Robot Vision

Two View Geometry & RGBD

Dr. Chen Feng

cfeng@nyu.edu

ROB-UY 3203, Spring 2024

Overview

- RGBD
  - Two-view geometry
    - Epipolar geometry
    - Fundamental/Essential matrix
  - RGBD sensors and principles
- Vanishing Points

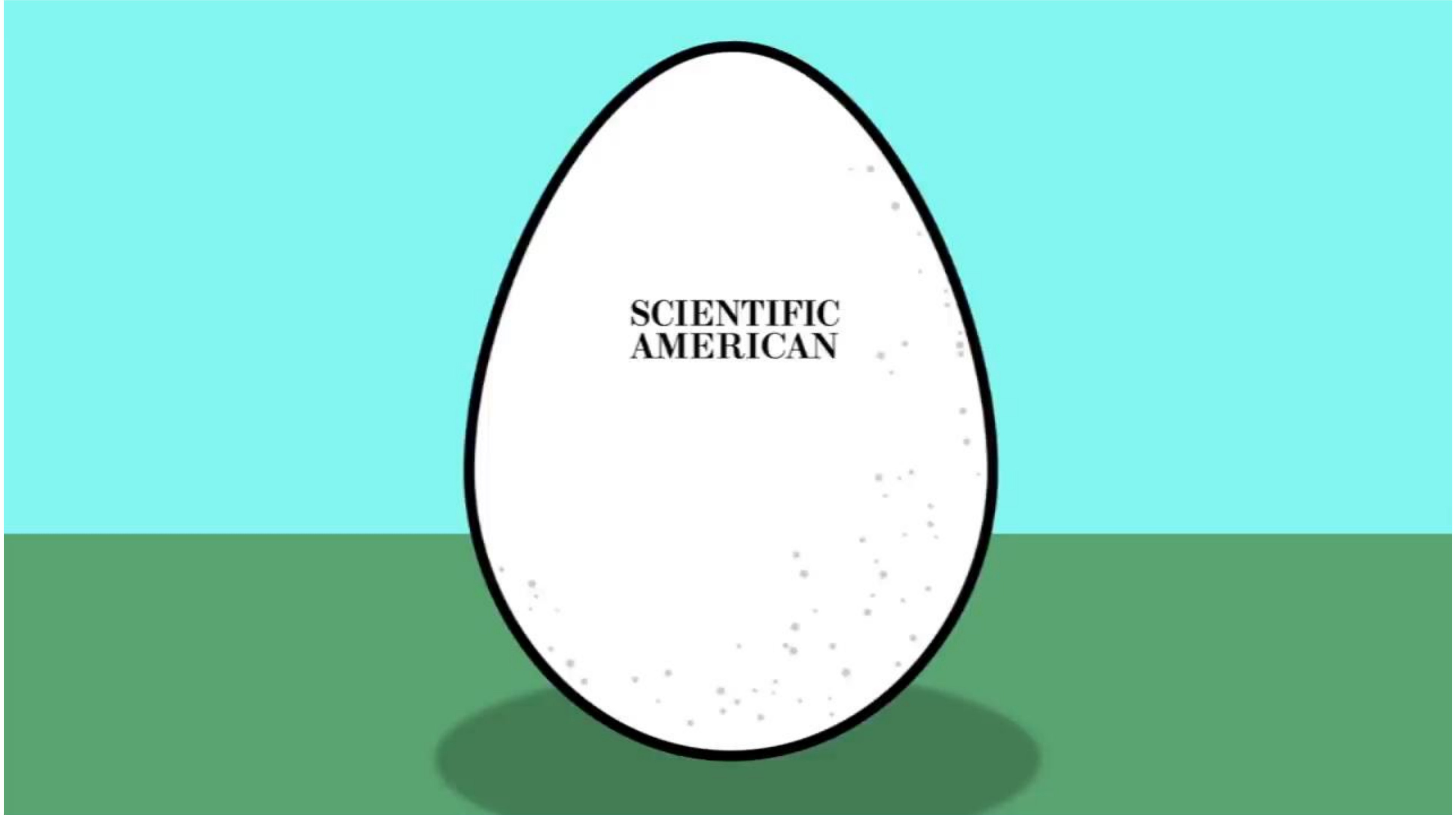
References

- Hartley & Zisserman 2003:
  - Section 9.1, 9.2, 9.5, 9.6
- Corke 2011:
  - Section 14.2, 14.3
- Forsyth & Ponce 2011
  - Chapter 7
- Szeliski 2011:
  - Section 11.1, 12.2, 12.3

Let us start from the movie: Lord of the Rings

- What is the paradox when comparing the two images?

Have you heard of the Ames Room?

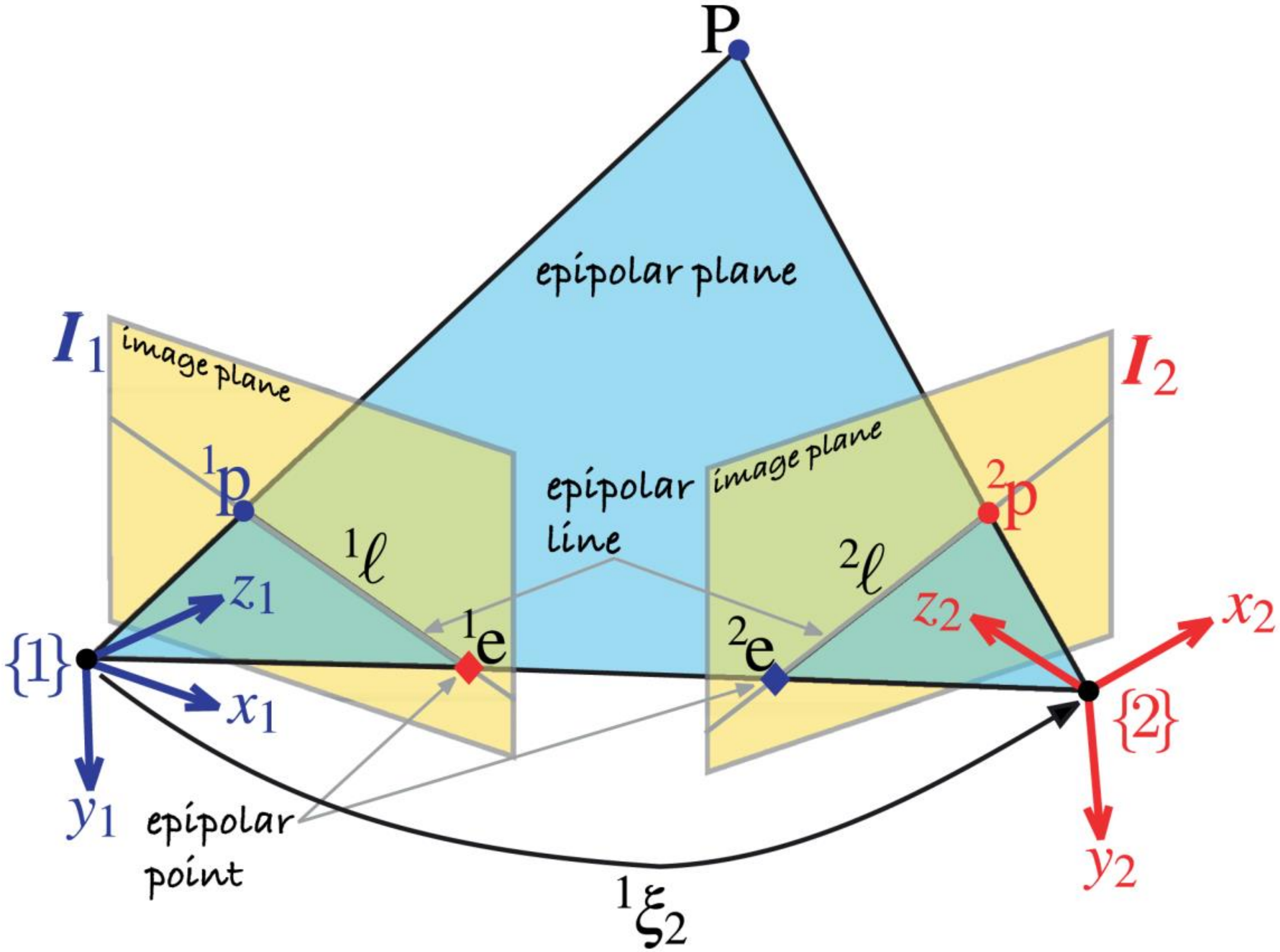


https://www.treehugger.com/optical-illusions-will-blow-your-mind-4864290#mntl-sc-block_1-0-37

