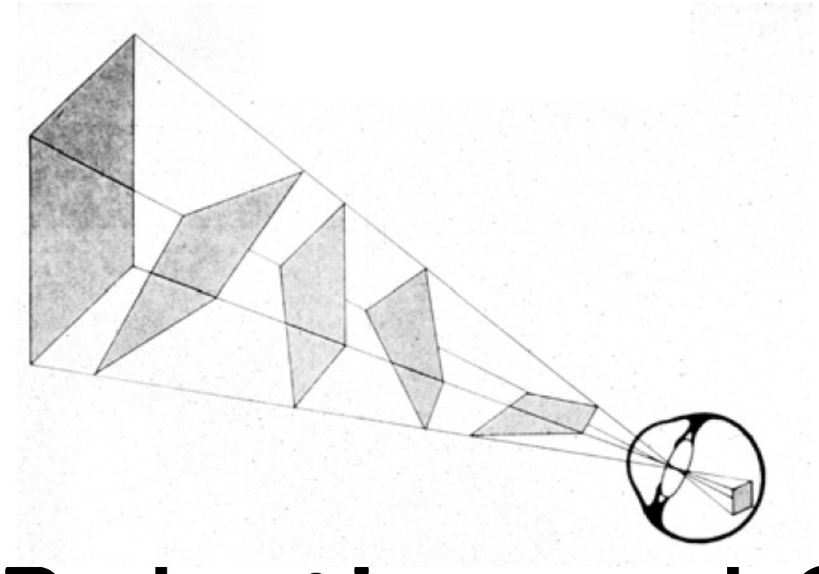


RoVi: Robotics and Computer Vision – Vision Introduction

Dirk Kraft

Program for Today

- Literature
- Course Structure
- Applications in Computer Vision
- Overview of course content
- Human Depth Perception
- Camera Geometry

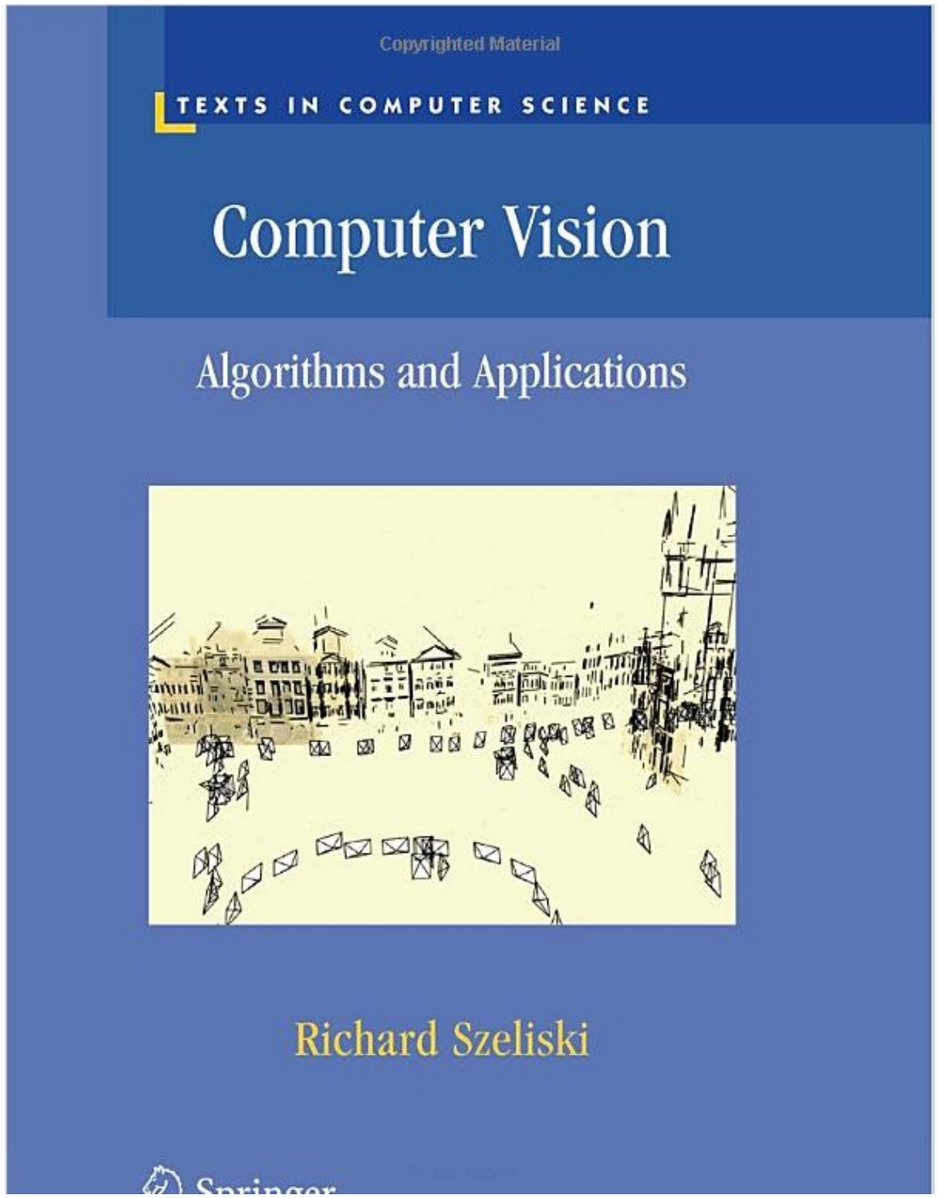
Program for Today

- Literature
- Course Structure
- Applications in Computer Vision
- Overview of course content
- Human Depth Perception
- Camera Geometry

Literature – General comments

- This is a master level course.
  - We expect you to read the stuff we tell you to read!
  - There will be holes that you need to fill yourself.
- One book not really covering all things well
- Different styles of books/material might be preferred by different people

Szeliski, Computer Vision: Algorithms and Applications, 2nd ed.



- We will use this as our main source
- The author offers a preprint pdf here: https://szeliski.org/Book/
- You have electronic access (via the SDU library) to pdfs of the final book's chapters.

Henrik Aanæs, Lecture Notes on Computer Vision

- Very good in the stuff he covers

Lecture Notes on Computer Vision

Henrik Aanæs
DTU informatics
e-mail: haa@imm.dtu.dk

January 30, 2015

Additional literature on itslearning

Files

Aanæs - 2015 - Lecture Notes on Computer Vision - Reduced.pdf

We will be using

- **R. Szeliski, Computer Vision: Algorithms and Applications, 2010, first edition. Free pre-print pdf here. Also available as an online resource from the SDU library.**

as our main textbooks and

- Henrik Aanæs, Lecture Notes on Computer Vision (see attached)

for some additional material.

OpenCV related books:

- Robert Laganière, OpenCV 2 computer vision application programming cookbook, 2011. The SDU library has two different online versions of this book available.
- Daniel Lélis Baggio et al., Mastering OpenCV with Practical Computer Vision Projects, 2012.

Other good books covering 3D comptuter vision:

- Richard Hartley and Andrew Zisserman, Multiple View Geometry in Computer Vision, 2004.
- Olivier Faugeras and Quang-Tuan Luong and Theo Papadopoulo, The Geometry of Multiple Images: The Laws That Govern the Formation of Multiple Images of a Scene and Some of Their Applications, 2004.
- Olivier Faugeras, Three-Dimensional Computer Vision, 1993.

Other good books covering image processing (or computer vision) are:

- M. Nixon and A. Aguado, Feature Extraction & Image Processing, 2nd edition, Academic Press, 2008.
- M. Sonka and V. Hlavac, Image Processing, Analysis, and Machine Vision, 3rd edition, Cengage Learning, 2007.
- L. G. Shapiro and G. C. Stockman, Computer Vision, Prentice Hall, 2001.
- D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*, Prentice Hall, 2002.

General image processing (or computer vision) online resources:

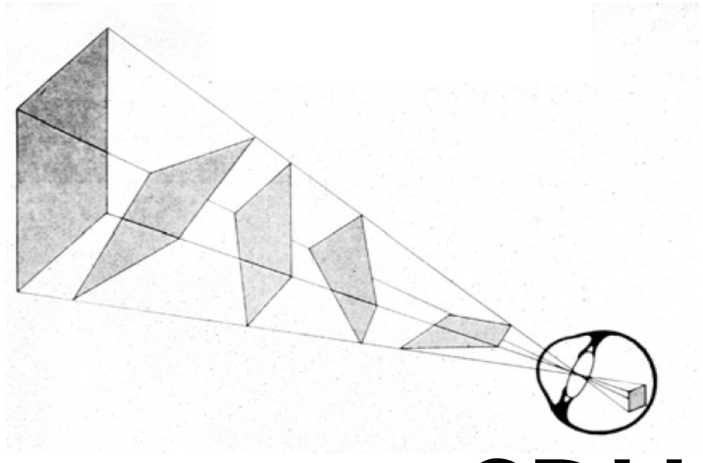
- CVonline
- The Computer Vision Homepage
- The Joy of Visual Perception: A Web Book

Program for Today

- Literature
- Course Structure
- Applications in Computer Vision
- Overview of course content
- Human Depth Perception
- Camera Geometry

Structure of Course

- Motivation
  - We live and act in a three-dimensional world.
  - One fundamental problem is that the visual information gathered by the eyes, or by cameras, is essentially two dimensional.
- Content
  - In this course we will study how to recover depth information.
  - We will also discuss how this visual information can be processed for the purpose of, e.g., pose and motion estimation.
  - We will hint at applications in
    - Robotics
      - Mobile Robotics
      - Grasping
    - Vision
      - Tracking
      - Reconstruction
      - Recognition
    - (Virtual Reality)



Course topics – vision
Subject to change

"Date"	Topic
Vision 1	Intro + Camera models + Calibration
Vision 2	Stereo + Triangulation + Epipolar Geometry
Vision 3	Matching + Dense Stereo
Vision 4	Other depth sensors + approaches
Vision 5	Point cloud processing
Vision 6	Uncertainty in Reconstruction (THMI)
Vision 7	Pose estimation, 3D -> 3D (FH)
Vision 8	Pose estimation, 2D -> 3D
Vision 9	Kalman filter
Vision 10	Particle filter
Subject to change – see itslearning plans under vision	

Program for Today

- Literature
- Course Structure
- Applications in Computer Vision
- Overview of course content
- Human Depth Perception
- Camera Geometry

Motivation

- Questions
  - How to recover three-dimensional information from images ?
  - How to use multiple views on the same scene ?
  - How to estimate motion and pose in space from visual data?
  - How to use this knowledge in applications?
- Motivation
  - We live and act in a three-dimensional world. How our visual perception provides us with such a rich, reliable and precise description of our environment has been the subject of important studies during the second half of the last century. One fundamental problem is that the visual information gathered by the eyes, or by cameras, is essentially two dimensional. In this course we will study how to recover depth information, focusing on multiple views, geometrical approaches. We will also discuss how this visual information can be processed for a variety of applications.

