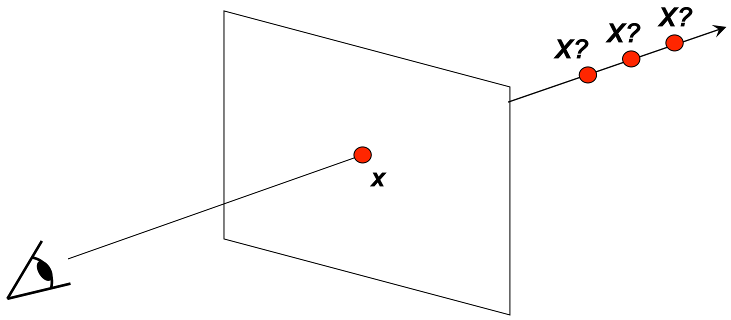
Stereo basics and Epipolar Geometry (Part 1)

Lecture's goals

- Stereo Analysis
  - Simple stereo geometry
  - Correspondence problem
  - Correlation-based stereo

Our Goal: Recovery of 3D Structure

- We will focus on perspective and motion
- We need *multi-view geometry* because recovery of structure from one image is inherently ambiguous



Slide credit: Svetlana Lazebnik

3

To Illustrate This Point...

- Structure and depth are inherently ambiguous from single views.

Slide credit: Svetlana Lazebnik

4

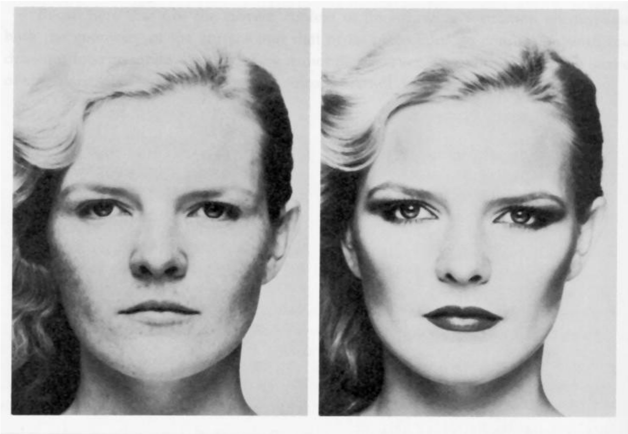
Recovering 3D from images

- Goal: Recovery of 3D structure
  - o What cues in the image allow us to do that?

5

Visual cues

- Shading



Merle Norman Cosmetics, Los Angeles

Slide credit: Steve Seitz

6

Visual cues

- Shading
- Texture

The Visual Cliff, by William Vandivert, 1960

Slide credit: Steve Seitz

7

Visual cues

- Shading
- Texture
- Focus

From *The Art of Photography*, Canon

Slide credit: Steve Seitz

8

Visual Cues

- Shading
- Texture
- Focus
- Perspective

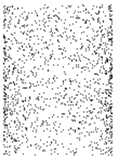
Slide credit: Steve Seitz

9

Visual Cues

- Shading
- Texture
- Focus
- Perspective
- Motion

Figures from L. Zhang



Slide credit: Steve Seitz

10

http://www.brainconnection.com/teasers/?main=illusion/motion-shape

What Is Stereo Vision?

- Generic problem formulation: given several images of the same object or scene, compute a representation of its 3D shape

Slide credit: Svetlana Lazebnik

11

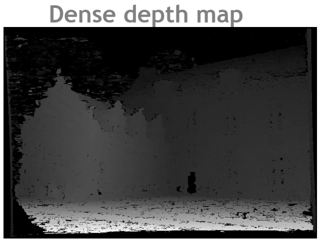
What Is Stereo Vision?