How did they implement the illusion in the movie?

https://youtu.be/QWMFpxkGO_s

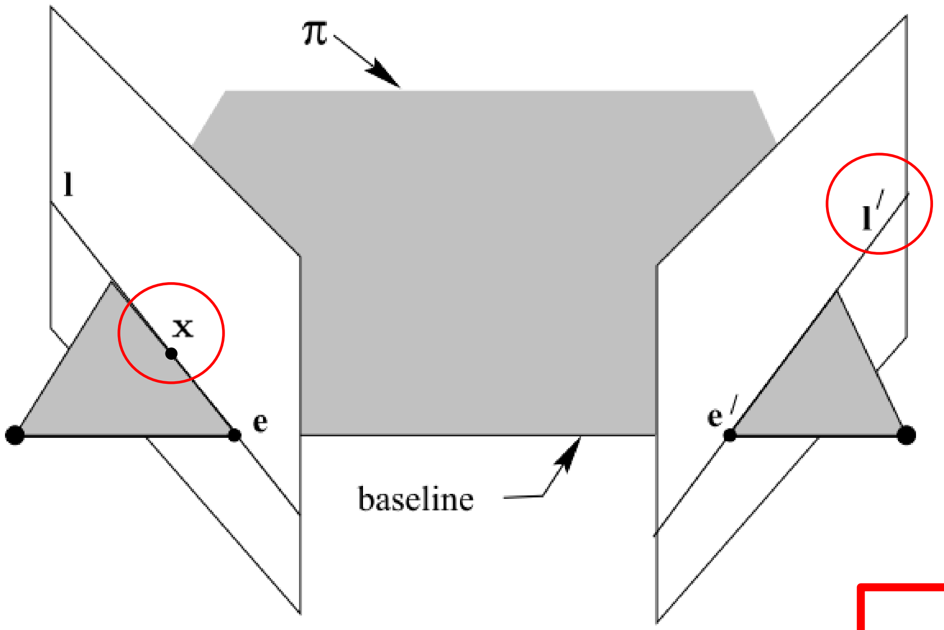
Two-view Geometry

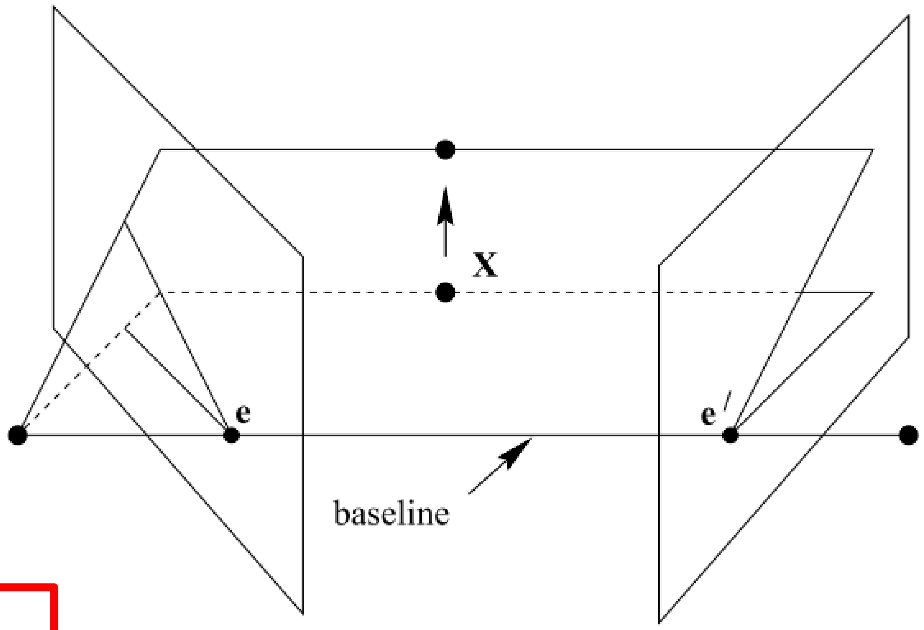


Corke 2011

Epipolar Geometry

- **Independent of scene structure**
- Only depends on the **cameras' internal parameters and relative pose**

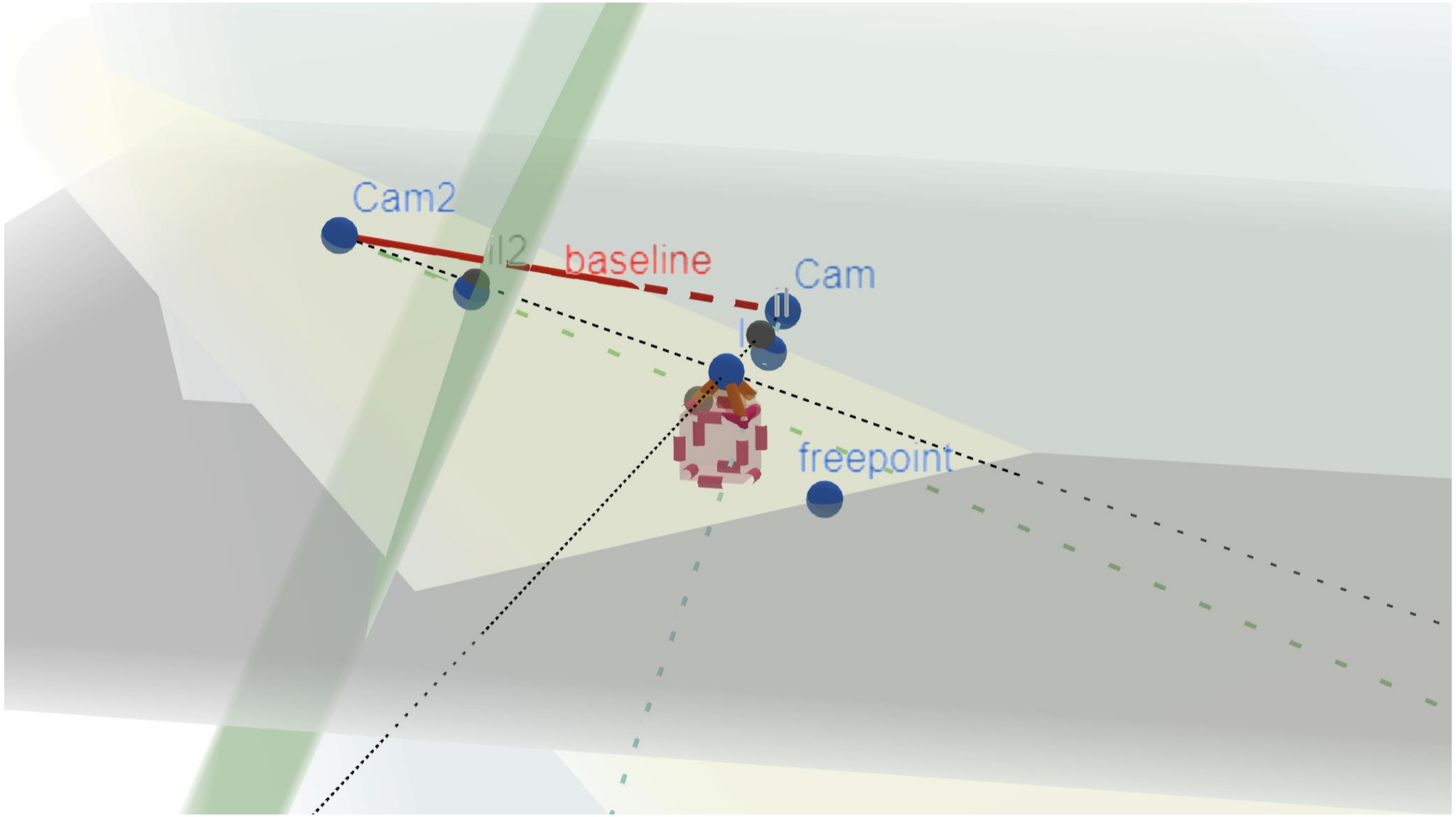




$$\boxed{\mathbf{x} \mapsto \mathbf{l}'}$$

Hartley & Zisserman 2003

Epipolar Geometry: Interactive Demo



https://www.geogebra.org/3d/qrqvgsfk

Fundamental Matrix

Result 9.3. *The fundamental matrix satisfies the condition that for any pair of corresponding points* $\mathbf{x} \leftrightarrow \mathbf{x}'$ *in the two images*