Microsoft Research

Search Microsoft Research

Videos | Projects | Publications | People | Downl

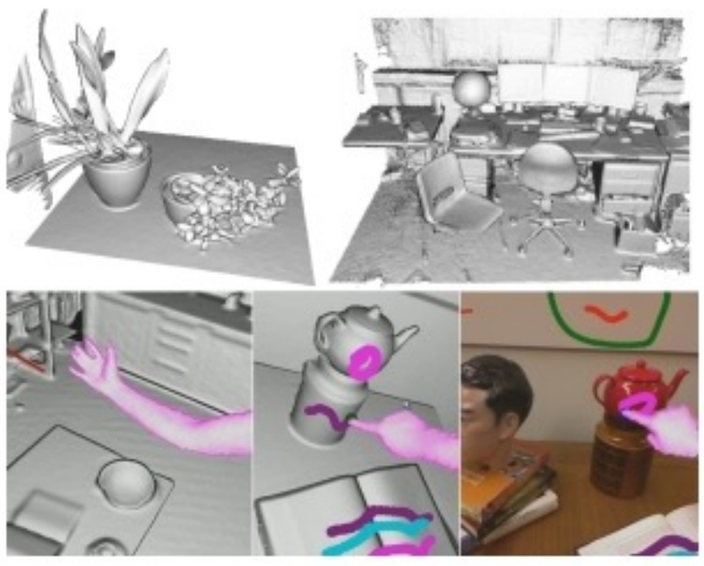
Home | Our Research | Connections | Careers | Hub

Worldwide Labs | Research Areas | Research Groups

Watch Rick Rashid's
TechFest Keynote

> Projects > 3D Surface Reconstruction

KinectFusion Project Page



This project investigates techniques to track the 6DOF position of handheld depth sensing cameras, such as Kinect, as they move through space and perform high quality 3D surface reconstructions for interaction. Other collaborators (missing from the list below): Richard Newcombe (Imperial College London); David Kim (Newcastle University & Microsoft Research); Andy Davison (Imperial College London)

Publications

- Shahram Izadi, David Kim, Otmar Hilliges, David Molyneaux, Richard Newcombe, Pushmeet Kohli, Jamie Shotton, Steve Hodges, Dustin Freeman, Andrew Davison, and Andrew Fitzgibbon, KinectFusion: Real-time 3D Reconstruction and Interaction Using a Moving Depth Camera, ACM Symposium on User Interface Software and Technology, October 2011
- Richard A. Newcombe, Shahram Izadi, Otmar Hilliges, David

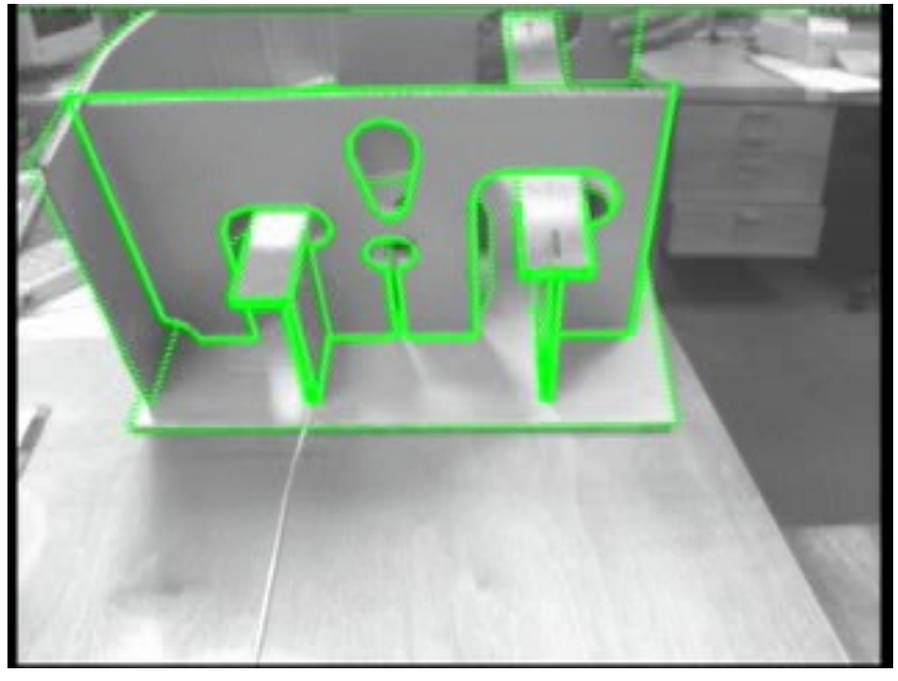
Videos

http://research.microsoft.com/en-us/projects/surfacerecon/

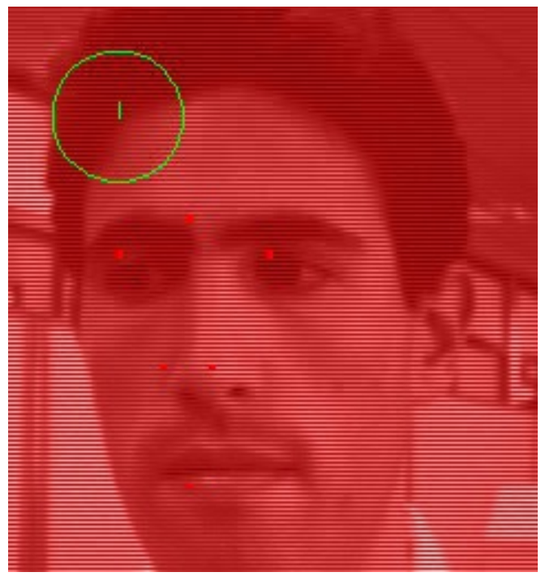
SDU ROBOTICS

Applications

- Pose Estimation using a CAD model



- Face Tracking



Grasping without Object Knowledge

Application II:
Complex Object Manipulation

Application III:
Reconstruction of Buildings

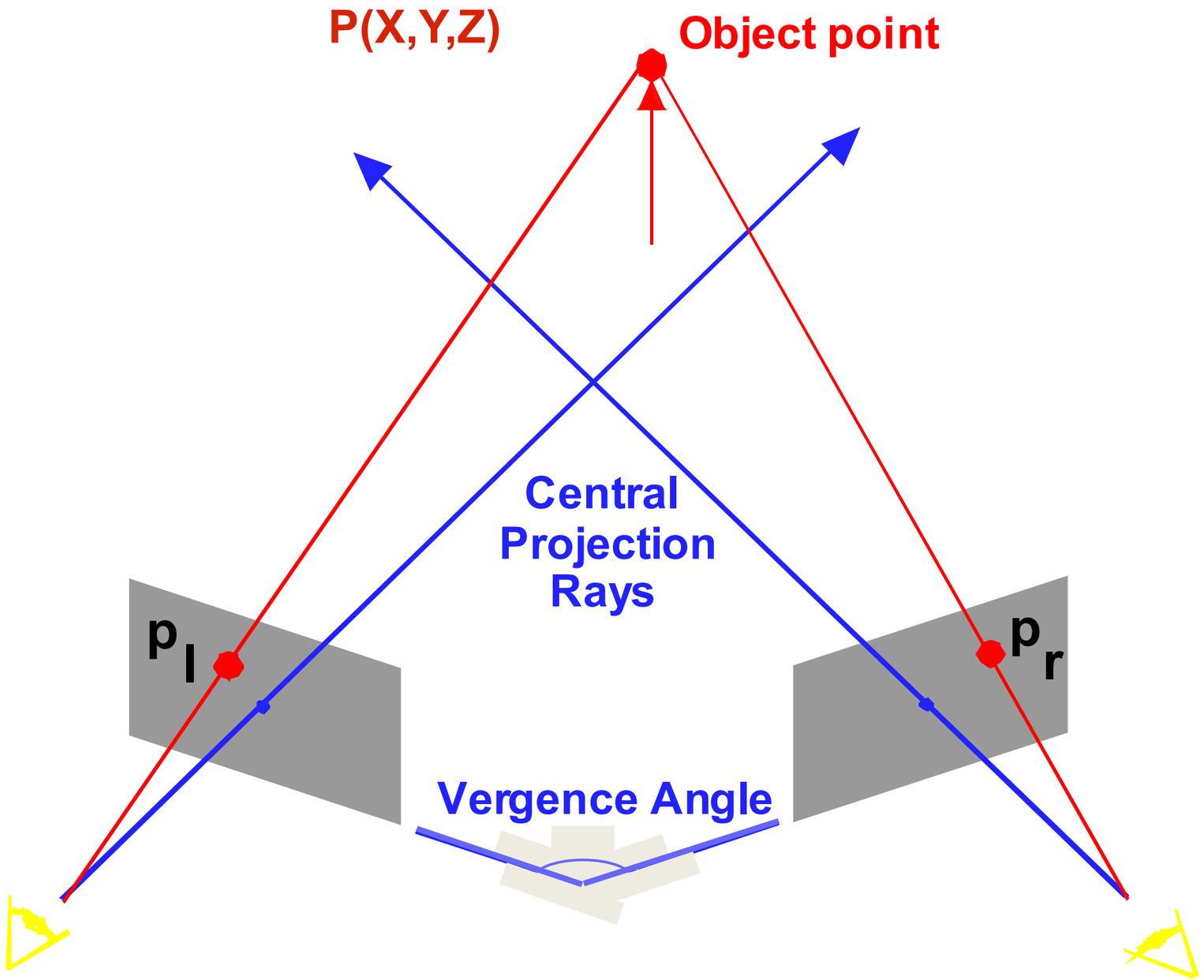
Applications 6: Virtual Reality merged with Reality