- Narrower formulation: given a calibrated binocular stereo pair, fuse it to produce a depth image

Image 1

Image 2

Dense depth map



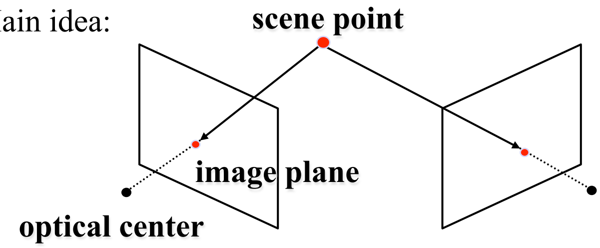
Slide credit: Svetlana Lazebnik

13

Estimating scene shape

- **Stereo**:
  - o infer 3d shape of scene from two (multiple) images from different viewpoints

Main idea:

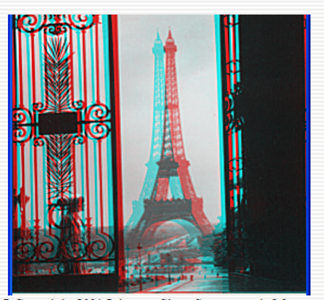


Stereo photography and stereo viewers

Take two pictures of the same subject from two slightly different viewpoints and display so that each eye sees only one of the images.

Invented by Sir Charles Wheatstone, 1838

Image from fisher-price.com

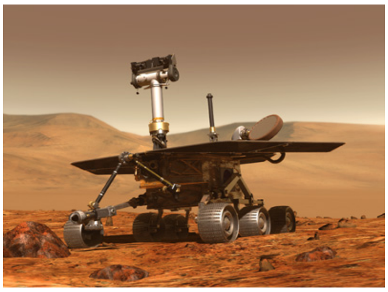


http://www.johnsonshawmuseum.org

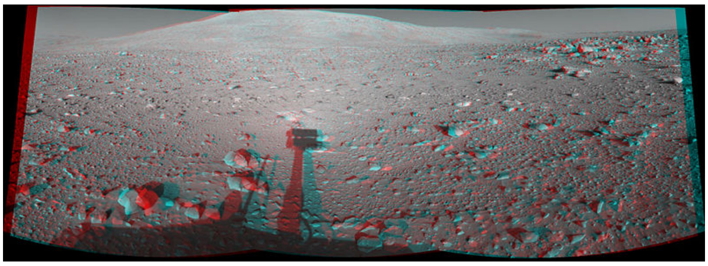
Public Library, Stereoscopic Looking Room, Chicago, by Phillips, 1923

Application of Stereo: Robotic Exploration

Nomad robot searches for meteorites in Antartica



Real-time stereo on Mars



Steve Seitz

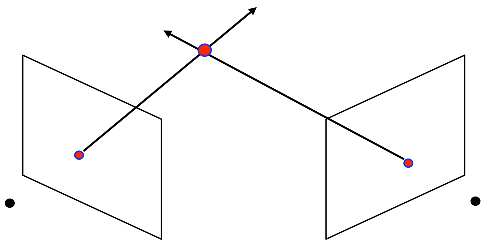
20

Topics of This Lecture

- **Geometric vision**
  - o Visual cues
  - o Stereo vision
- **Epipolar geometry**
  - o Depth with stereo
  - o Geometry for a simple stereo system
  - o Case example with parallel optical axes
  - o General case with calibrated cameras
- **Stereopsis & 3D Reconstruction**
  - o Correspondence search
  - o Additional correspondence constraints
  - o Possible sources of error
  - o Applications

21

Depth with Stereo: Basic Idea



- Basic Principle: Triangulation
  - o Gives reconstruction as intersection of two rays
  - o Requires
    - Camera pose (calibration)
    - Point correspondence