$$\mathbf{x}'^{\mathsf{T}} \mathrm{F} \mathbf{x} = 0.$$

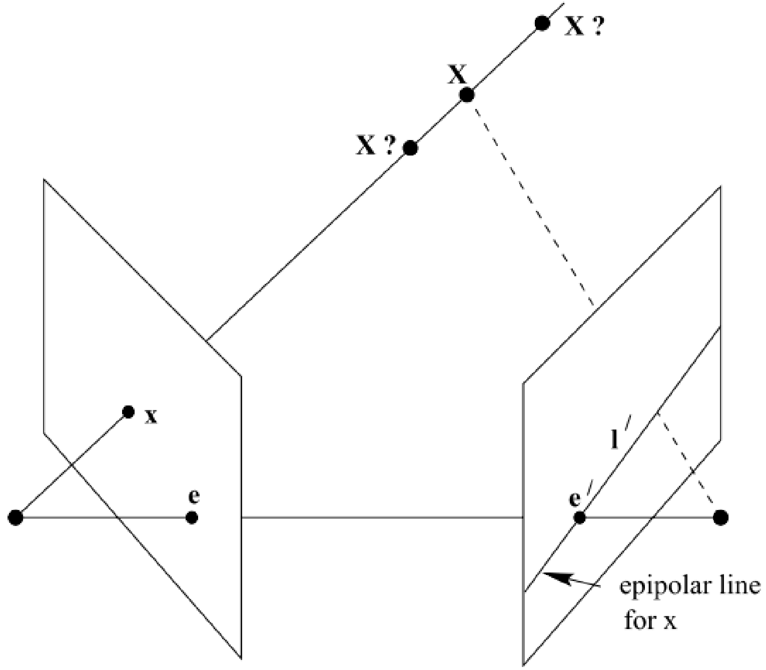
[LonguetHiggins-81] H. C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293:133–135, September 1981.

Hartley & Zisserman 2003

Fundamental Matrix

Result 9.3. *The fundamental matrix satisfies the condition that for any pair of corresponding points* $\mathbf{x} \leftrightarrow \mathbf{x}'$ *in the two images*

$$\mathbf{x}'^{\mathsf{T}} \mathrm{F} \mathbf{x} = 0.$$



$$\boxed{\mathbf{x} \mapsto \mathbf{l}'}$$

Hartley & Zisserman 2003

Normalized Coordinate and Essential Matrix

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

$$\mathtt{P} = \mathtt{K}[\mathtt{R} \mid \mathbf{t}]$$

Normalized coordinate $\hat{\mathbf{x}} = \mathtt{K}^{-1}\mathbf{x} = [\mathtt{R} \mid \mathbf{t}]\mathbf{X}$

$$\hat{\mathbf{x}}'^{\mathsf{T}}\mathtt{E}\hat{\mathbf{x}} = 0$$

Result 9.17. *A 3×3 matrix is an essential matrix if and only if two of its singular values are equal, and the third is zero.*

$$\mathtt{E} = \mathtt{K}'^{\mathsf{T}}\mathtt{F}\mathtt{K} \qquad \boxed{\mathtt{E} = [\mathbf{t}]_{\times}\mathtt{R}}$$

Hartley & Zisserman 2003

Estimating F-matrix

- Find multiple $\mathbf{X} \leftrightarrow \mathbf{X}'$ correspondences **(≥ 7)** between two images

$$\mathbf{x}'^{\mathsf{T}}\mathbf{F}\mathbf{x} = 0$$

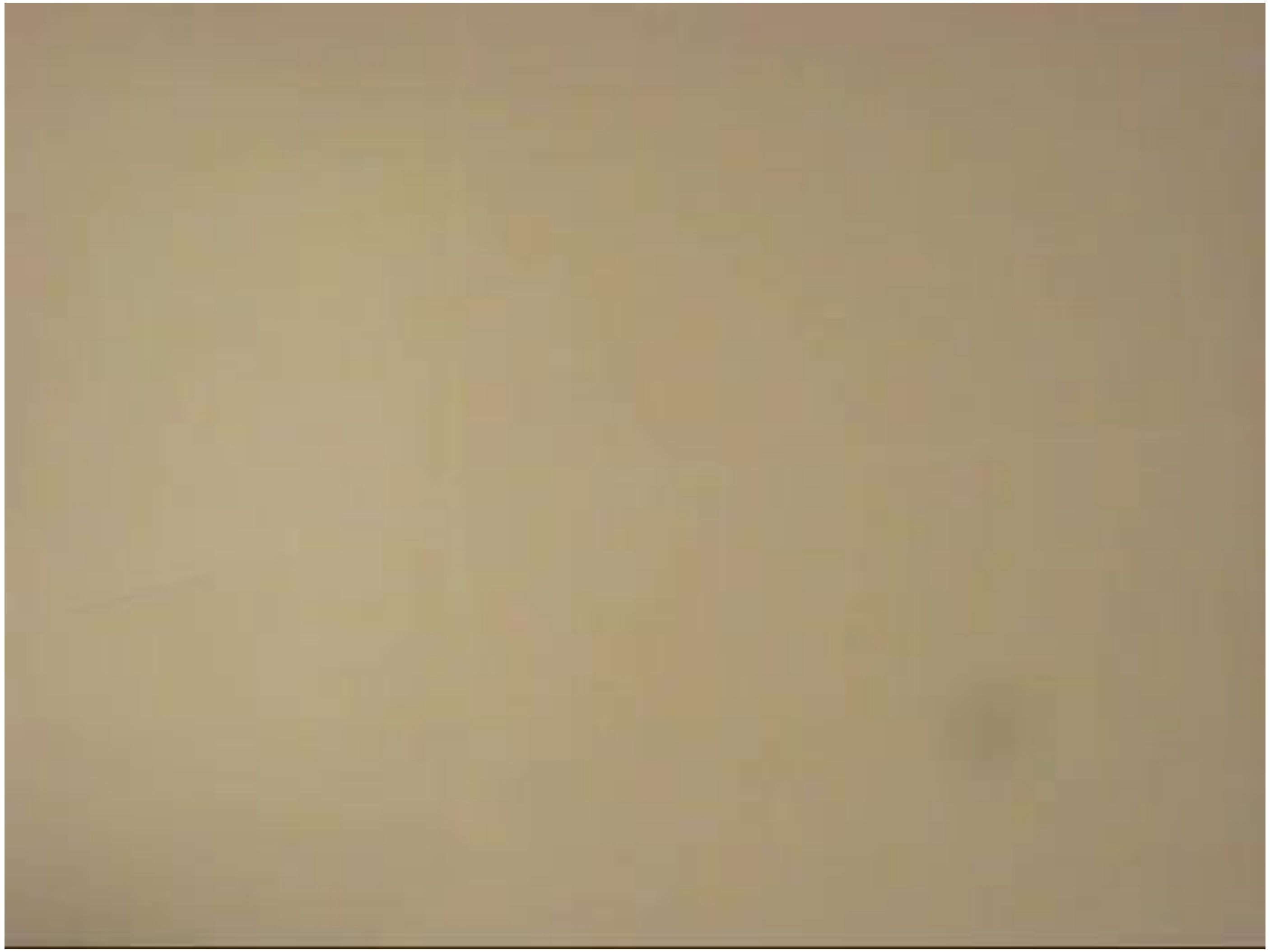
$$x'xf_{11} + x'yf_{12} + x'f_{13} + y'xf_{21} + y'yf_{22} + y'f_{23} + xf_{31} + yf_{32} + f_{33} = 0$$

$$\mathbf{Af} = \begin{bmatrix} x'_1x_1 & x'_1y_1 & x'_1 & y'_1x_1 & y'_1y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ x'_nx_n & x'_ny_n & x'_n & y'_nx_n & y'_ny_n & y'_n & x_n & y_n & 1 \end{bmatrix} \mathbf{f} = \mathbf{0}$$