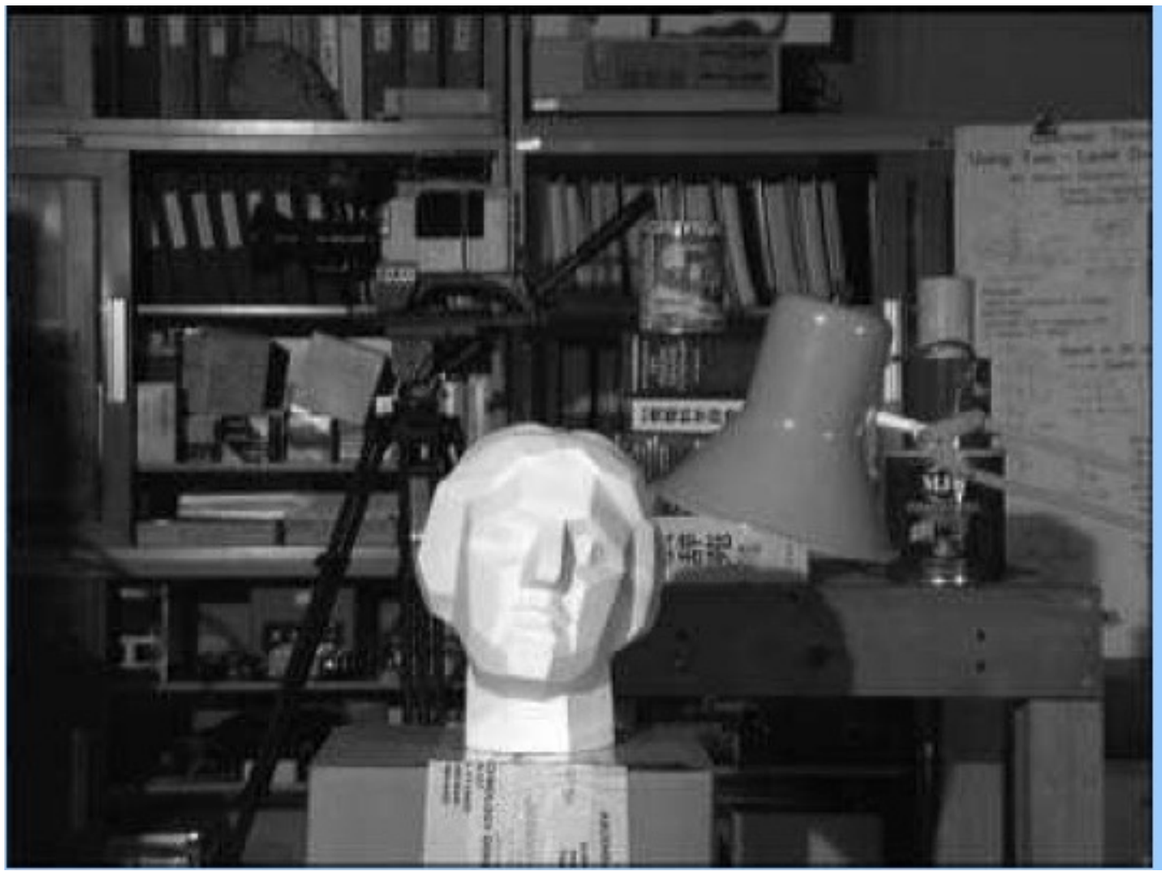
Program for Today

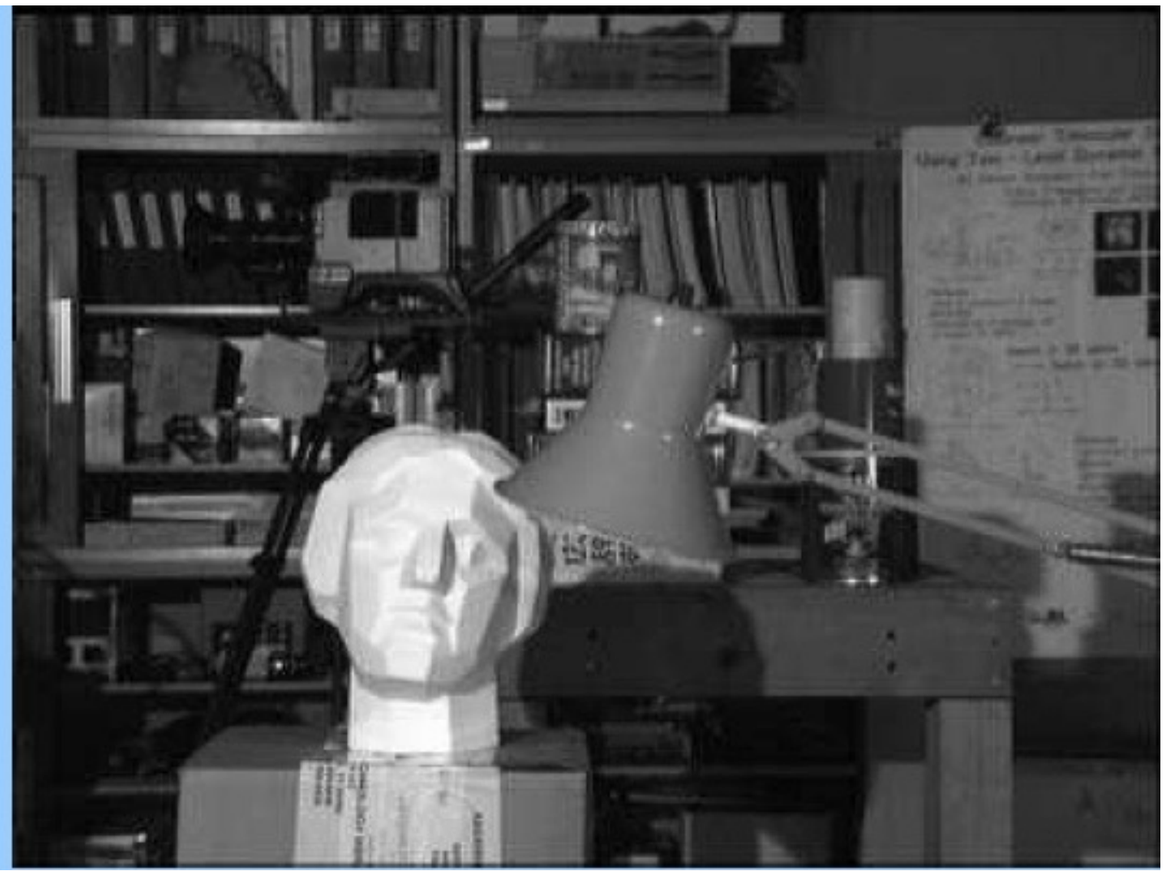
- Literature
- Course Structure
- Applications in Computer Vision
- Overview of course content
- Human Depth Perception
- Camera Geometry

Stereopsis

Stereopsis


P(X,Y,Z)
Object point
Central
Projection
Rays
p_l
p_r
Vergence Angle



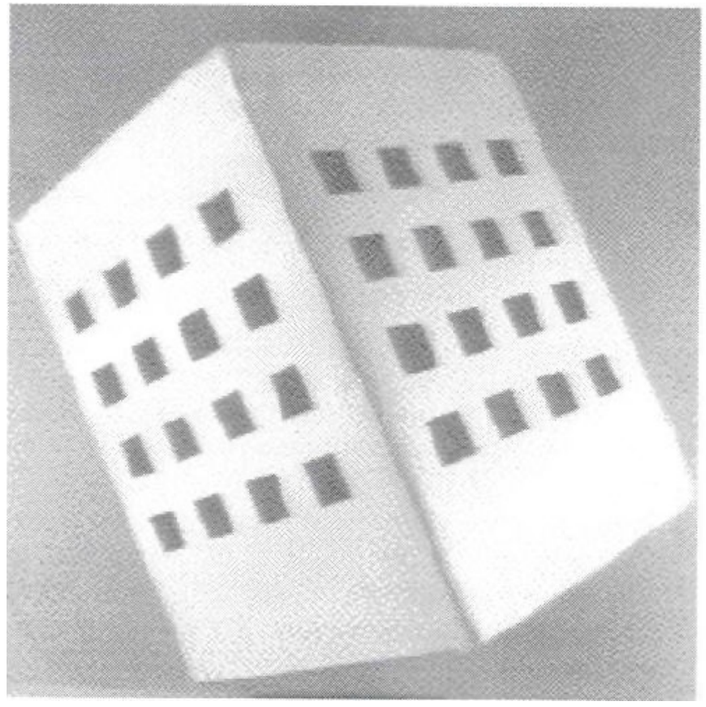


What are the Problems of Stereo

- What do you think?
  - Calibration
  - Matching
  - Reconstruction

Calibration

- Given one or more images of a calibration pattern,
- Estimate
  - The intrinsic parameters
  - The extrinsic parameters, or
  - BOTH