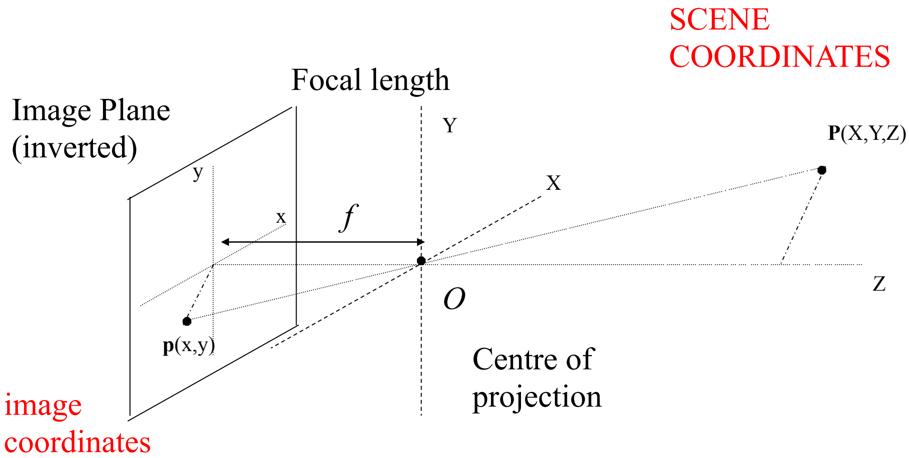
22

Geometry for a simple stereo system

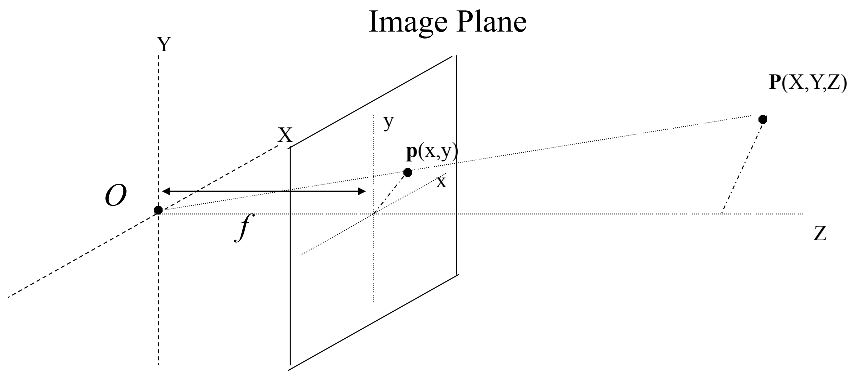
23

Camera Model

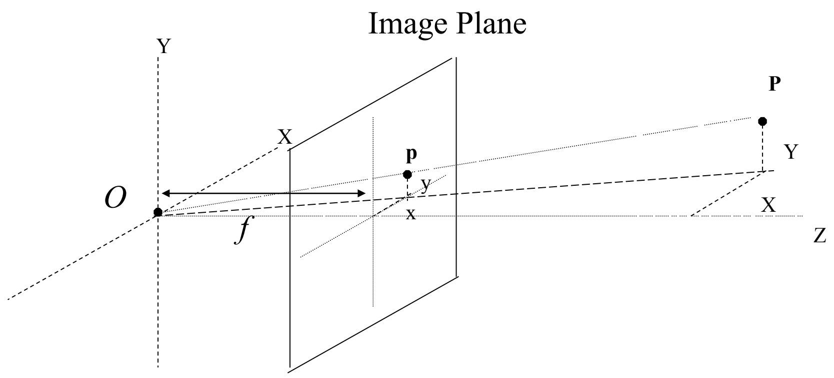
- Pinhole camera
- Perspective transform



