THE UNIVERSITY OF BRITISH COLUMBIA

CPSC 425: Computer Vision

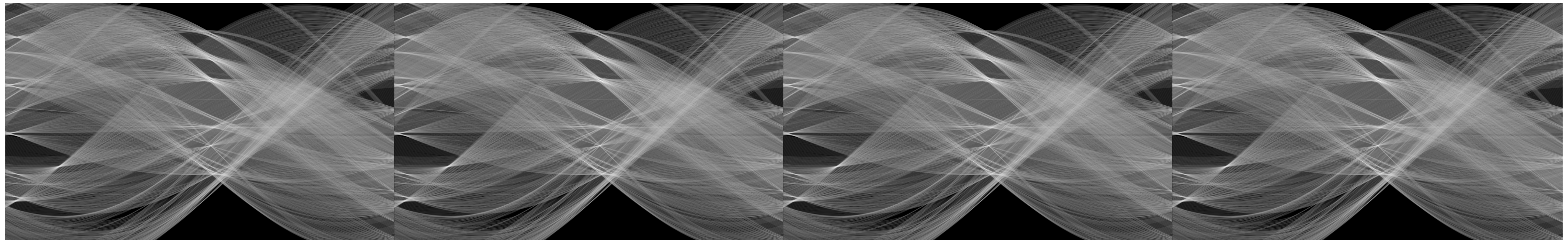


Image Credit: Ioannis (Yannis) Gkioulekas (CMU)

Lecture 14: Stereo

Menu for Today

Topics:

- 3D Correspondence, **Epipolar** Geometry
- **Stereo** Vision
- **Quiz 4**

Readings:

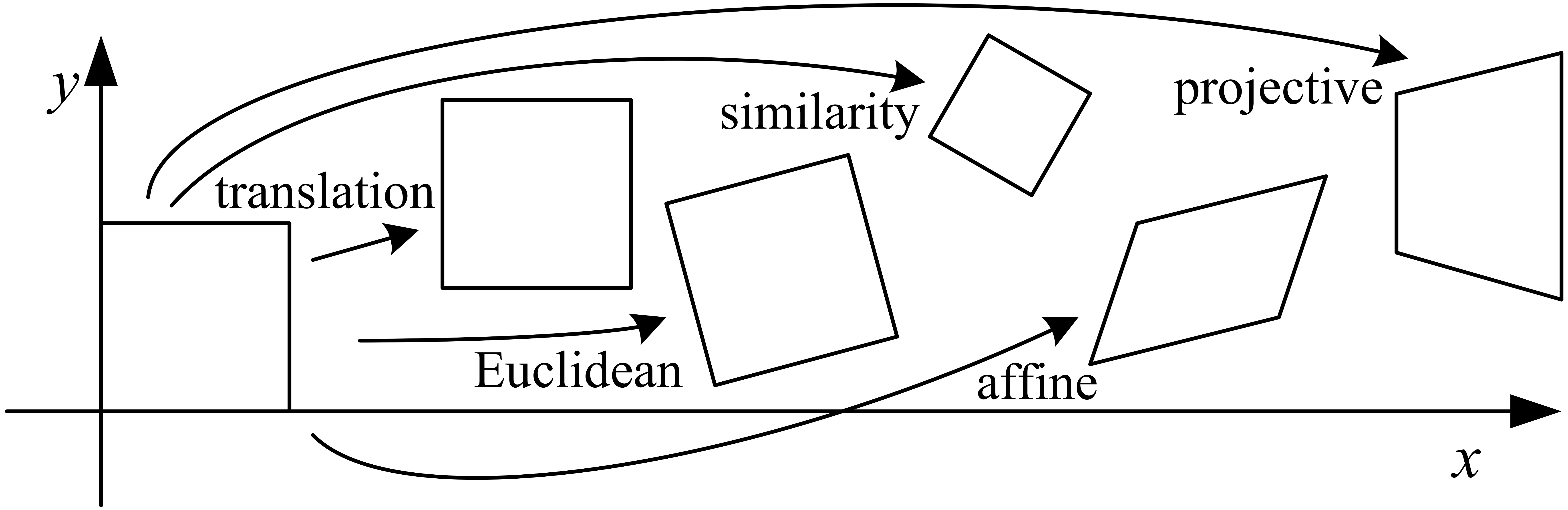
- **Today's** Lecture: Szeliski 12.1, 12.3-12.4, 9.3

Reminders:

- **Assignment 4**: RANSAC and Panoramas due **November 9th**

Recap: 2D Transformations

- We will look at a family that can be represented by 3x3 matrices

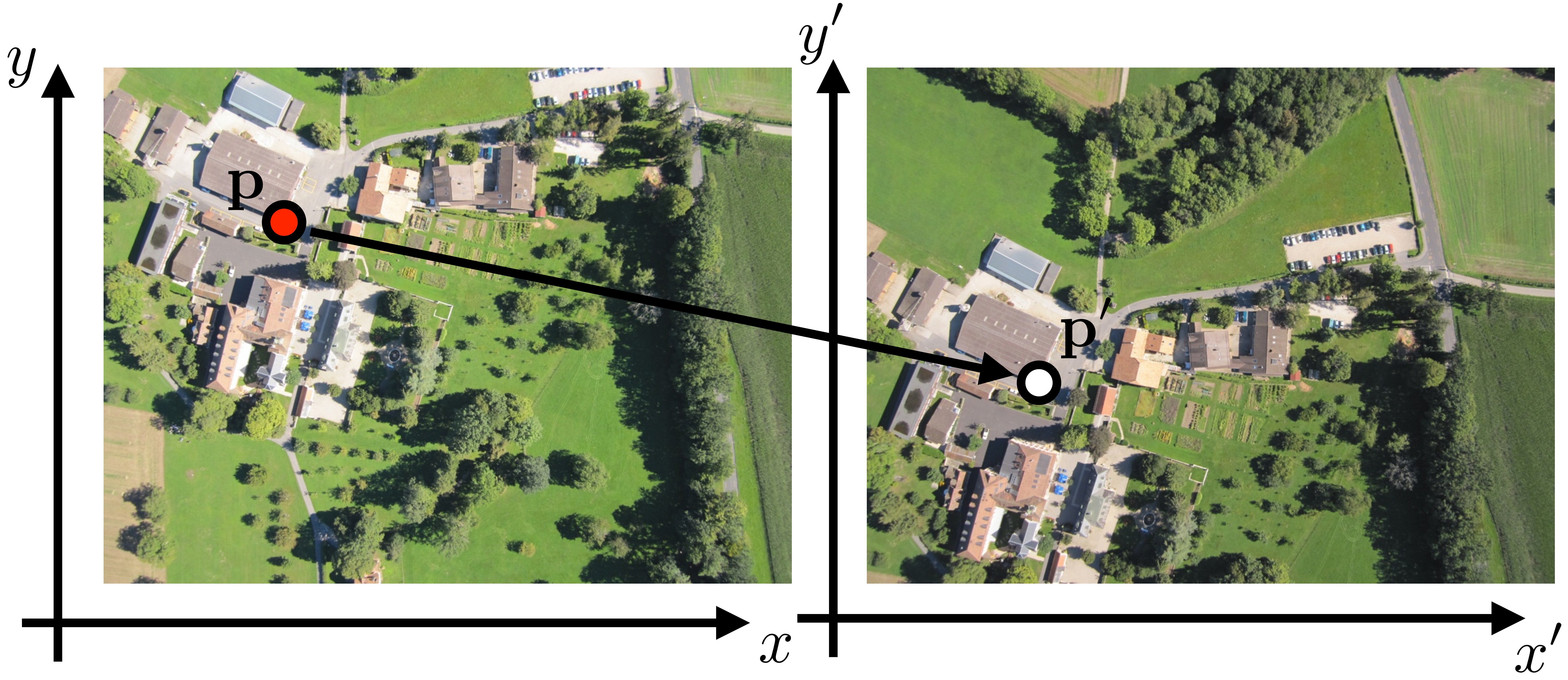


This group represents perspective projections of **planar surfaces** in the world

[Szeliski 2.1]