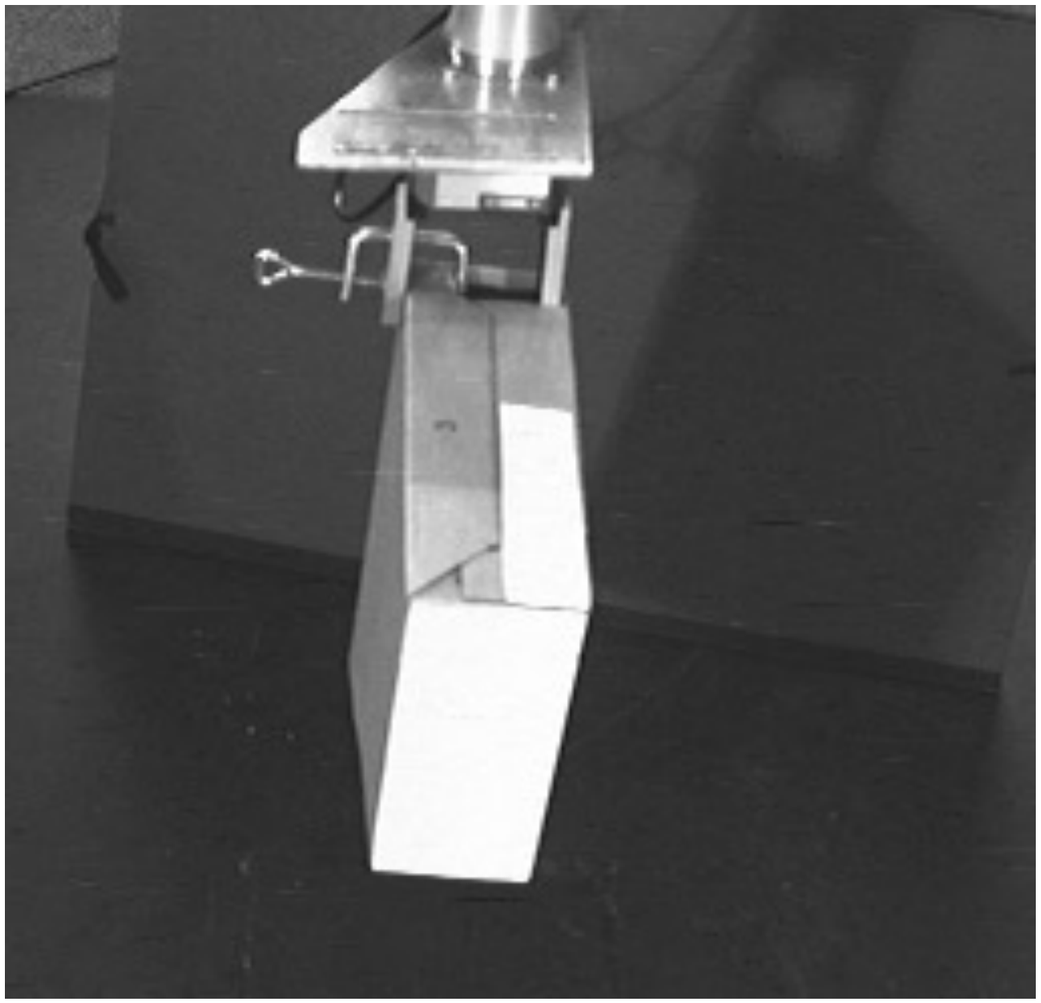


Matching

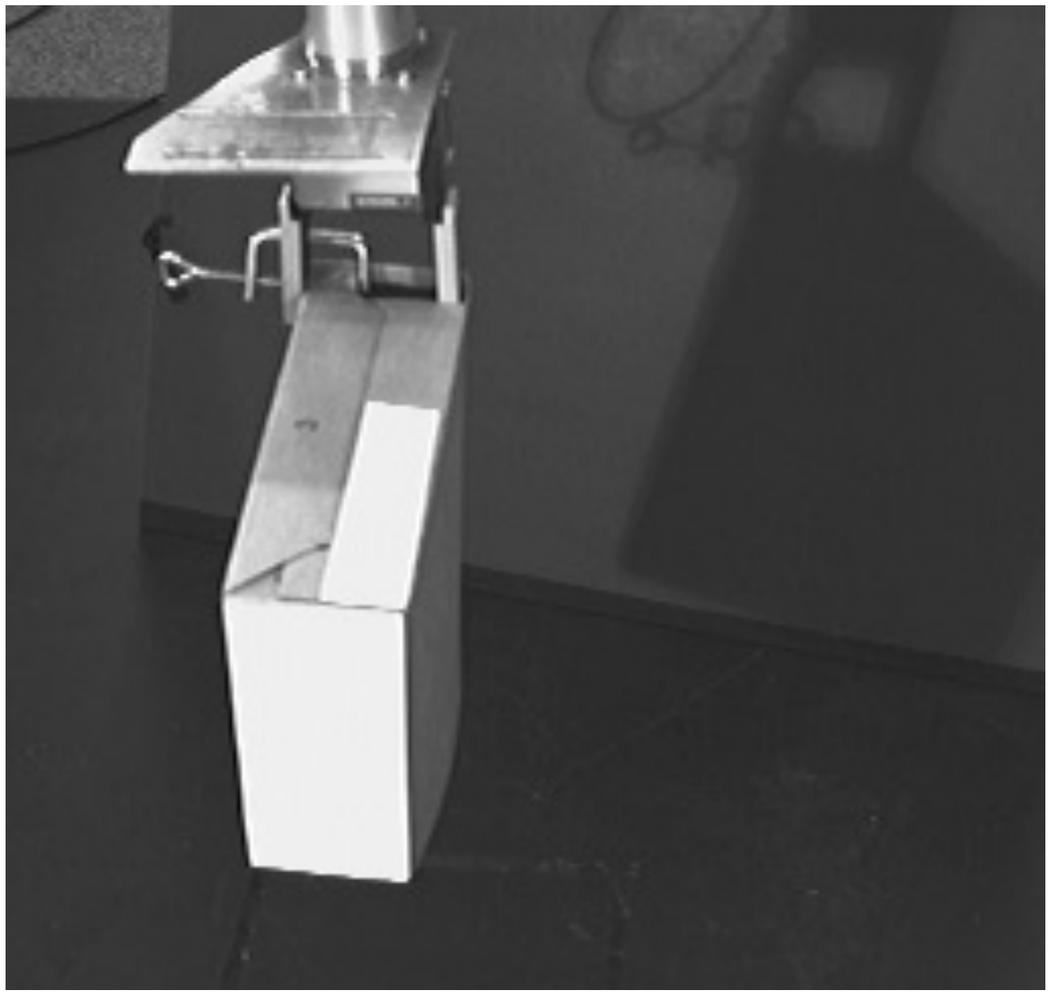
- Correspondence Problem
  - Define!
  - What is the problem?
  - How can things be made easier?

What are the involved problems?