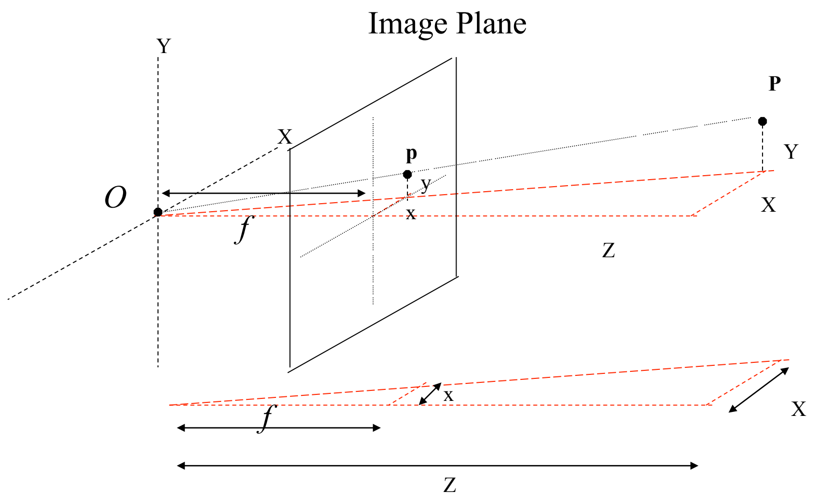
Pinhole Camera



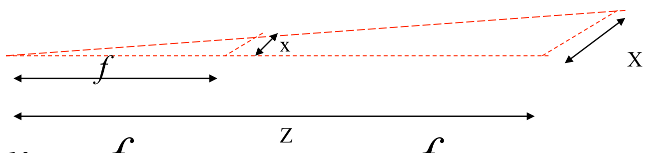
Pinhole Camera



Perspective Camera



Perspective Projection

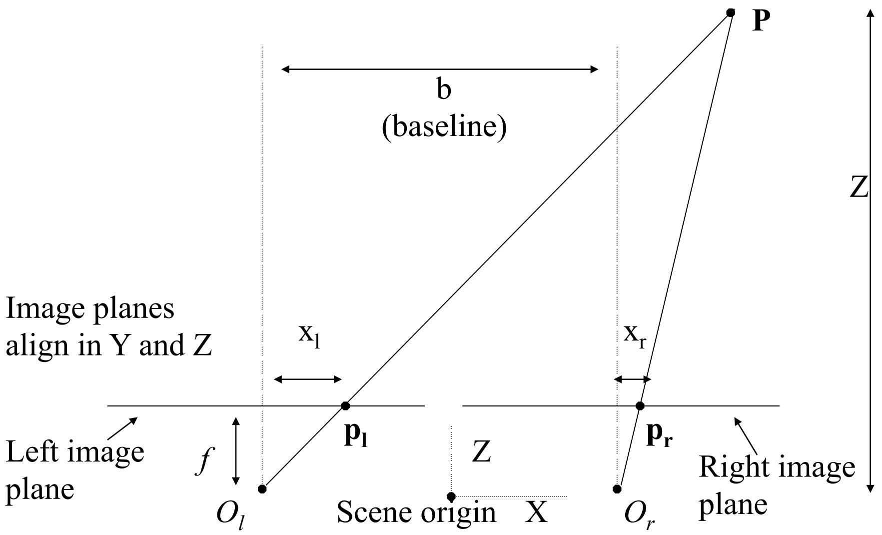


Similar triangles $$\frac{x}{X} = \frac{f}{Z}$$ $$x = \frac{f}{Z}X$$

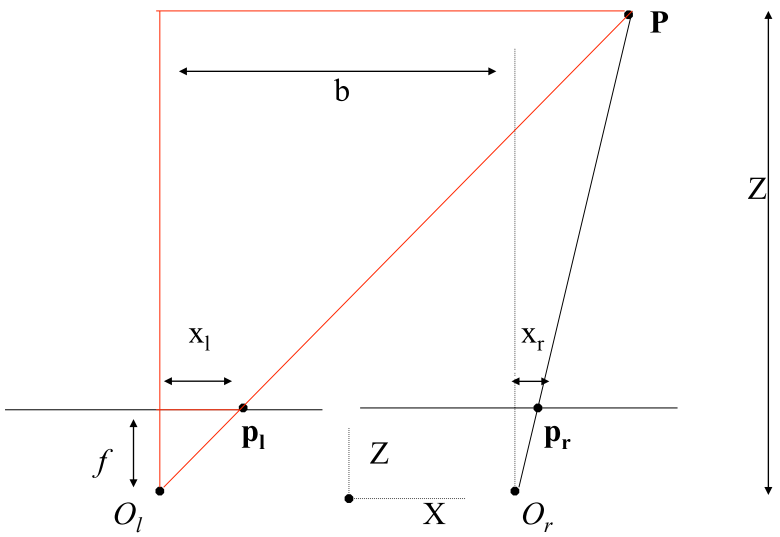
Similarly $$y = \frac{f}{Z}Y$$

Perspective transform

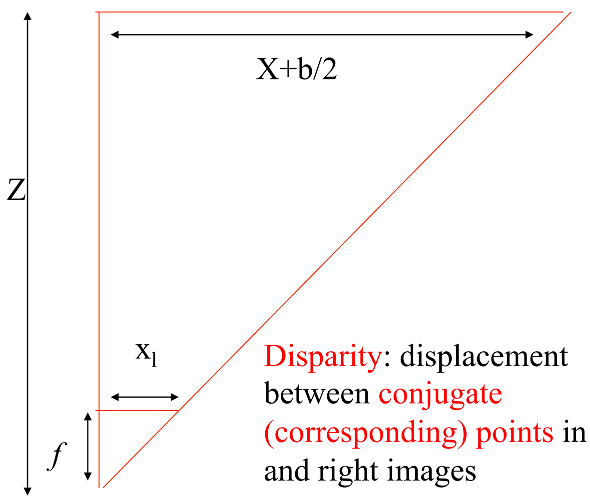
A simple Stereo System



A simple Stereo System



Depth by triangulation



Disparity: displacement between conjugate (corresponding) points in left and right images