Recap: 2D Transformations

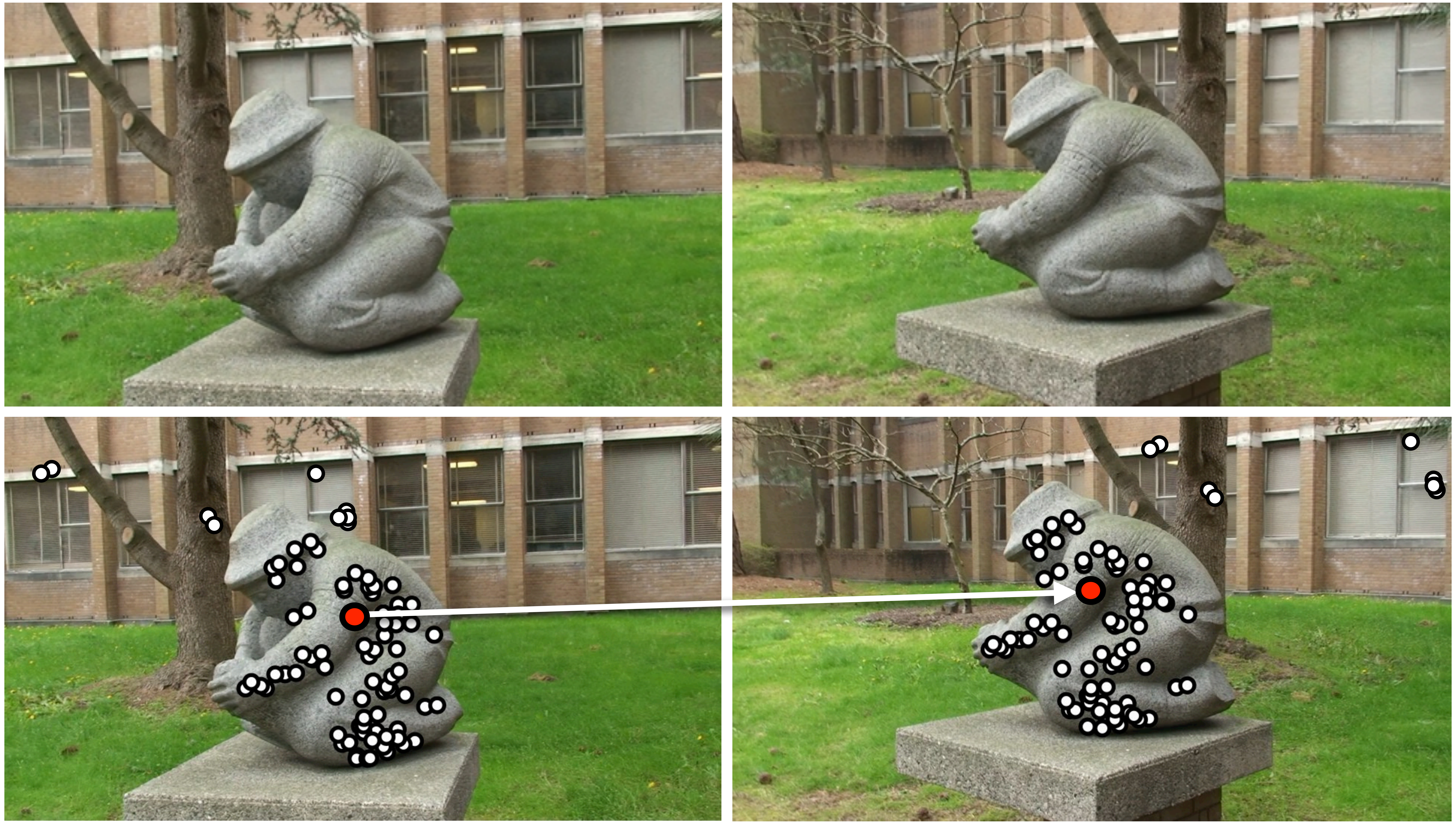
- 3x3 matrix defines one-to-one mapping between views



$$\begin{bmatrix} x_1' \\ y_1' \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}$$