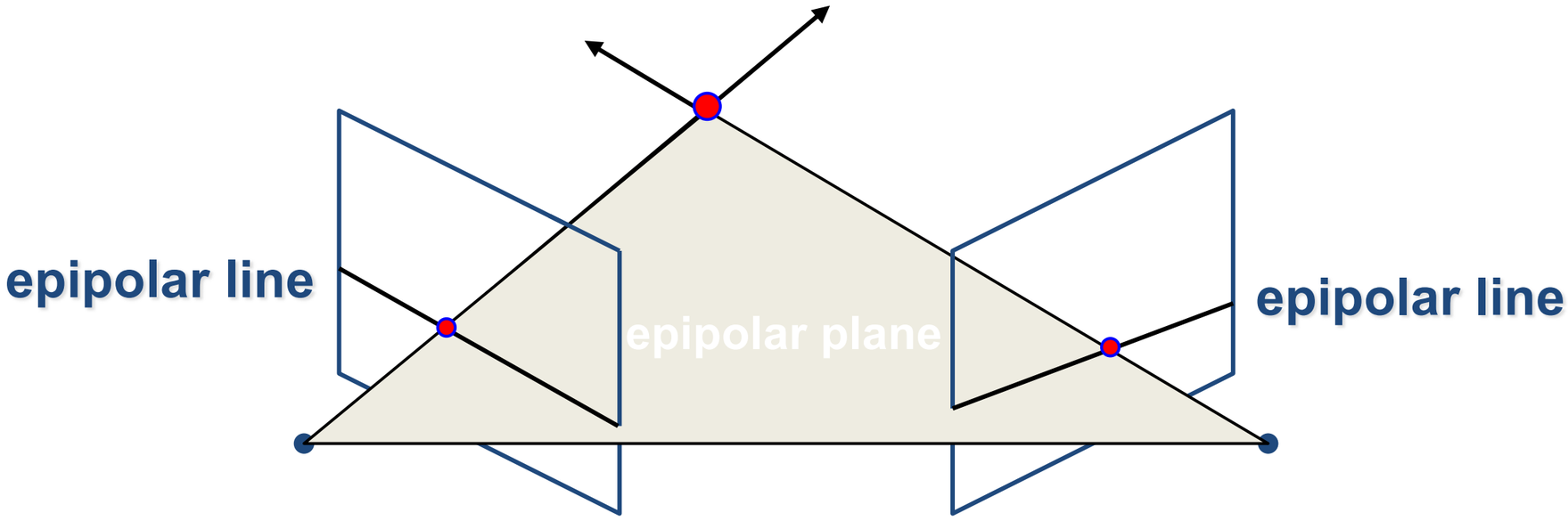
Left



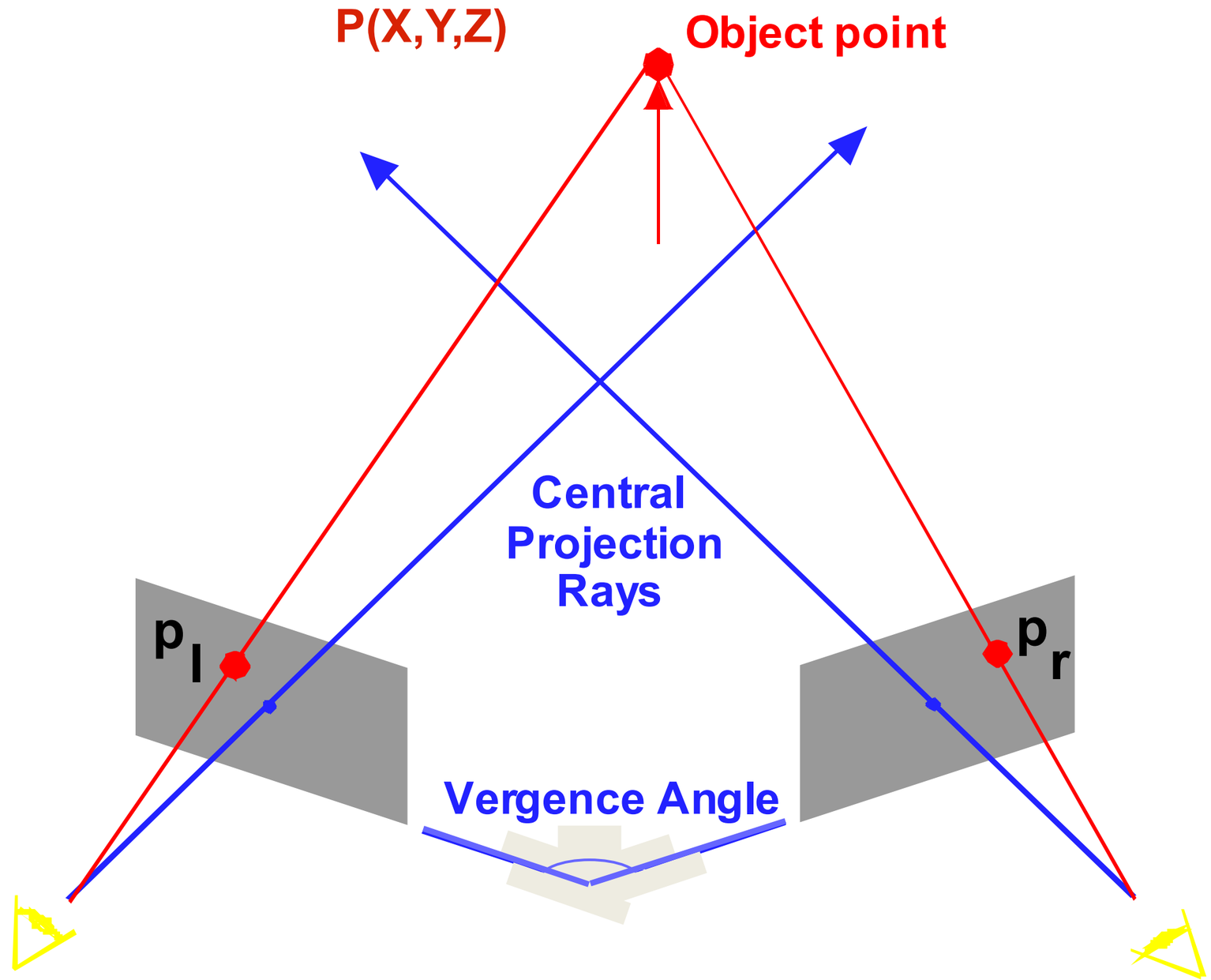
Right



Epipolar Geometry



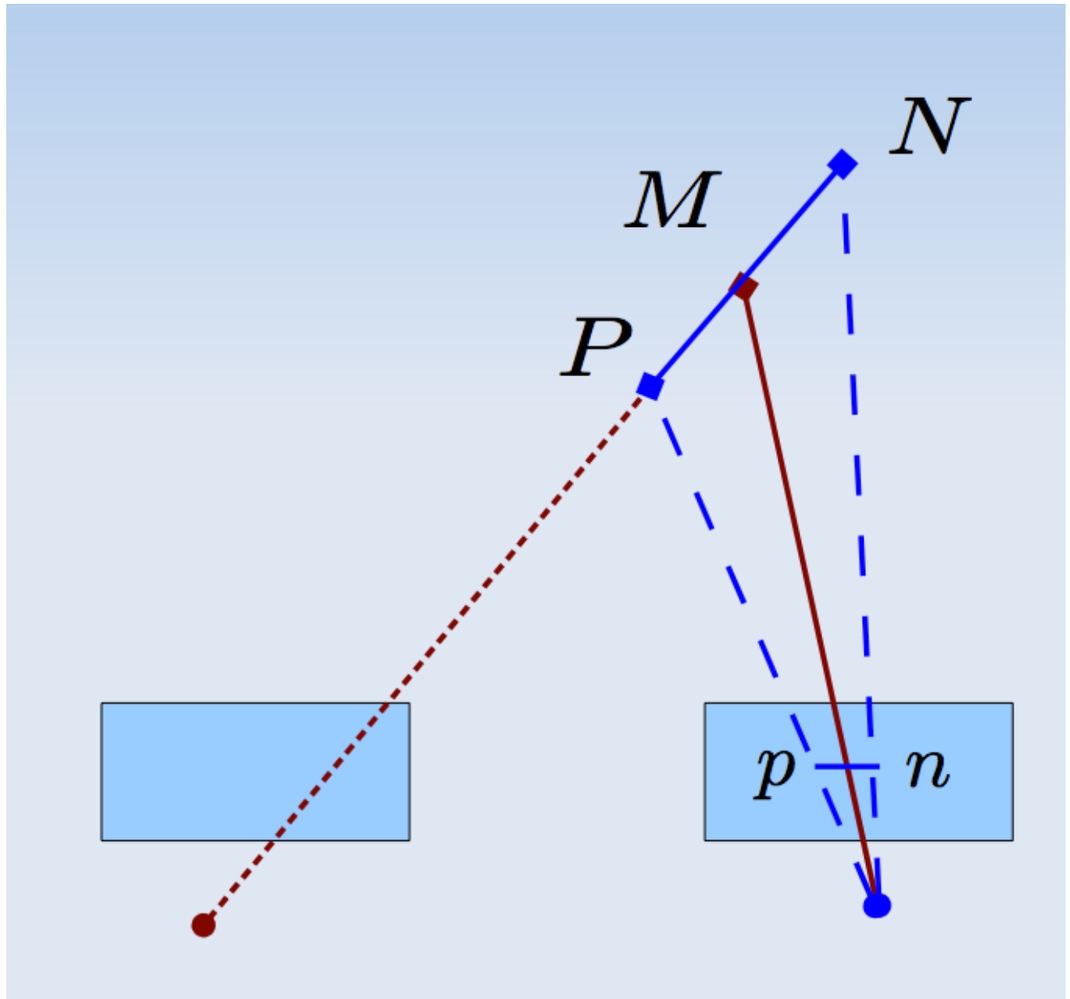
Reconstruction



What has been learned:

- You will
  - know how to calibrate a stereo rig and to interpret the projection matrices
  - be able to use the epipolar geometry to facilitate matching
  - be able to reconstruct from correspondences

Uncertainty in Reconstruction



Applications

- Pose estimation
- Tracking

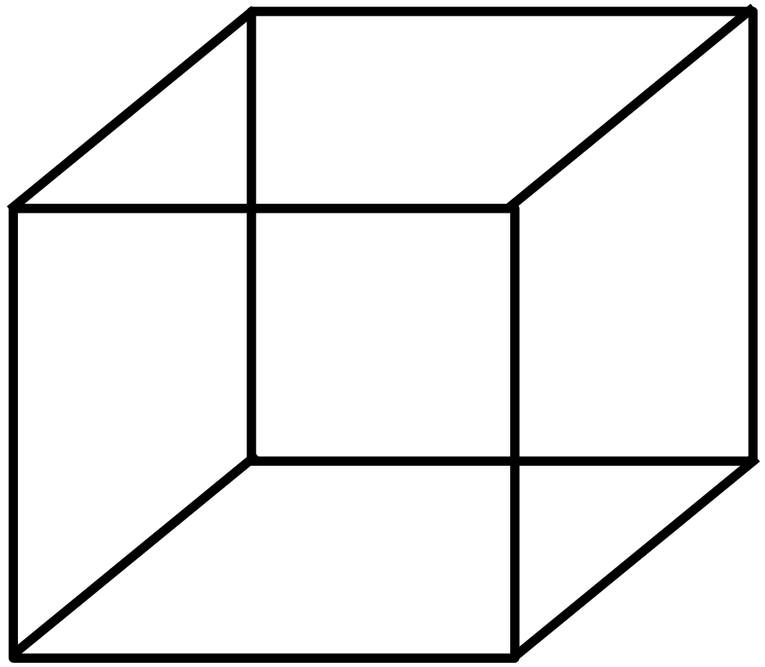
Program for Today

- Literature
- Course Structure
- Applications in Computer Vision
- Overview of course content
- Human Depth Perception
- Camera Geometry

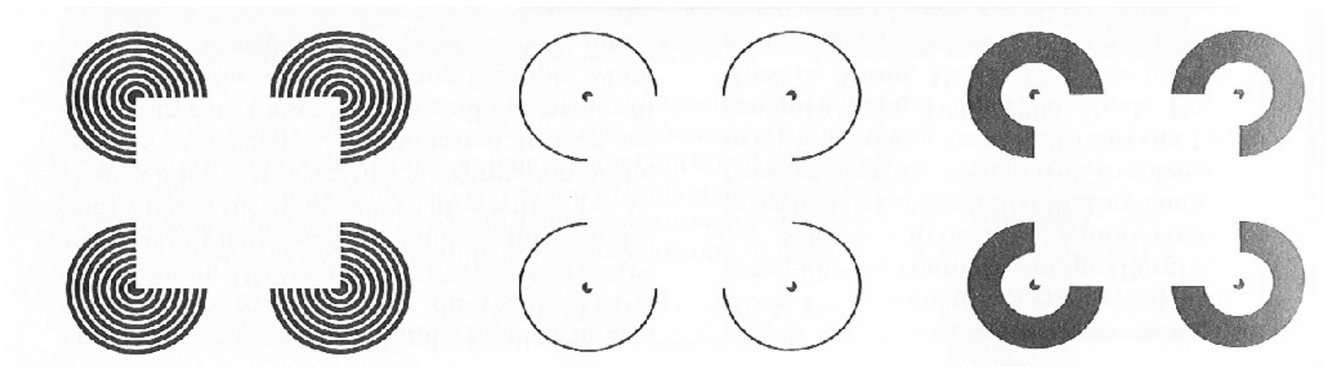
What happens next?