$$\frac{x_l}{f} = \frac{X + b/2}{Z}$$

Similar triangles

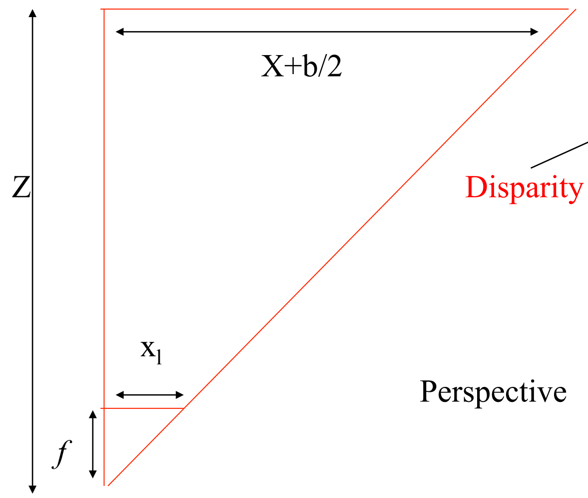
Similarly

$$\frac{x_r}{f} = \frac{X - b/2}{Z}$$

$$\frac{x_l - x_r}{f} = \frac{b}{Z}$$

$$Z = \frac{bf}{\underbrace{x_l - x_r}_{\text{disparity}}}$$

Scene Coordinates



$$Z = \frac{bf}{x_l - x_r}$$

$$\frac{x_l}{f} = \frac{X + b/2}{Z}$$

Perspective

$$\begin{cases} X = Z\dfrac{x_l}{f} - b/2 \\ Y = Z\dfrac{y_l}{f} \end{cases}$$

Depth from disparity

image I(x,y)

Disparity map D(x,y)

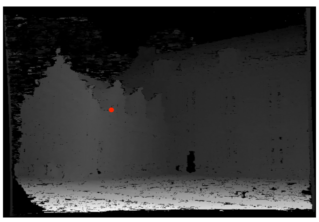


image I′(x′,y′)

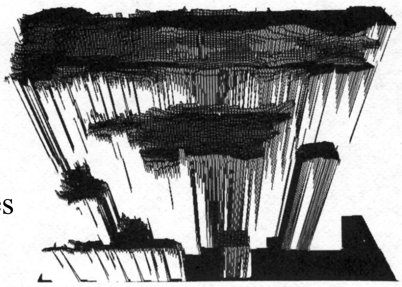
(x′,y′)=(x+D(x,y), y)

So if we could find the **corresponding points** in two images, we could **estimate relative depth**…

Stereo from Mars

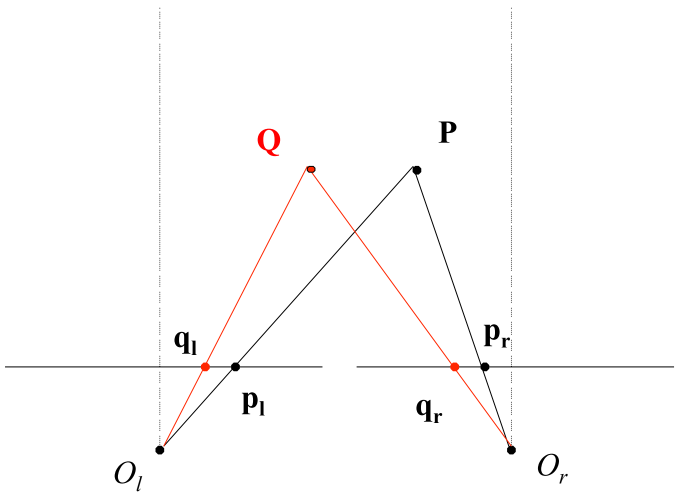
Stereo views from
Viking lander

Disparity map (interpolated)