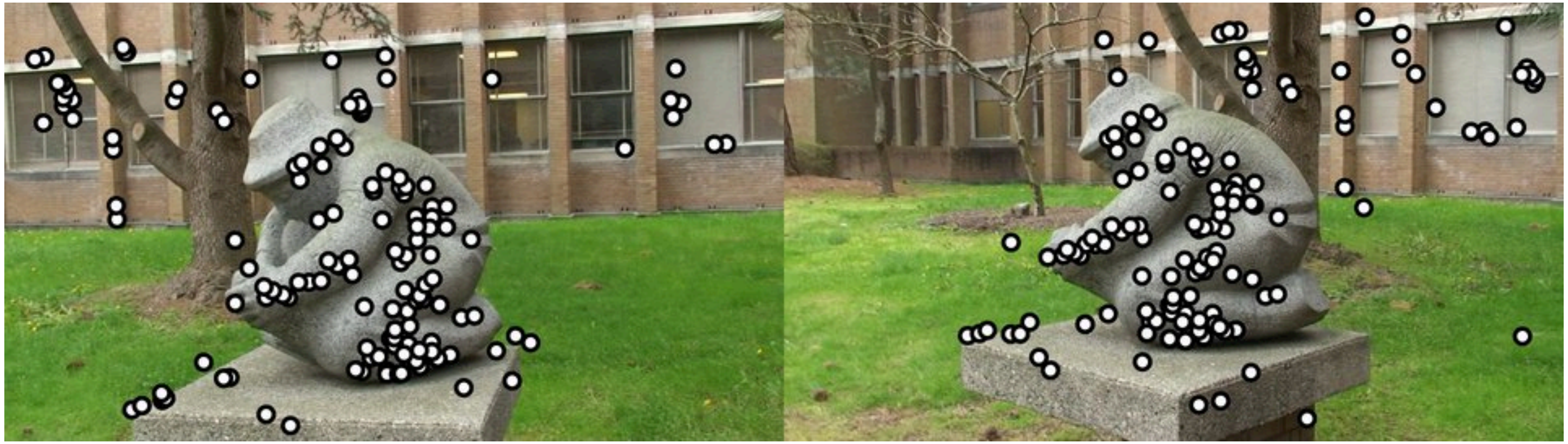
Correspondences in 3D

- Find all matches between views



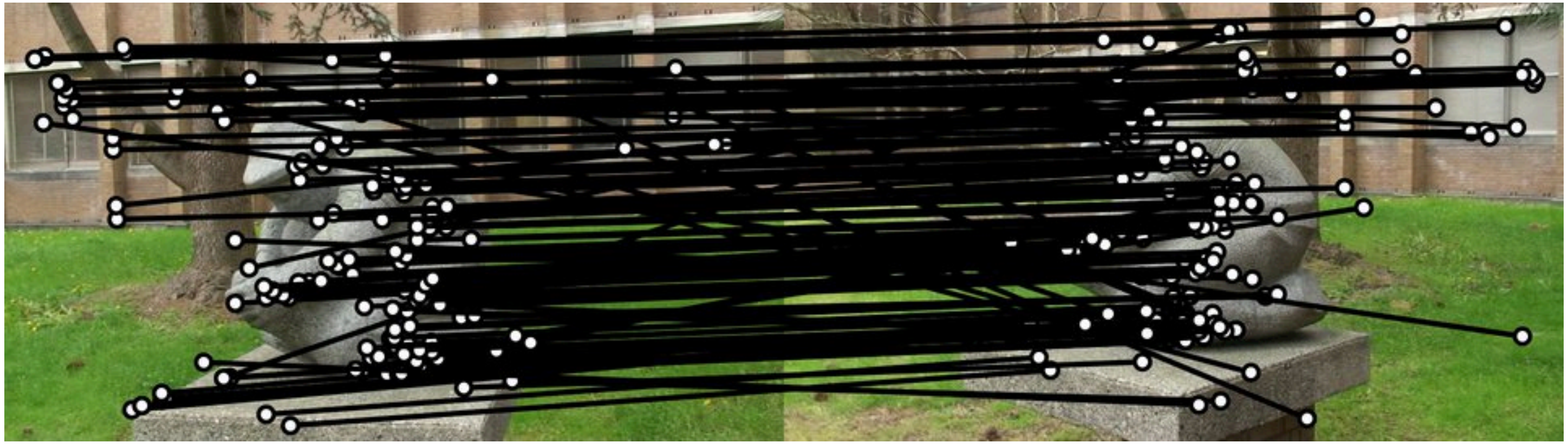
Correspondences in 3D

- Find subset of matches that are consistent with a geometric transformation



Correspondences in 3D

- Find subset of matches that are consistent with a geometric transformation

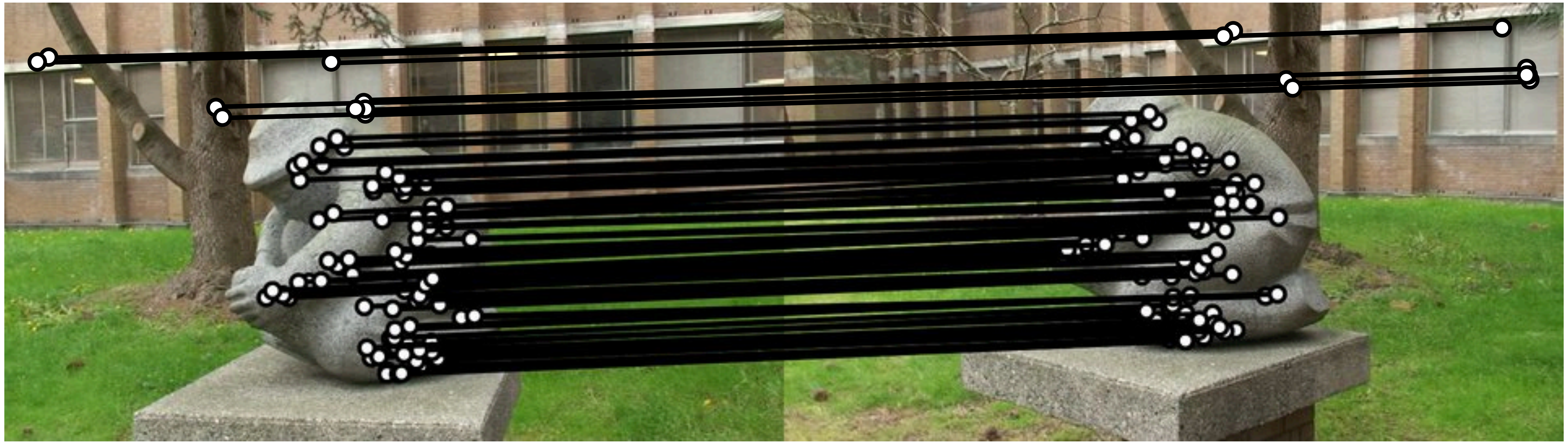