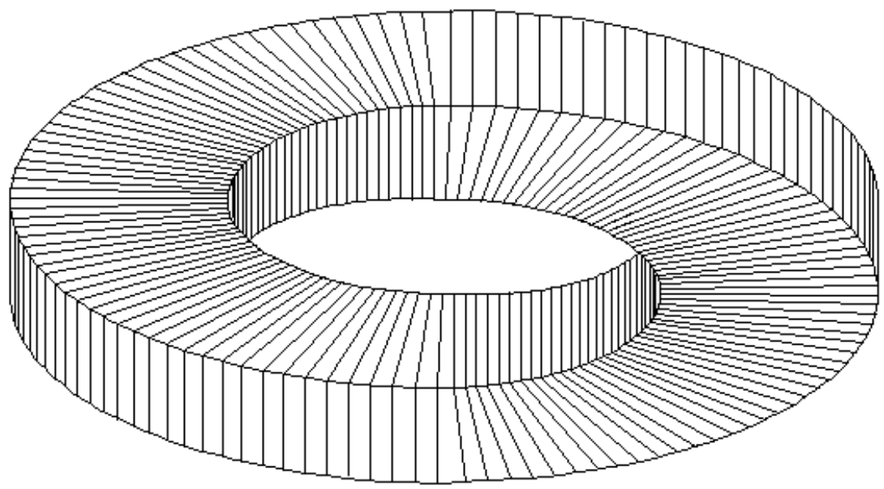
What really happens

Some Depth Illusions



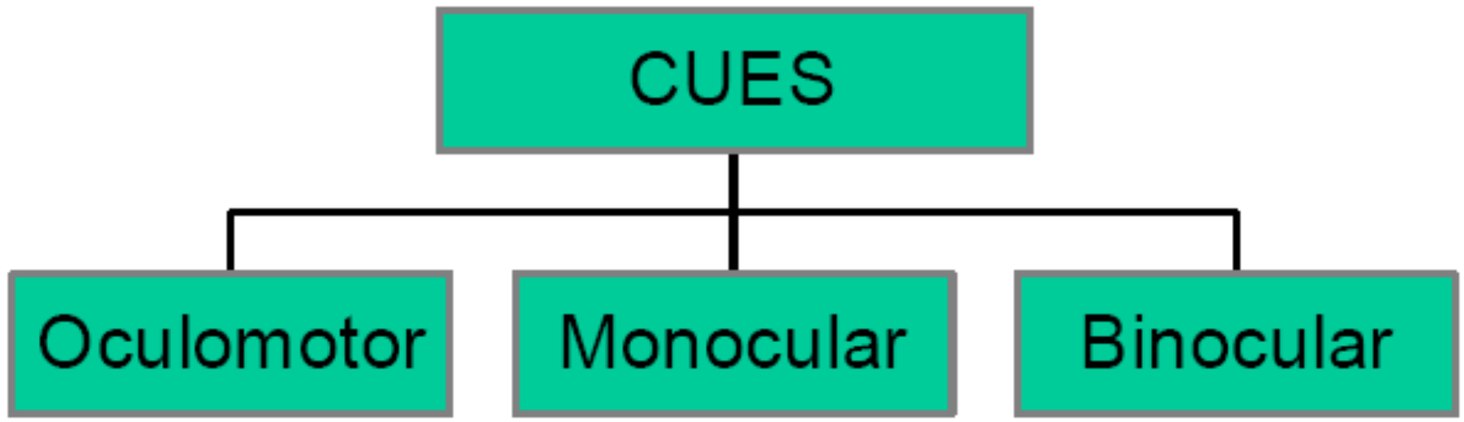
Necker Cube





DEPTH CUES

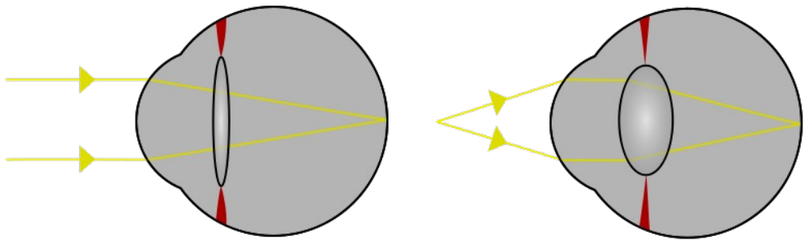
Albert Yonas: God must have loved depth cues because she made so many of them