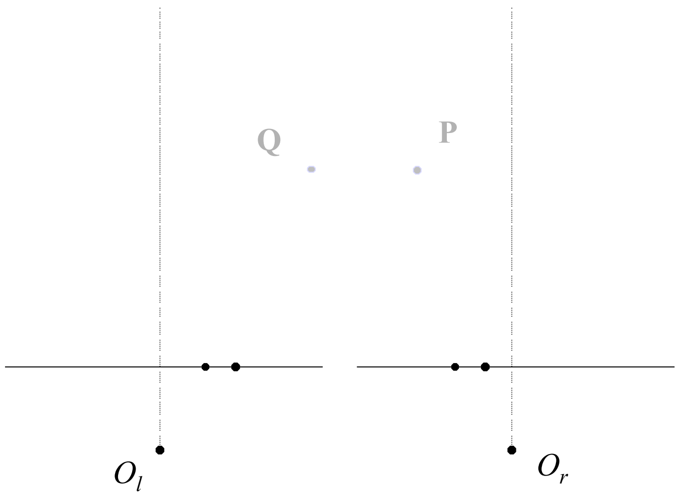
Despite the appearance of flat rocky terrain, disparity maps show a series of planes separated by gaps. This is indicative of a series of dunes Receding from the viewer.

Reconstructing from Image Points

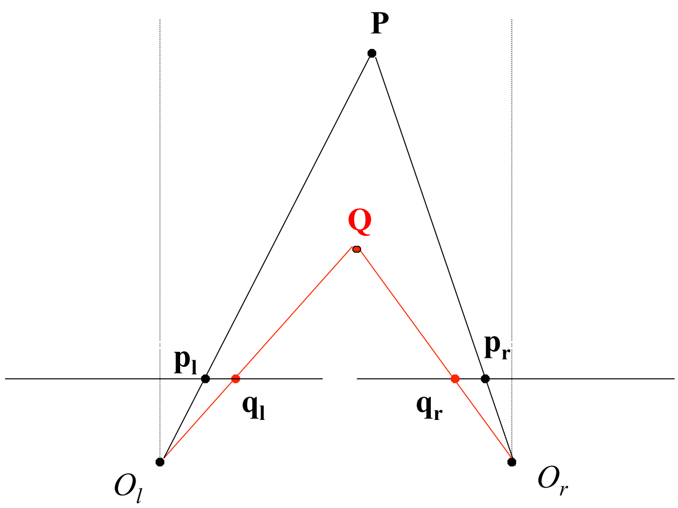


Why is this impossible?

Because we don’t know how to make the correspondences.



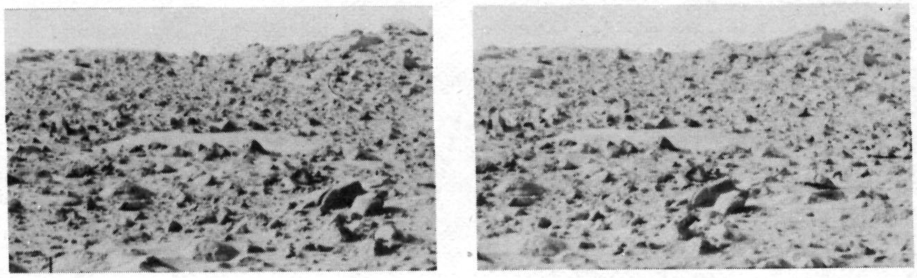
The Correspondence Problem



Components of Stereo Analysis

- Find correspondences
  - Conjugate pairs of points
  - Potentially hard – lots of pairs
- Reconstruction
  - Calculate scene coordinates (X,Y,Z)
  - Easy, once you have done…
- Calibration
  - Calculate parameters of cameras (e.g. b, f, …)

Stereo from Mars



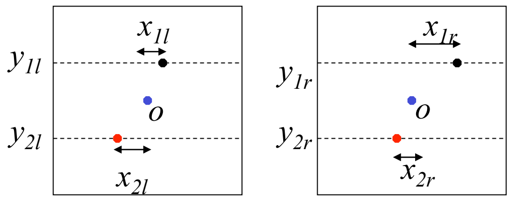
Stereo views from
Viking lander

Finding Correspondences

- Difficult to find correspondences at every point
  - Sparse set of points
- Assume most scene points are visible in both views
- Assume corresponding points are similar
- Find correspondences at "interesting" points
  - Need recognisable image structure.