- Enforce rank-2 constraint by SVD

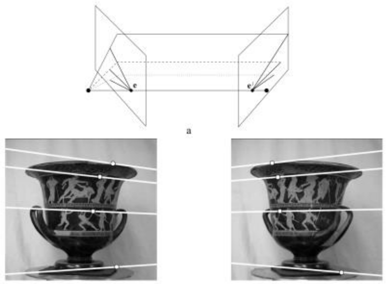
Hartley & Zisserman 2003

The Fundamental Matrix Song – Daniel Wedge



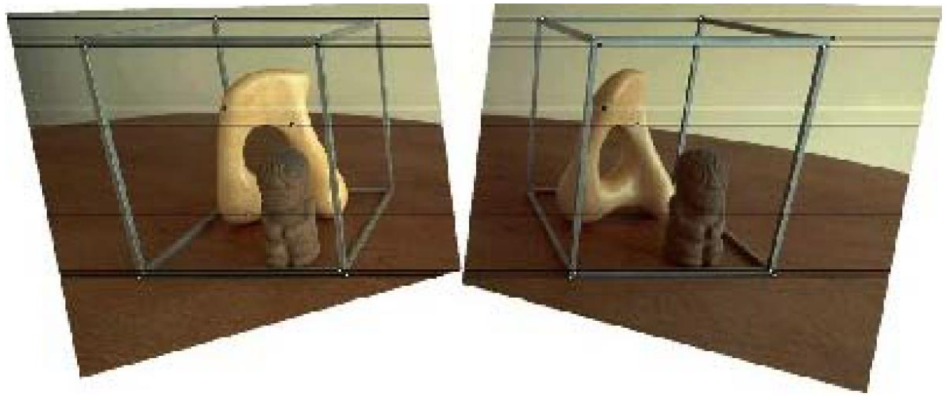
Video from: https://youtu.be/DgGV3l82NTk

Depth from Stereo Images



Hartley & Zisserman 2003

- Epipolar geometry: stereo calibration
- Finding disparity by searching along epipolar lines
    - Low texture: bad
    - Repeated texture: bad

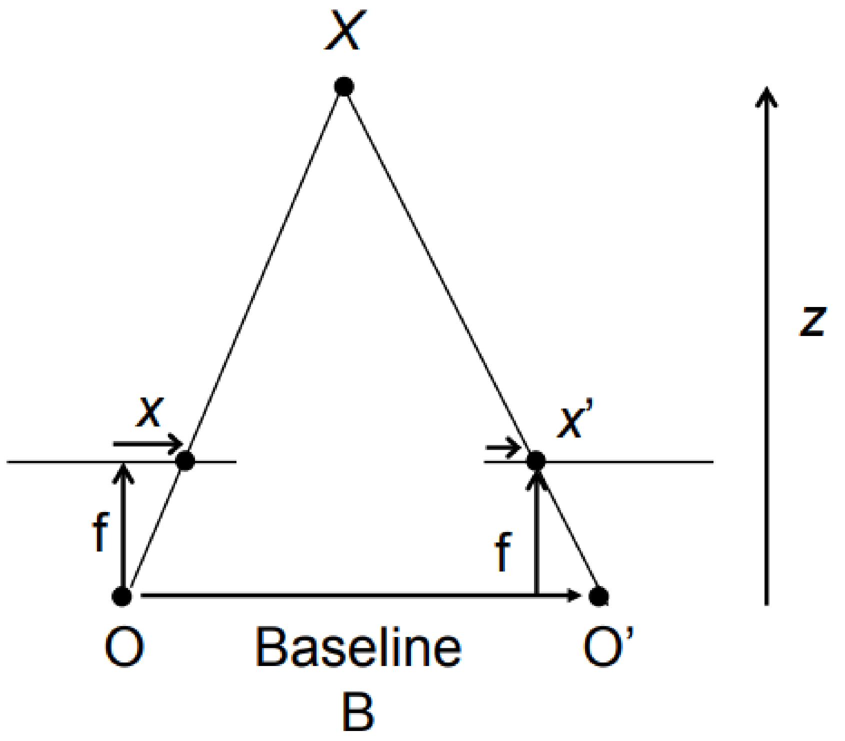


Szeliski 2011

Depth from Stereo Images

- Disparity and depth
  - Inversely related
- Stereo camera infers pixel depth from disparity
- Longer baseline==better depth accuracy

$$\frac{x - x'}{O - O'} = \frac{f}{z}$$

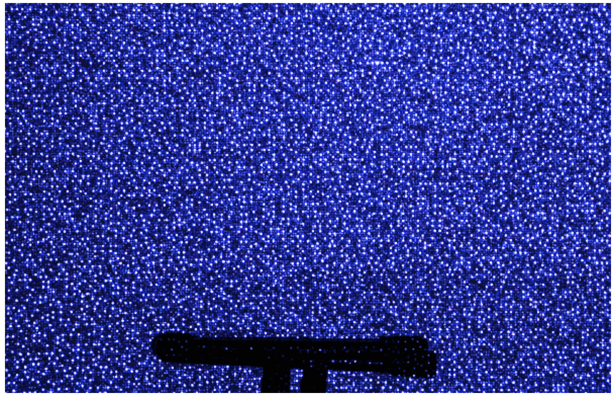
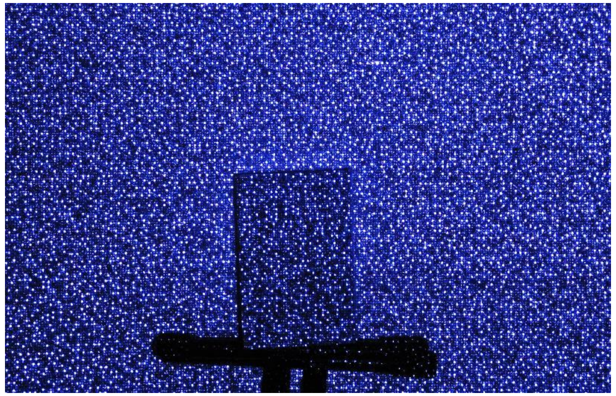


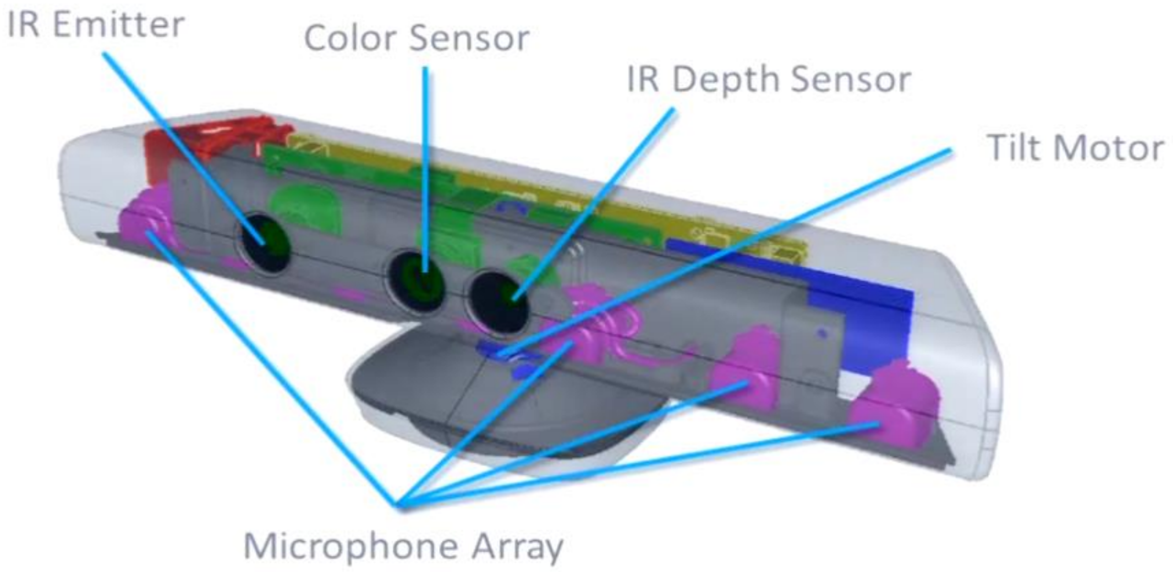
$$disparity = x - x' = \frac{B \cdot f}{z} \qquad z = \frac{B \cdot f}{x - x'}$$