Correspondences in 3D

- Find subset of matches that are consistent with a geometric transformation

Correspondences in 3D

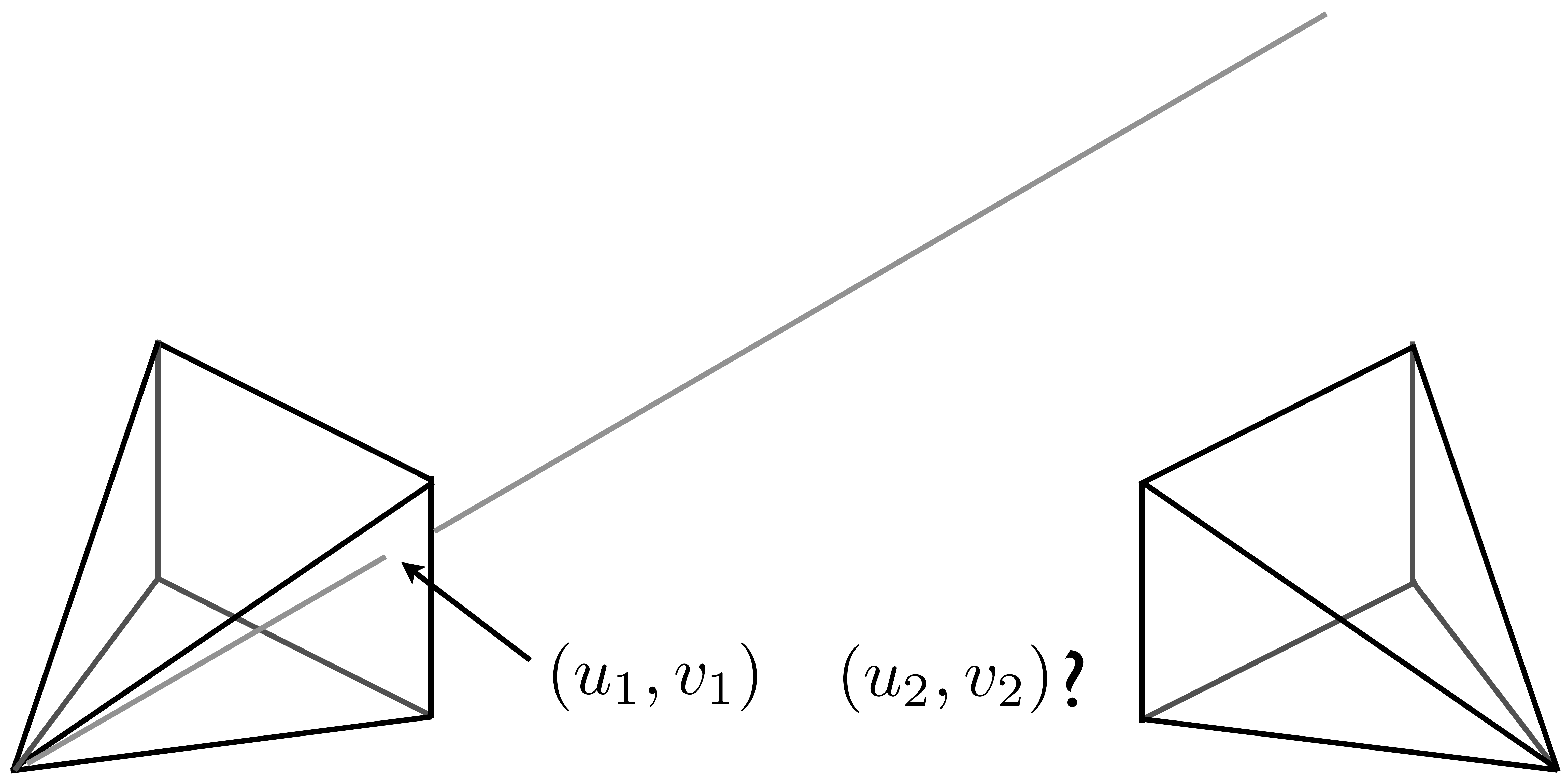
- Find subset of matches that are consistent with a geometric transformation



Consistent matches can be used for subsequent stages, e.g., 3D reconstruction, object recognition etc.