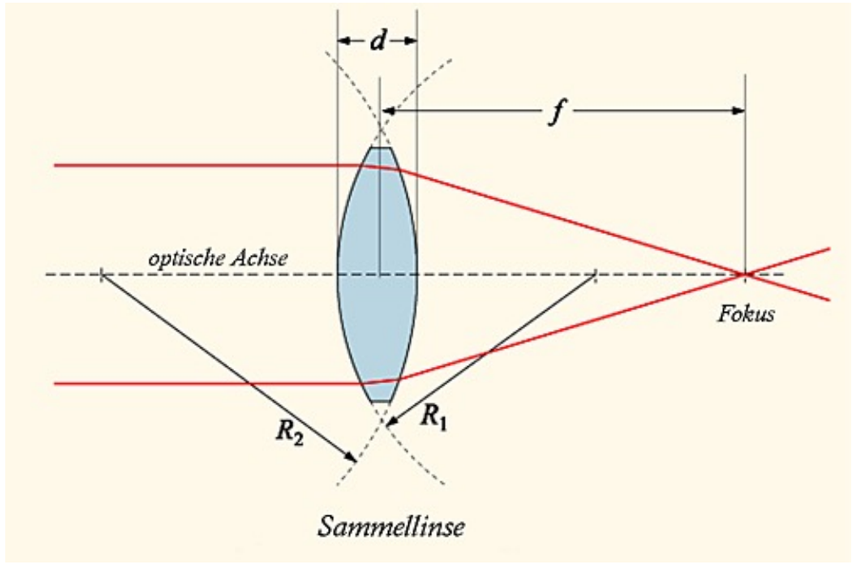
Oculumotor cues

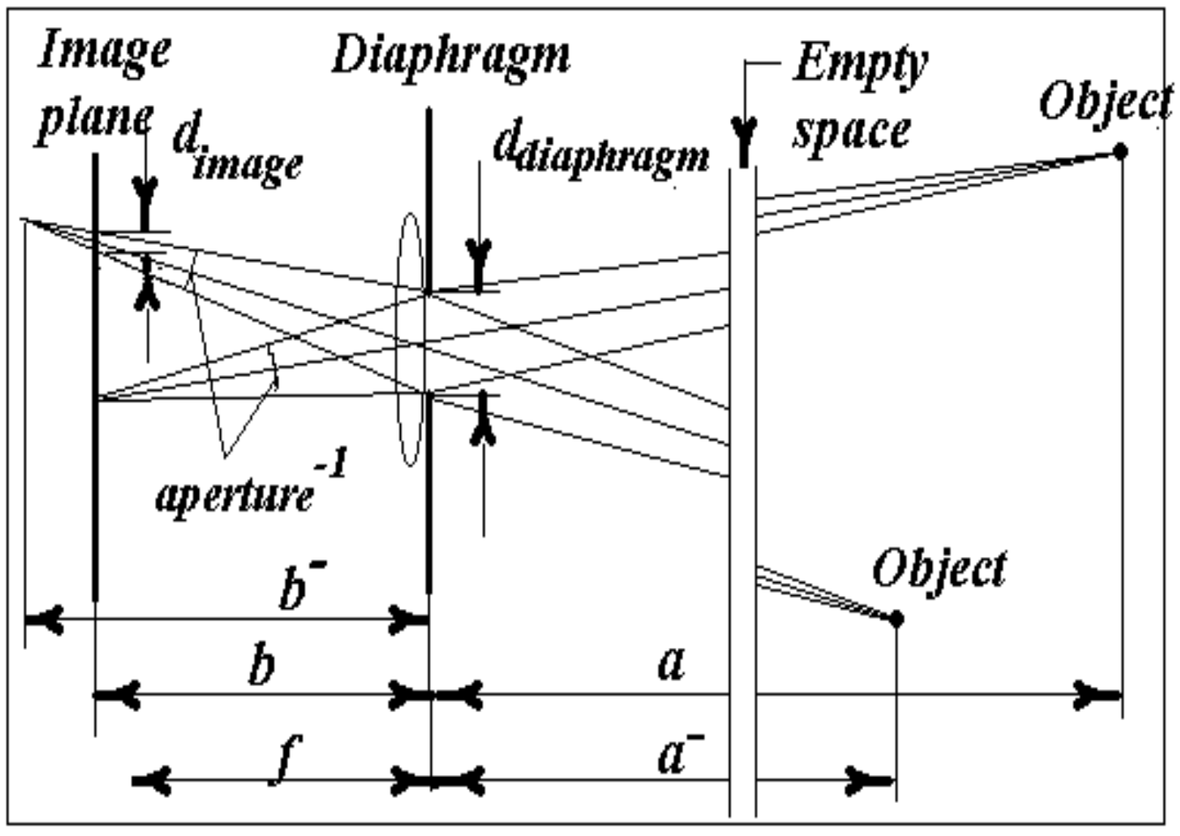
Accommodation

$$1/a+1/b=1/f$$

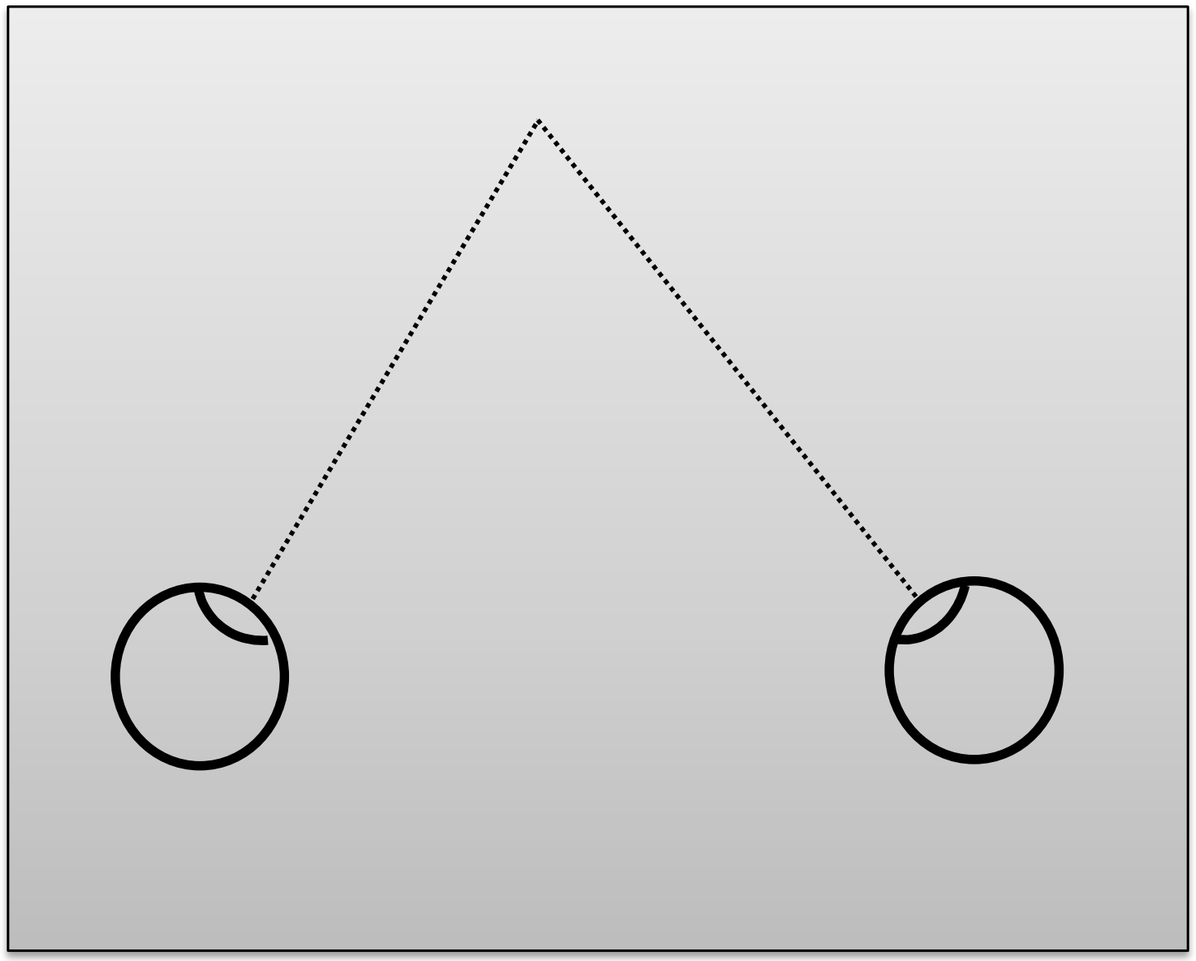


a=object distance
b=distance to image plane
f=focal length





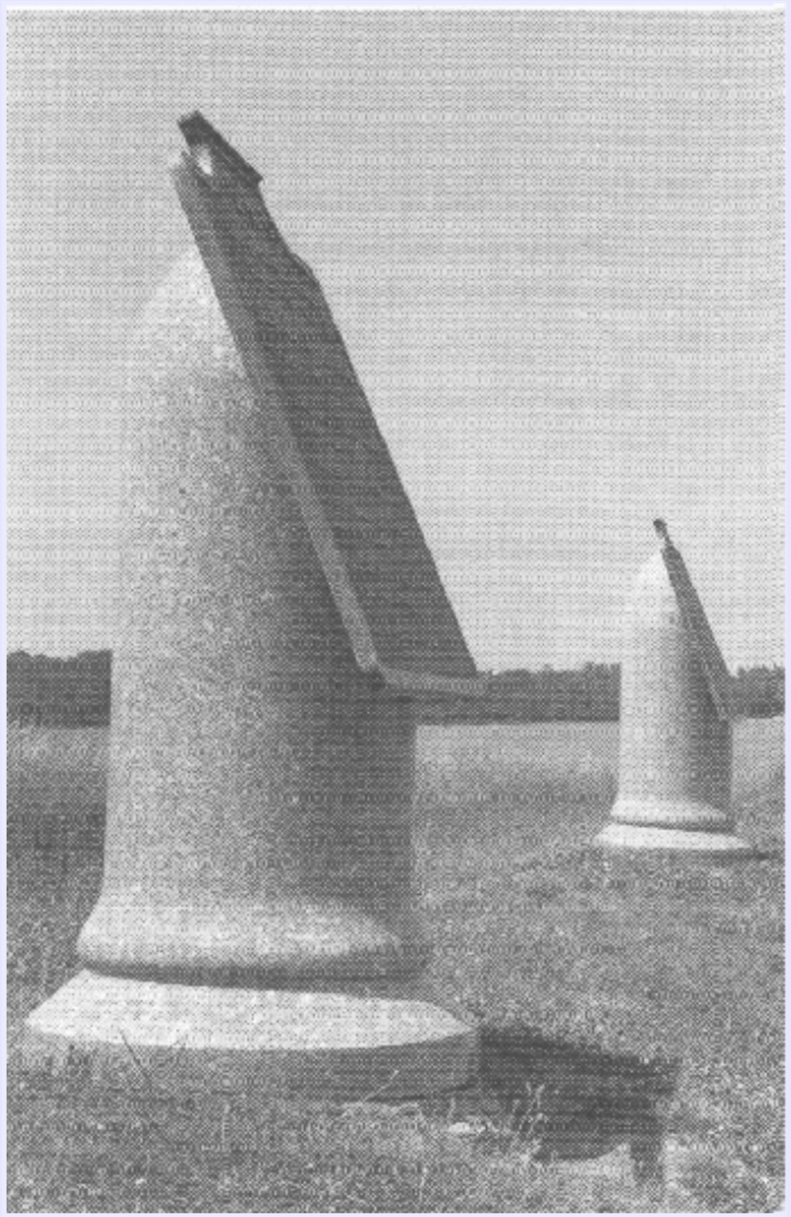
Convergence



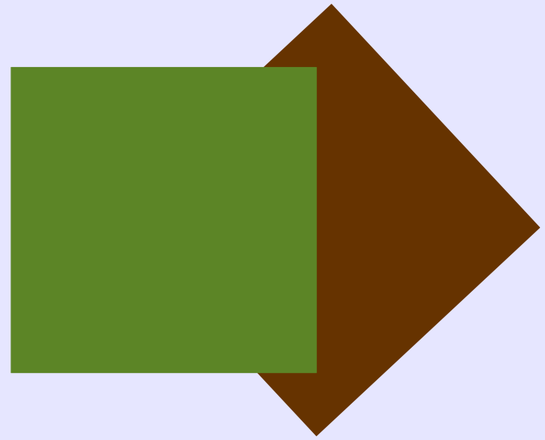
Monocular Cues

1. Static (pictorial) cues
    1. Occlusion
    2. Relative size
    3. Aerial (Atmospheric)
    4. Shading
2. Perspective
    1. Linear perspective
    2. Texture gradient
    3. Relative Height
    4. Cast shadows/shading

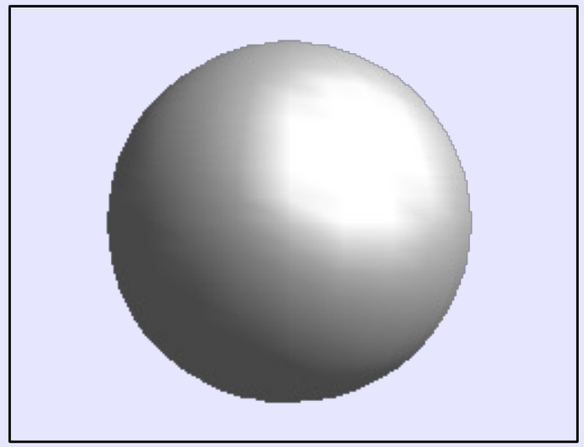
Depth Cues II (Pictorial Depth Cues)



Familiar Size



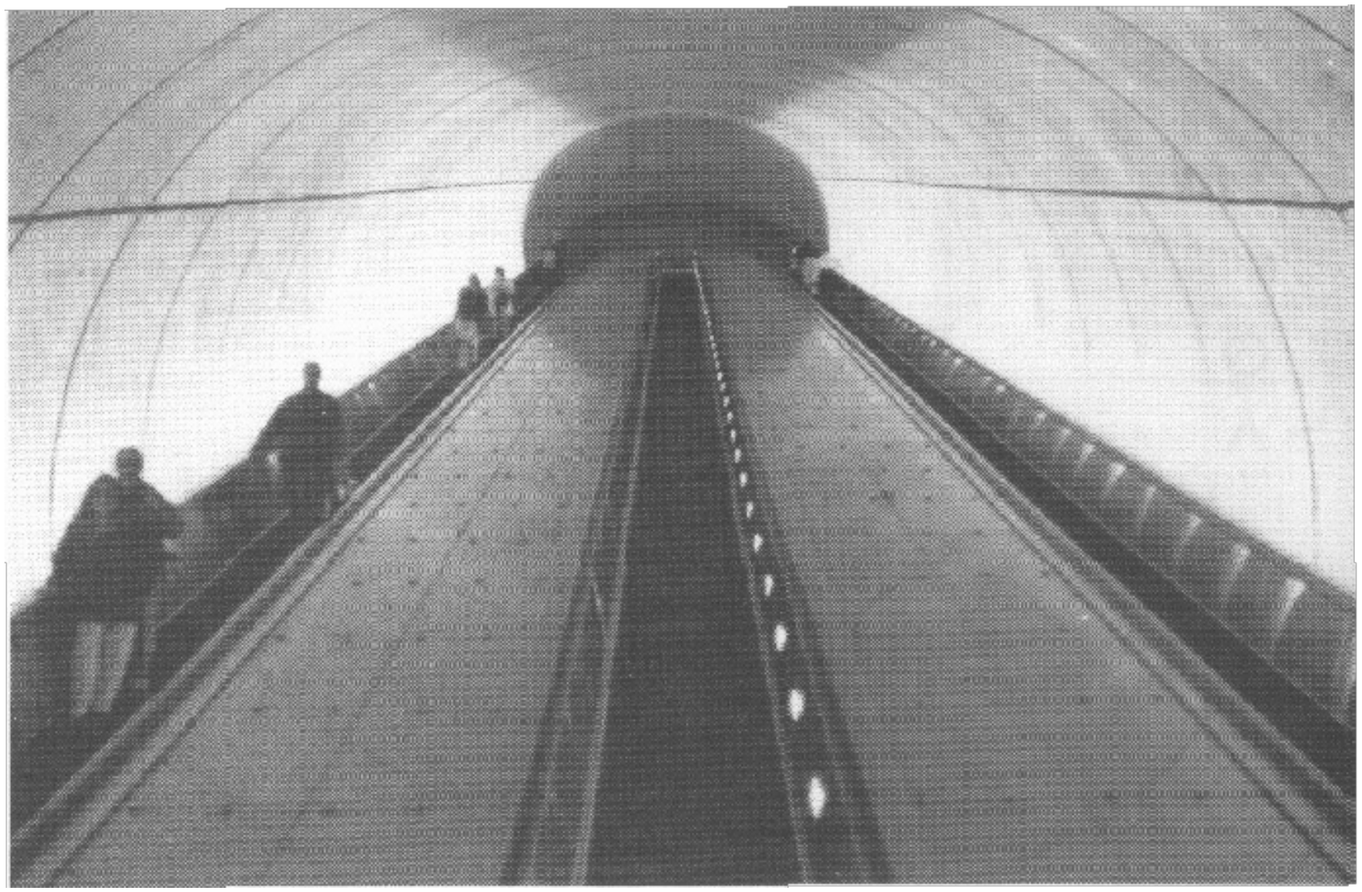
Occlusion



Shading

Atmospheric

Depth Cues III (Pictorial Depth Cues)

