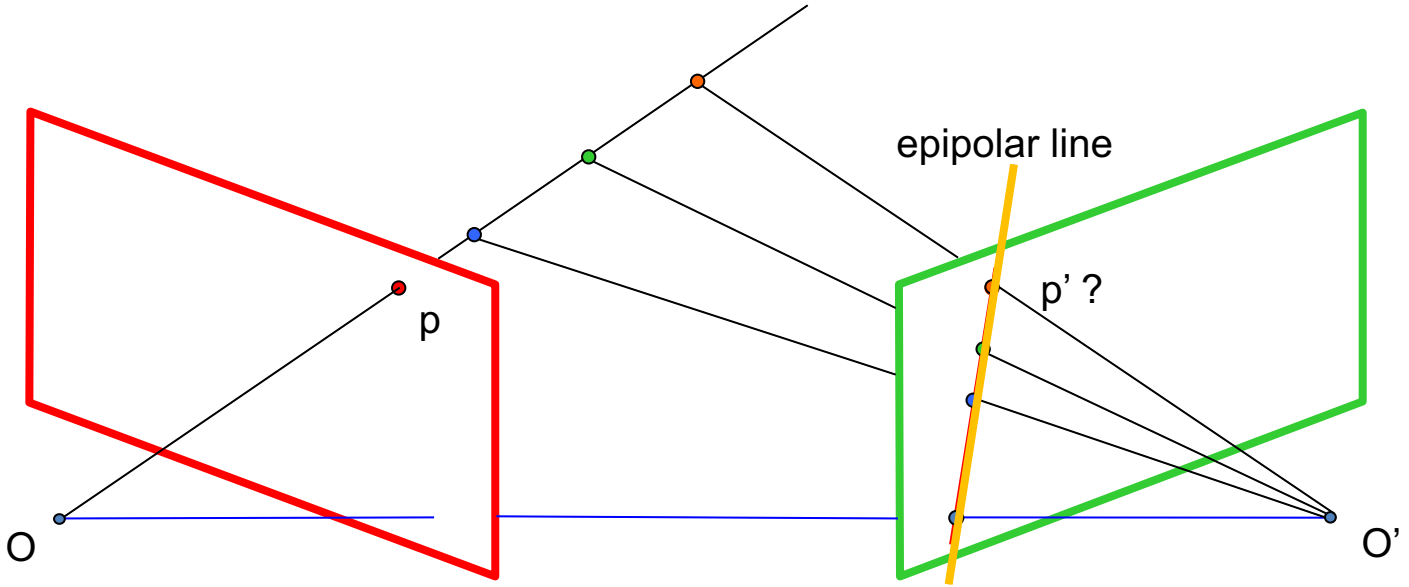
Summary

- We discussed about what Stereo Vision is.



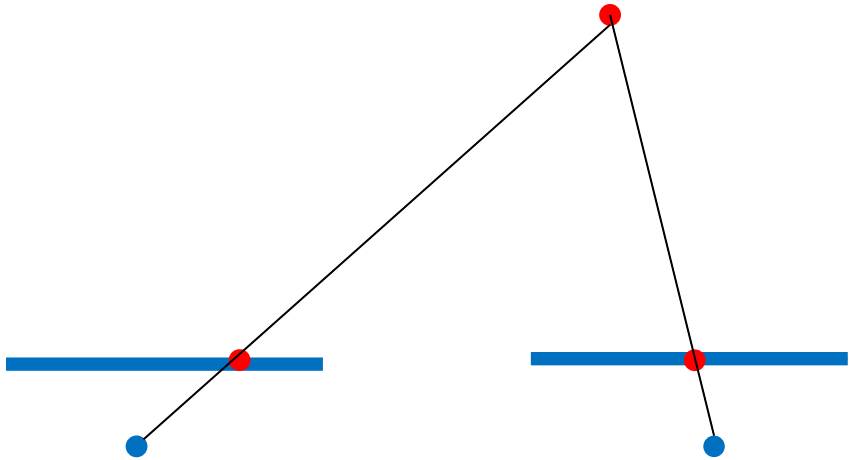
Summary

- We discussed about what Stereo Vision is.
- We learned about :
  – Stereo/Epipolar Geometry



Summary

- We discussed about what Stereo Vision is.
- We learned about :
  - Stereo/Epipolar Geometry
  - Rectified Stereo Case



Summary

- We discussed about what Stereo Vision is.
- We learned about :
  - Stereo/Epipolar Geometry
  - Rectified Stereo Case
  - Depth from Stereo Matches

Reference Image

Disparity Map

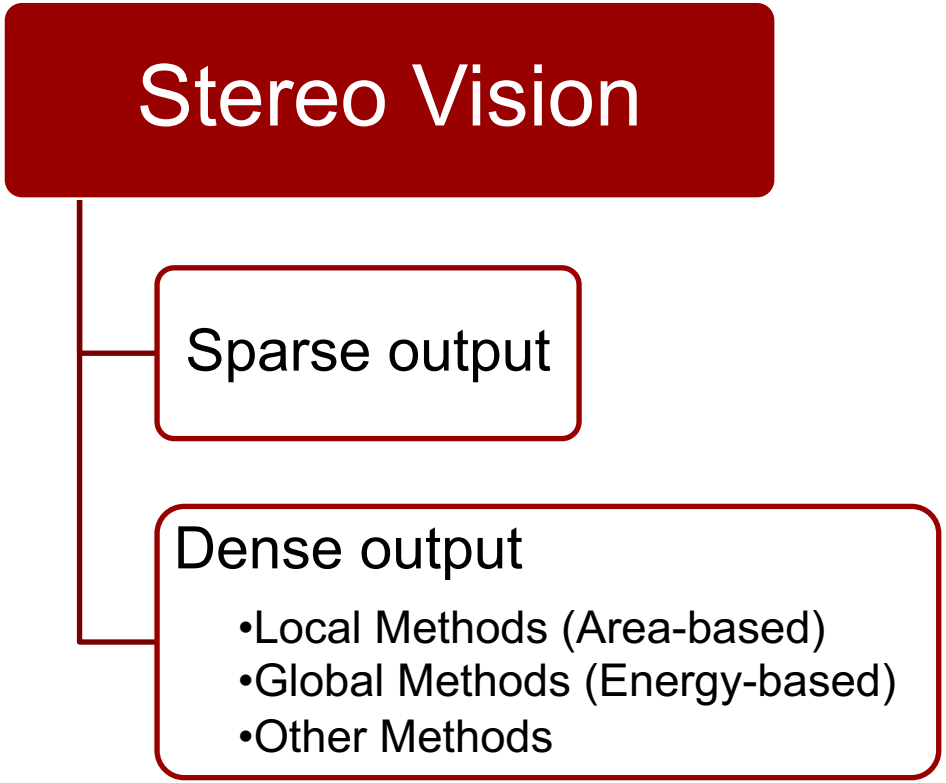
$$Z = f \frac{B}{X_L - X_R}$$

Disparity

Summary

- We discussed about what Stereo Vision is.

- We learned about :
  - Stereo/Epipolar Geometry
  - Rectified Stereo Case
  - Depth from Stereo Matches
- Correspondence Problem
  - **Dense** vs Sparse Correspondence
  - **Local** vs Global Correspondence
  - (Dis)-Similarity Measures



Lazaros Nalpantidis

Stereo Vision

DTU Electrical Engineering