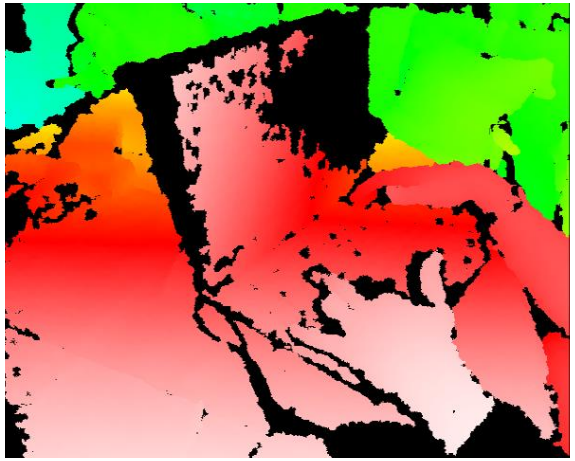
http://pages.cs.wisc.edu/~dyer/cs534/slides/17_kinect.pdf

Depth from Structured Light

- projector + camera
  - Microsoft Kinect
  - Texture in the environment: generally not important
  - But cannot be used under direct sunlight



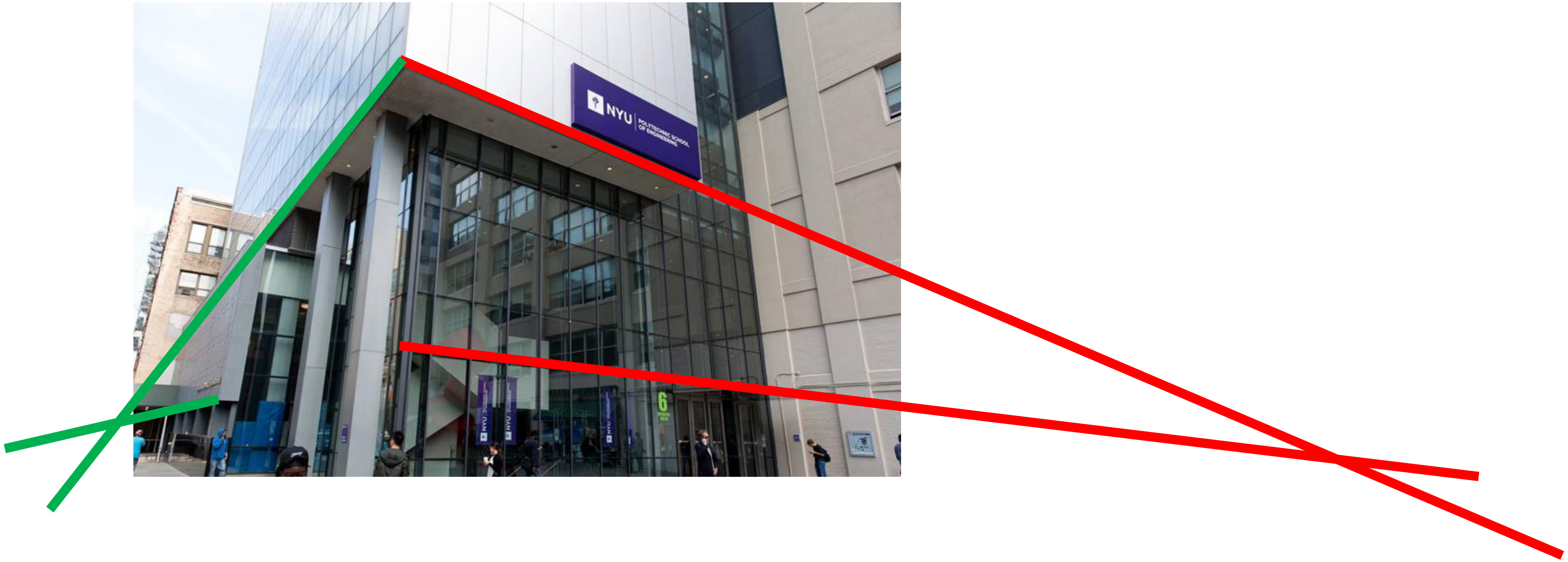
Other RGBD Sensor Alternatives

- Intel RealSense
- Asus Xtion Pro
- Microsoft Kinect V2
- Structure Sensor

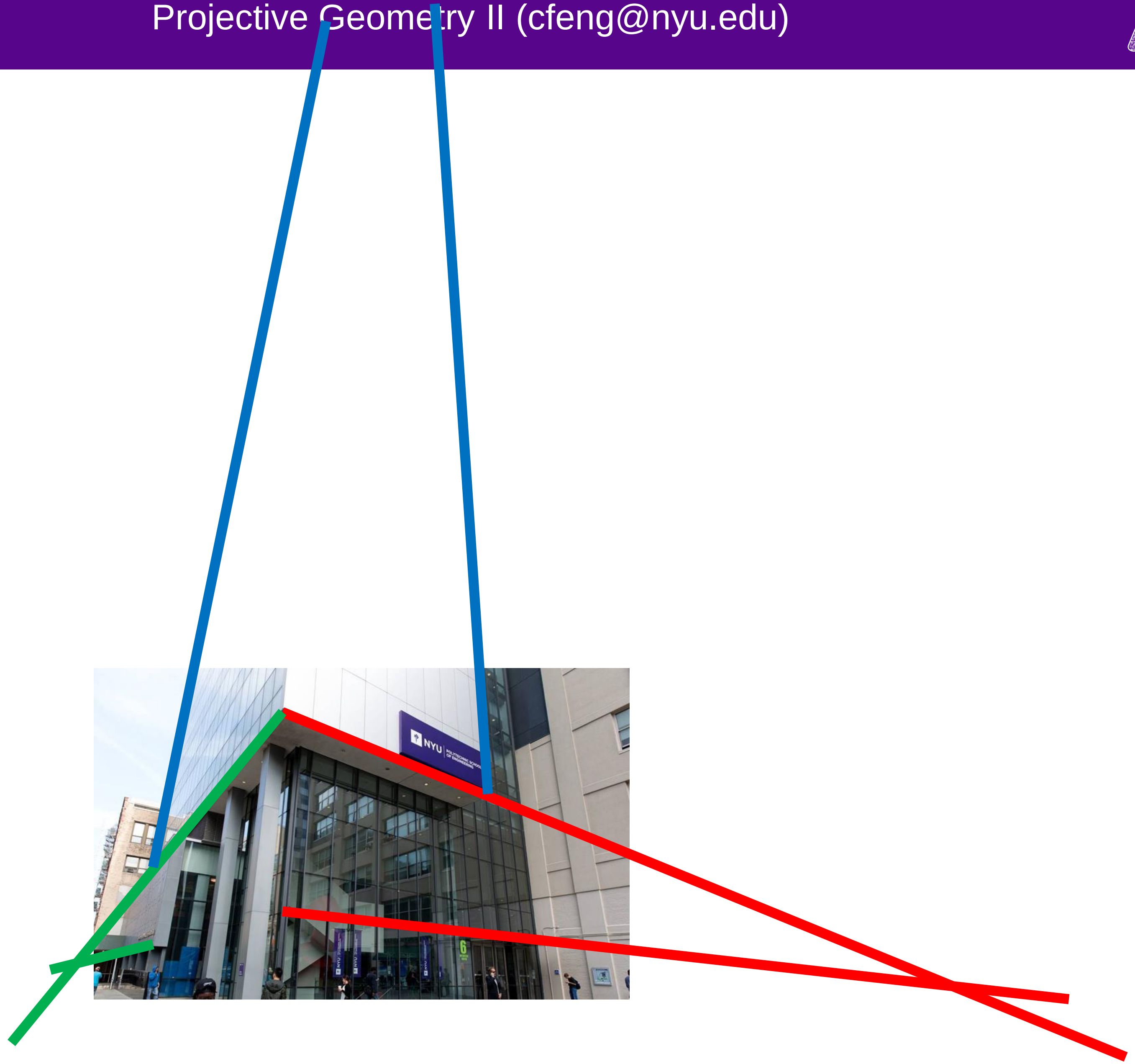
- If you do not have RGBD sensor
  - http://www.michaelfirman.co.uk/RGBDdatasets/