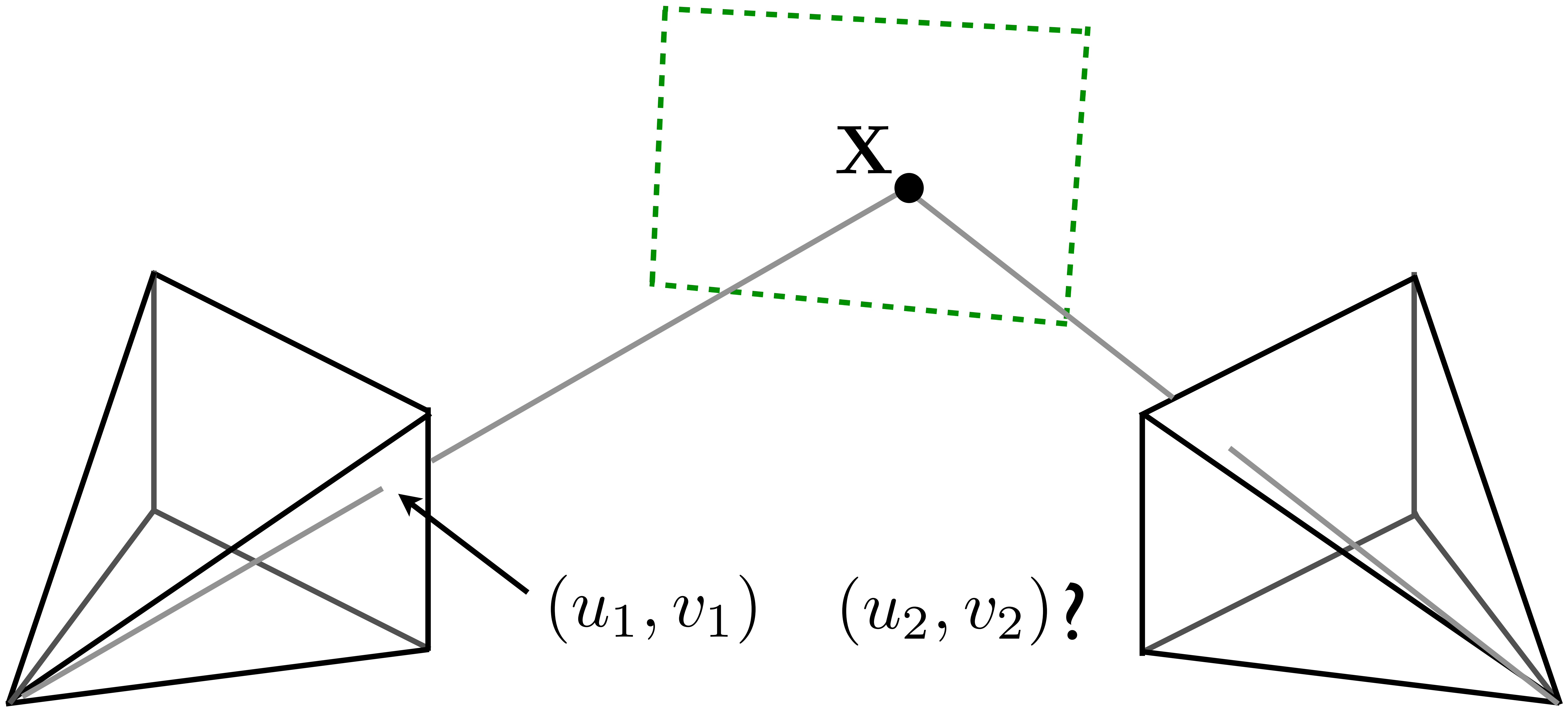
2-view Geometry

- How do we transfer points between 2 views?



2-view Geometry

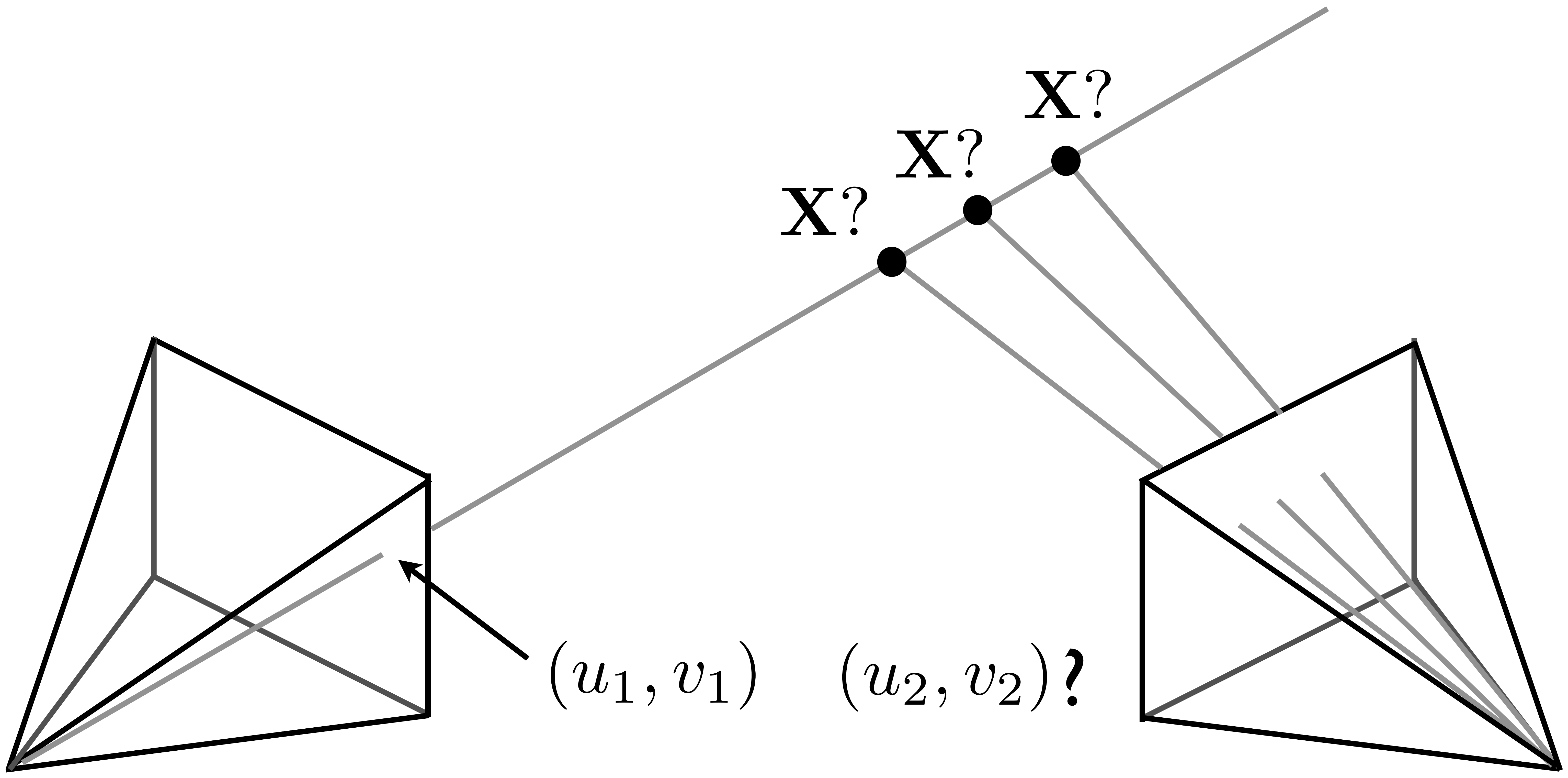
- How do we transfer points between 2 views? (planar case)