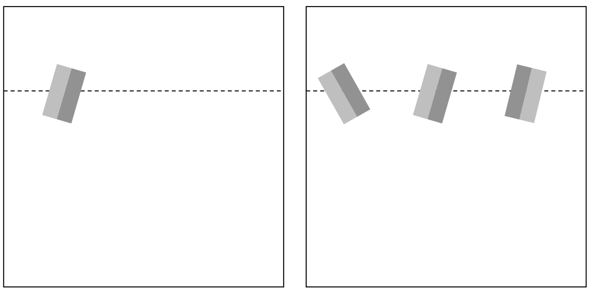
The Epipolar Constraint
(will talk in detail about this later)

- Potentially large search space for each candidate point.
- Match for a given (x_l, y_l) lies on the line $y_r = y_l$
    - o At least for our simple system
    - o 1D search problem
- Epipolar lines



Matching Edges

- Edges are convenient places to match
    - o Correspond to significant structure
    - o Small number of points to match
    - o Edge polarity and direction provide cues for matching

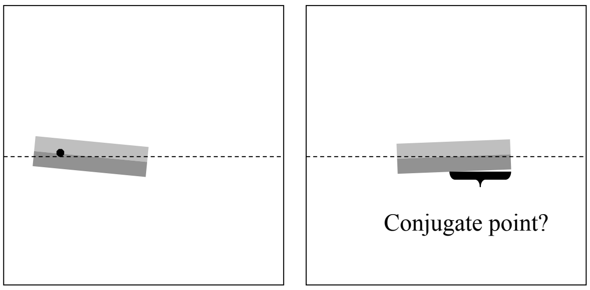


Edge Detection

- Canny Detector
  - o Multiple scale (coarse – fine search)
    - Detect edges at coarse scales first
    - Smaller number: fewer false matches
    - Starting point for search at finer scales.
  - o Accuracy
    - Non-maximal suppression can locate edges to sub-pixel precision.
    - High accuracy is important for accurate depth estimation

Matching Edges

- Image gradients at corresponding points may not be equally high
  - o Shadows, occlusions, illumination differences
- Horizontal edges are difficult to match
  - o Match points are poorly localised along epipolar lines



Edges are useful candidates for correspondence matching because...

- They represent significant image structure
- There aren't usually too many of them (combinatorics)
- We can use image features (polarity, direction) to verify matches
- We can locate them accurately (Canny – sub-pixel localisation)
- Multi-scale location (coarse to fine search)

Disadvantages of using edges for correspondence matching...

- Not all significant structure lies on edges
- Edge magnitude features may not be reliable for matching
- Near-horizontal edges do not provide good localisation.

Correspondence by Correlation Matching

- Similar points do not necessarily lie on salient edges
- Corresponding points should look similar
  - o Not identical – different viewpoints
- Cross-correlation search to find candidate matches
  - o 1-D search along epipolar lines
- Which points to match?
- "Interesting" points (see previous lectures !)

Interest Operators

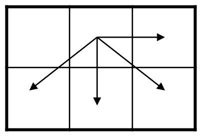
- We need *locally* distinct points.
- Edge matches could be obtained at neighbouring points along an edge
  - o Never find matches on near-horizontal edges. No disparity values here.
- "Interest" operators seek isolated distinct points
- Moravec operator
- Corners (Harris) (already covered - see previous lectures)
- LoG (DoG) (already covered)

Moravec Operator

- Non-linear filter
- Over some neighbourhood (e.g. 5X5) calculate

$$\sum \left(I_{(i,j)} - I_{(i+1,j)}\right)^2 \quad \sum \left(I_{(i,j)} - I_{(i-1,j+1)}\right)^2$$

$$\sum \left(I_{(i,j)} - I_{(i+1,j+1)}\right)^2 \quad \sum \left(I_{(i,j)} - I_{(i,j+1)}\right)^2$$



Pixel differences

- Output value is *minimum* of these values
- Suppress non-maxima of the filter output
    - o Isolate local maxima to get distinct points
- Finds points where intensity is varying quickly
    - o Taking minimum eliminates edges as candidates

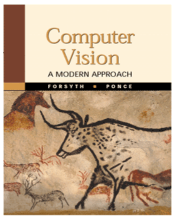
Example

Corner points detected on two views

References and Further Reading

- Background information on epipolar geometry and stereopsis can be found in Chapters 10.1-10.2 and 11.1-11.3 of

D. Forsyth, J. Ponce,
Computer Vision - A Modern Approach.
Prentice Hall, 2003



53