Vanishing Point

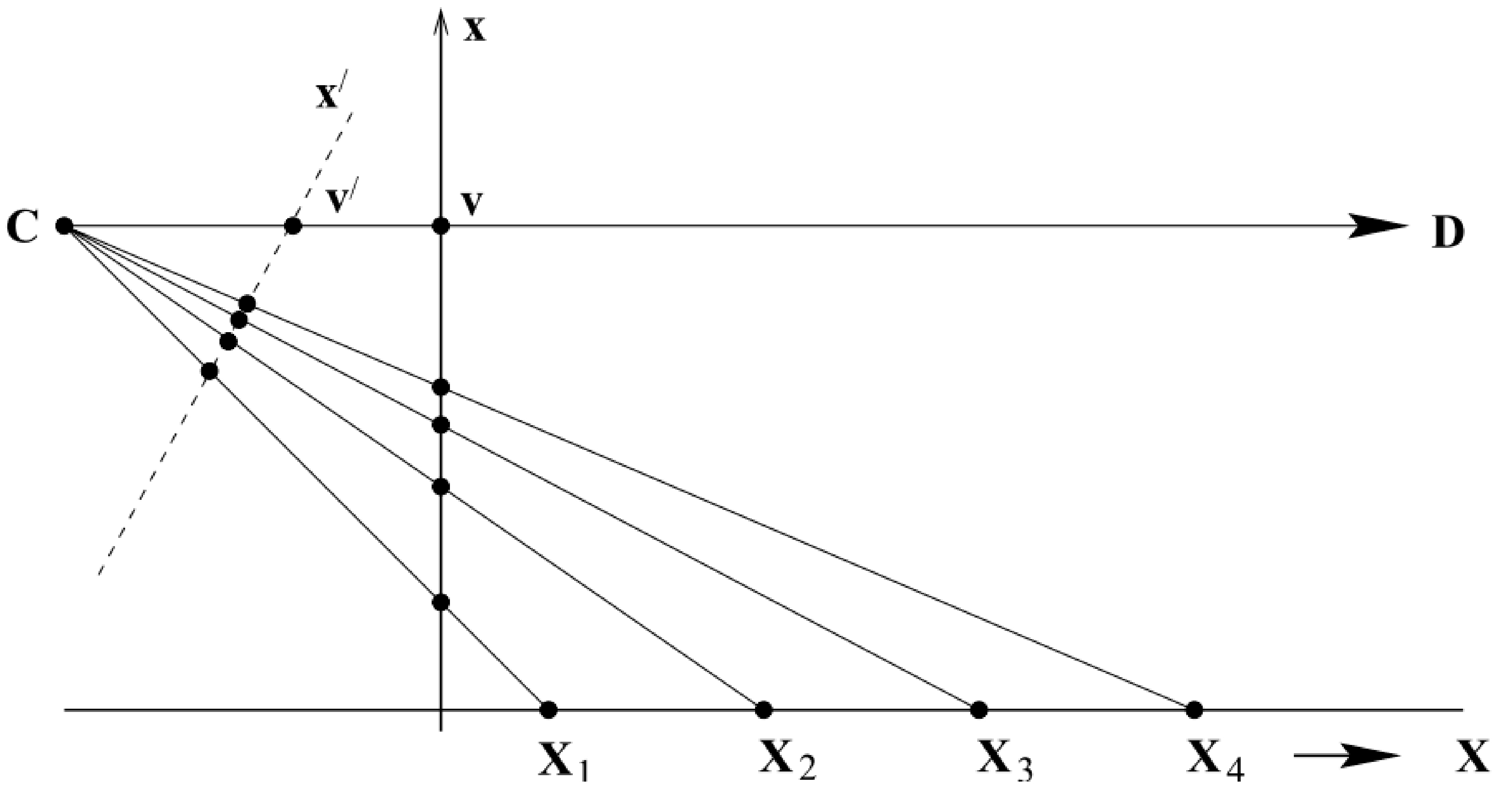
Vanishing Points



Vanishing Points



Vanishing Point – 2D

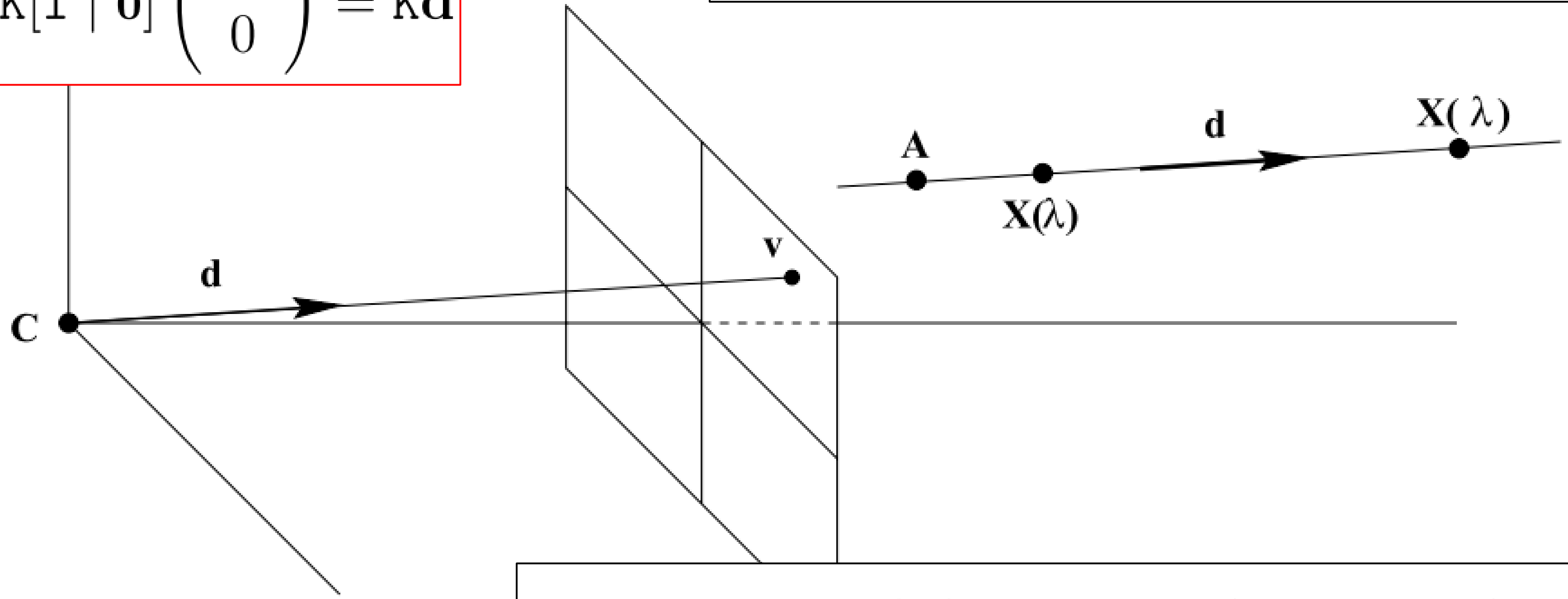


Hartley & Zisserman 2003

Vanishing Point – 3D

$$\mathbf{v} = \mathtt{P}\mathbf{X}_\infty = \mathtt{K}[\mathtt{I} \mid \mathbf{0}]\begin{pmatrix}\mathbf{d}\\0\end{pmatrix} = \mathtt{K}\mathbf{d}$$

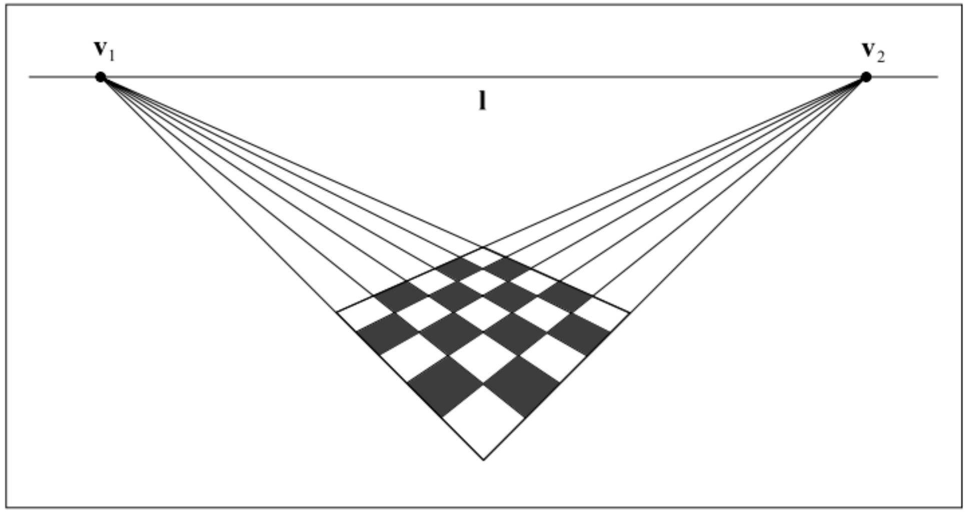
$$\mathbf{x}(\lambda) = \mathtt{P}\mathbf{X}(\lambda) = \mathtt{P}\mathbf{A} + \lambda\mathtt{P}\mathbf{D} = \mathbf{a} + \lambda\mathtt{K}\mathbf{d}$$