Planar case: one-to-one mapping via plane (Homography)

2-view Geometry

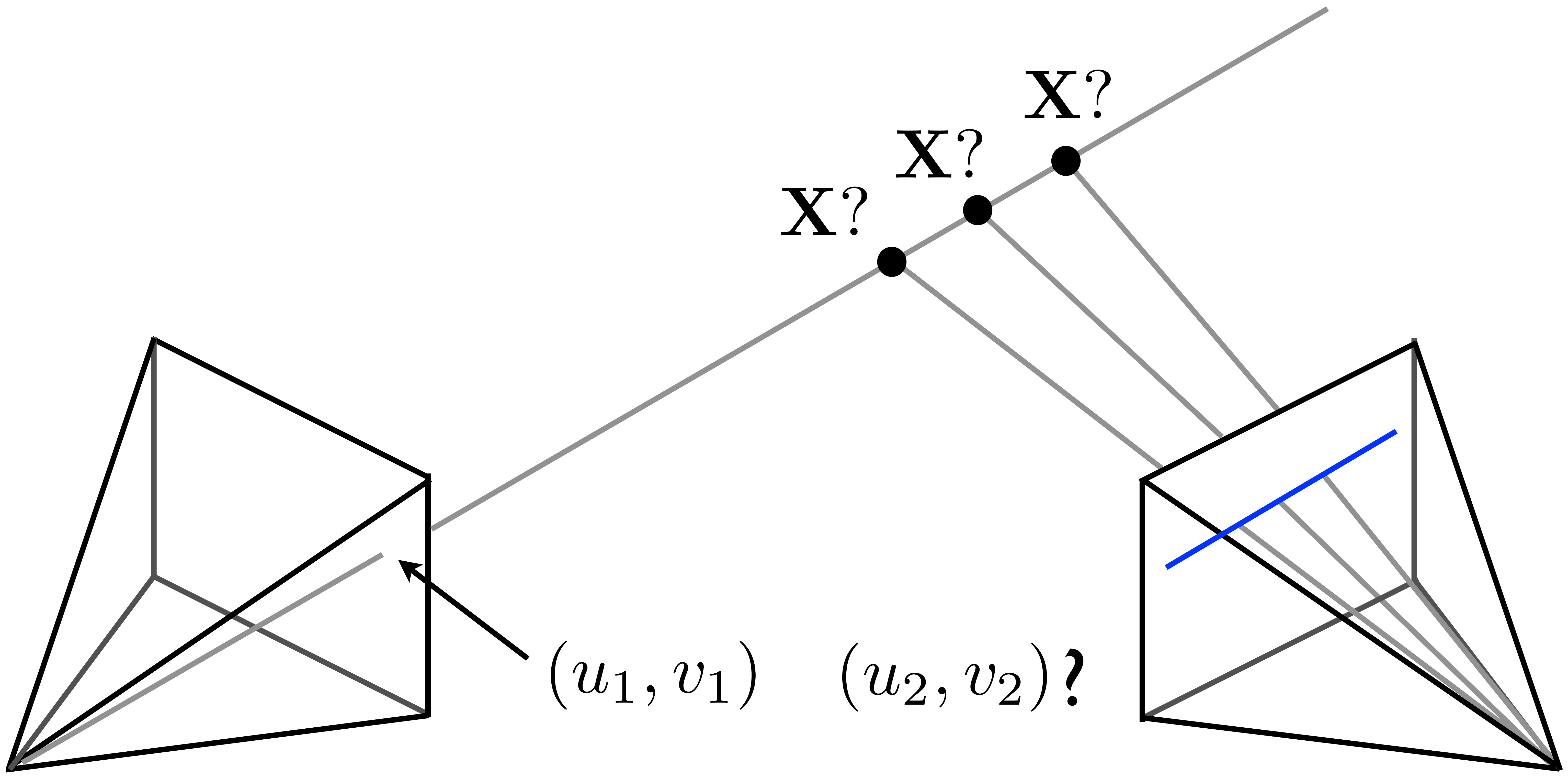
- How do we transfer points between 2 views? (non-planar)