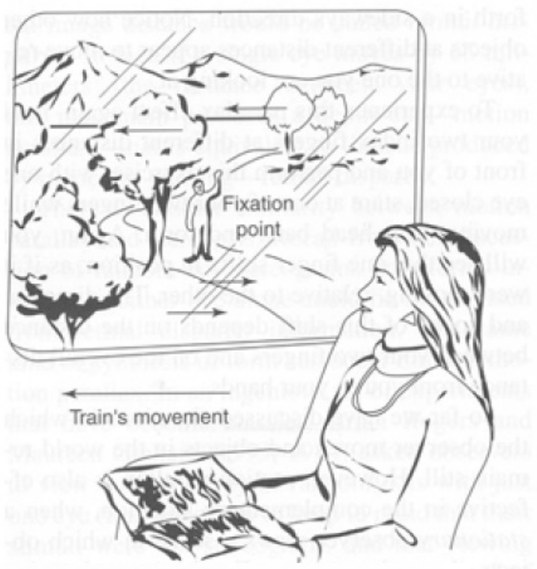


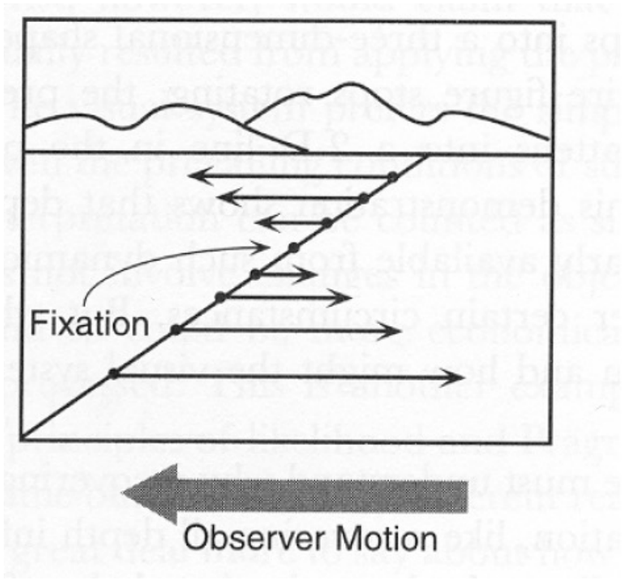
Linear Perspective

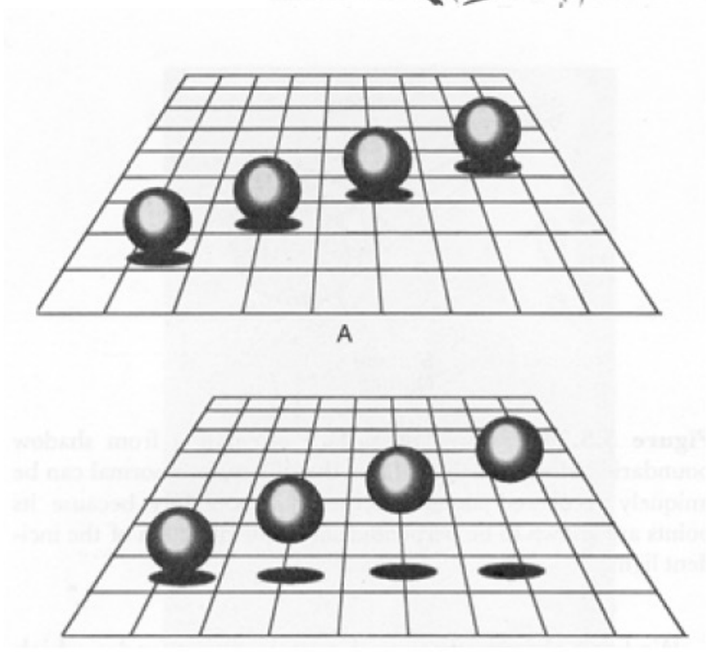
Texture Gradient

Kinematic Information

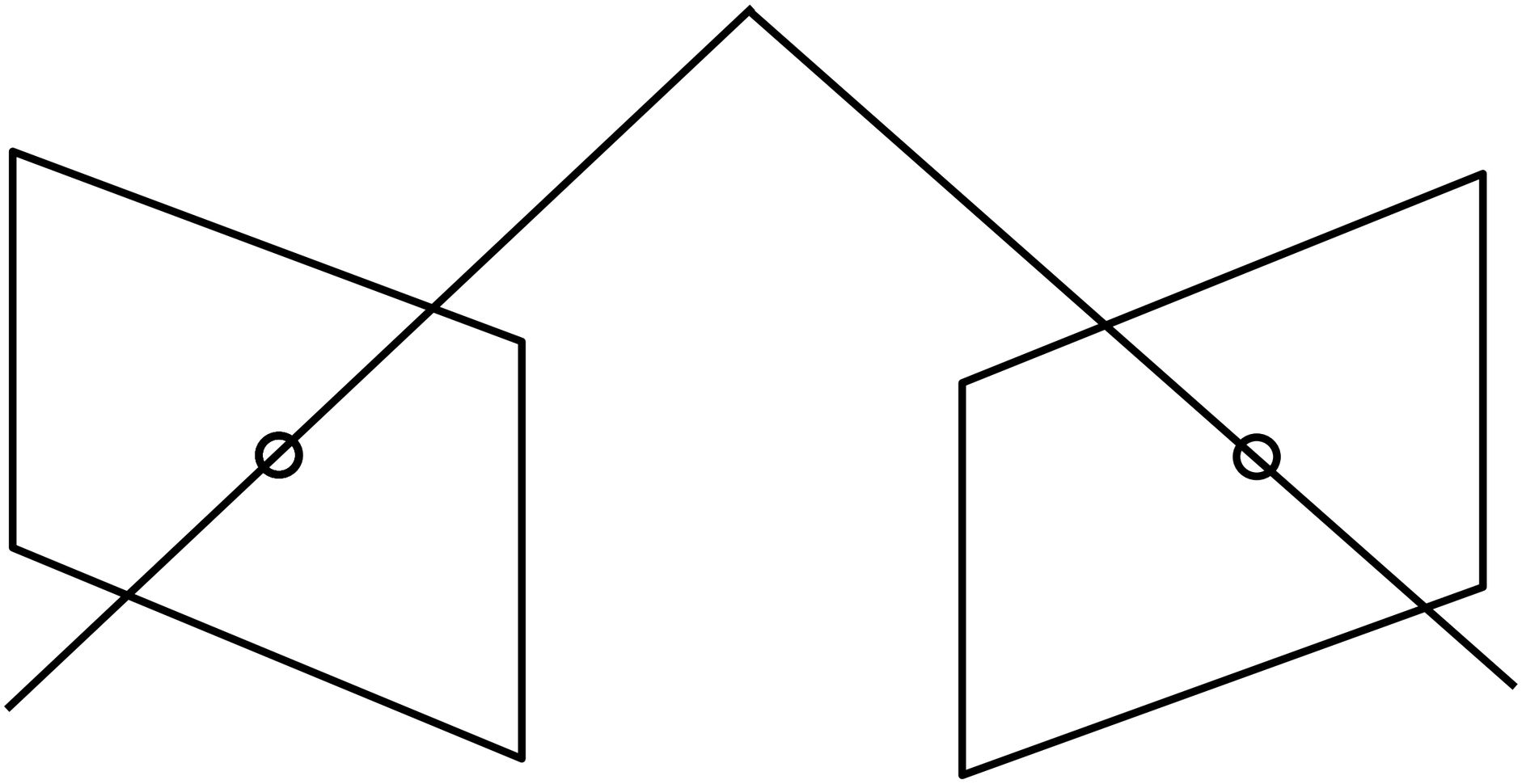
Motion produced monocular depth cue: Motion parallax







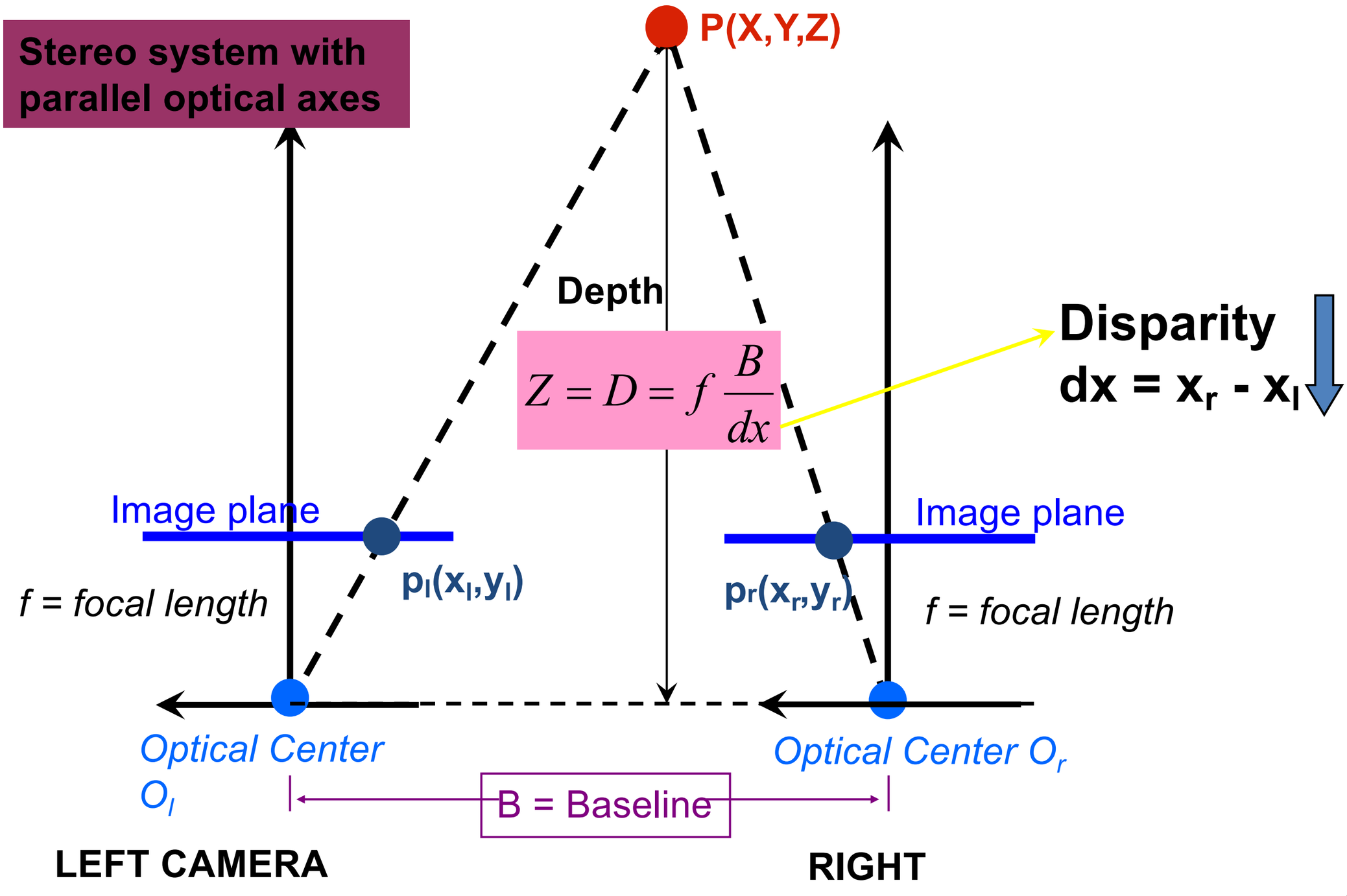
Binocular Cue: Stereo



Definition: Extracting 3D information from two different views of a Scene

SDU ROBOTICS

Stereo system with parallel optical axes



Program for Today

- Literature
- Course Structure
- Applications in Computer Vision
- Overview of course content
- Human Depth Perception
- Camera Geometry (other slide set)