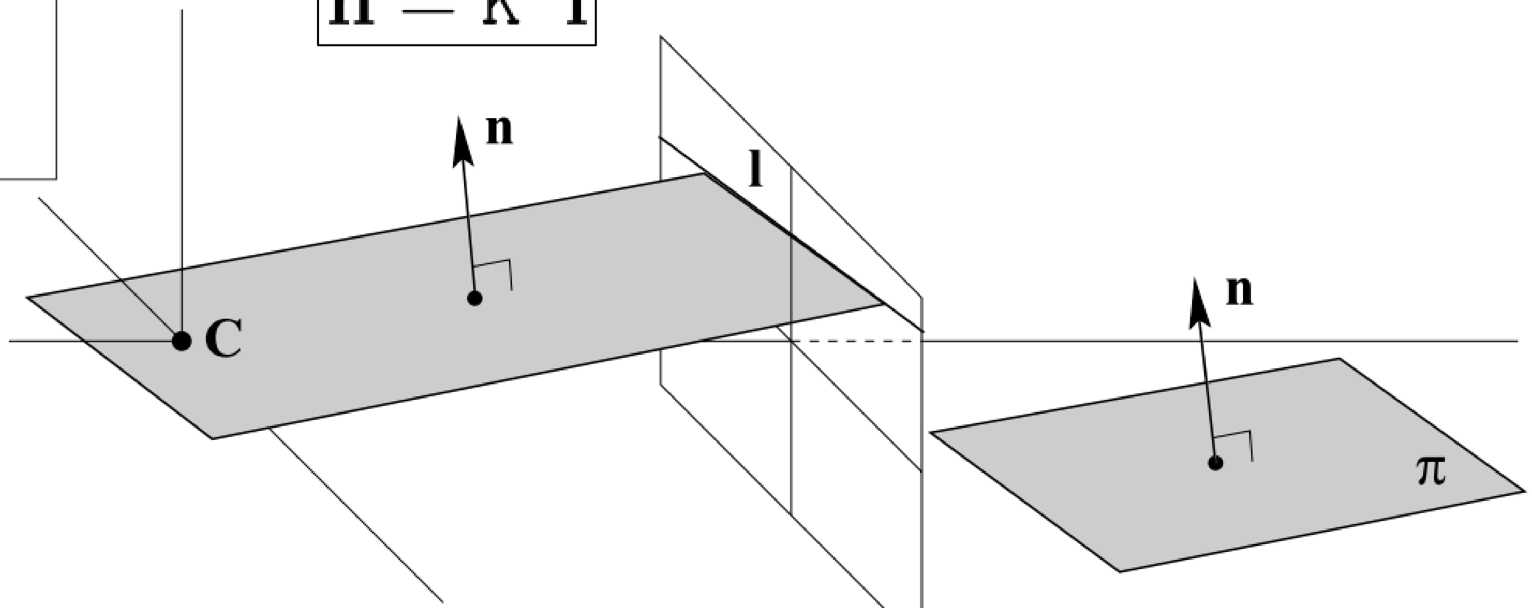
$$\mathbf{v} = \lim_{\lambda\to\infty} \mathbf{x}(\lambda) = \lim_{\lambda\to\infty} (\mathbf{a} + \lambda\mathtt{K}\mathbf{d}) = \mathtt{K}\mathbf{d}$$

Hartley & Zisserman 2003

Vanishing Line

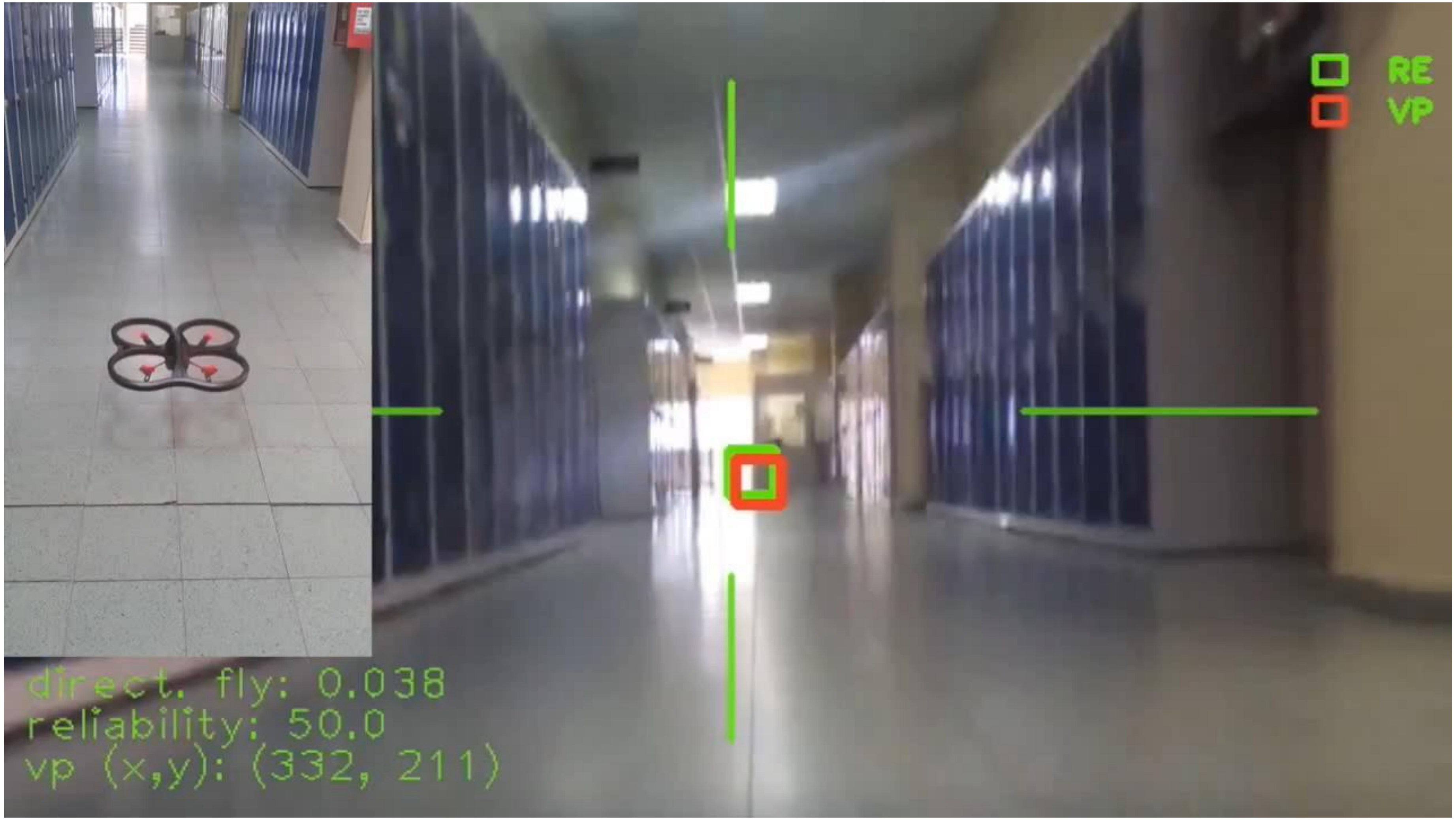


$$\boxed{\mathbf{n} = K^{\top}\mathbf{l}}$$



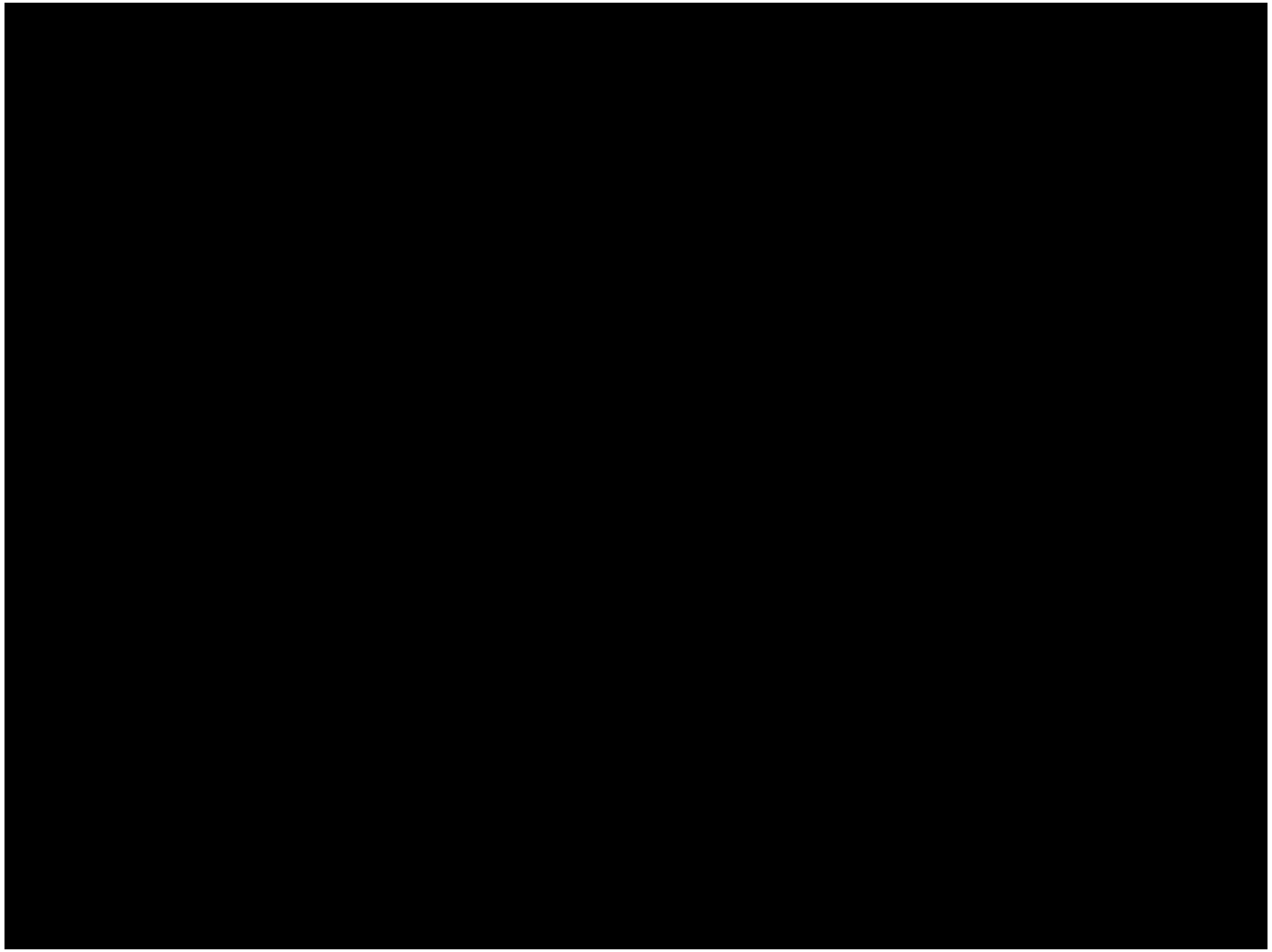
Hartley & Zisserman 2003

Vanishing Point for Robots



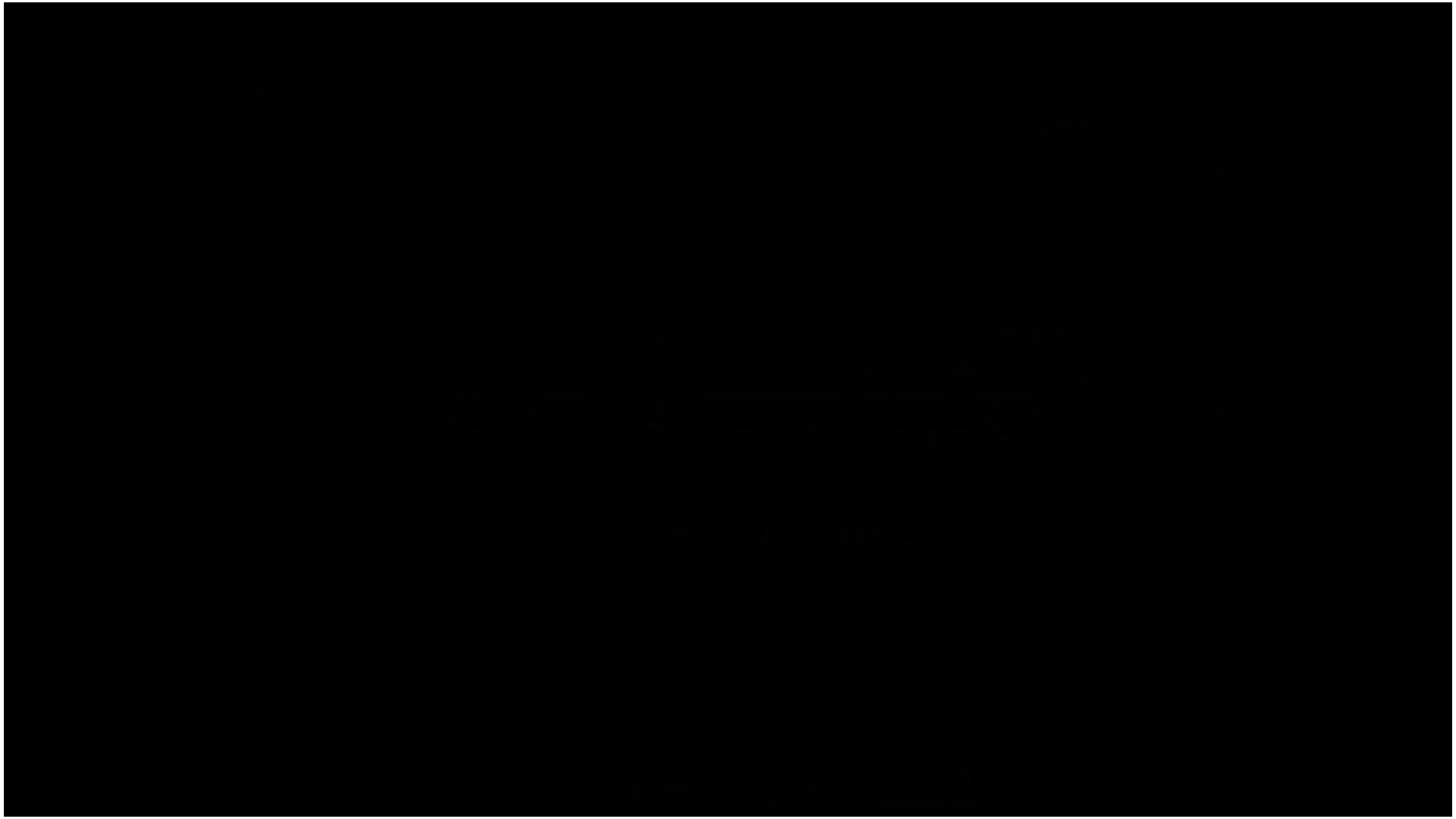
Video from: https://youtu.be/Te4AHOBHbiY

Vanishing Point for Autonomous Driving



Video from: https://youtu.be/Iws6-q9ji0o

Vanishing Point for Traffic Surveillance



Video from: https://youtu.be/S3msCdn3fNM

References for Next Week

- Forsyth & Ponce 2011
  - Chapter 8
- Szeliski 2011:
  - Chapter 11
- Corke 2022:
  - Section 14.4
- Hartley & Zisserman 2003:
  - Section 5.2, 18.1, A6

- Chen Feng, Vineet R. Kamat, and Carol C. Menassa. "Marker-Assisted Structure from Motion for 3D Environment Modeling and Object Pose Estimation." *(2016)*.