Non-planar case: depends on the depth of the 3D point

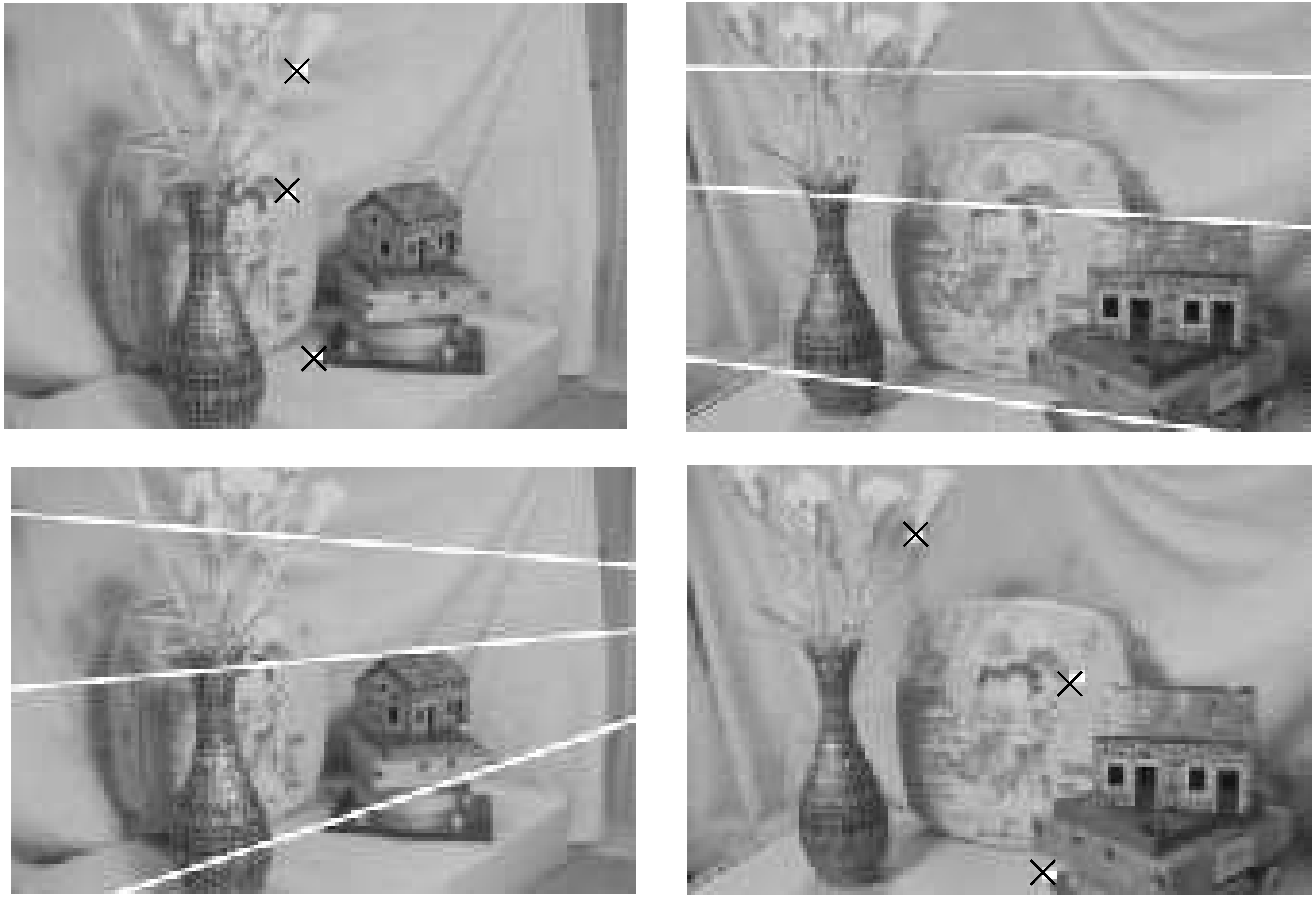
Epipolar Line

- How do we transfer points between 2 views? (non-planar)



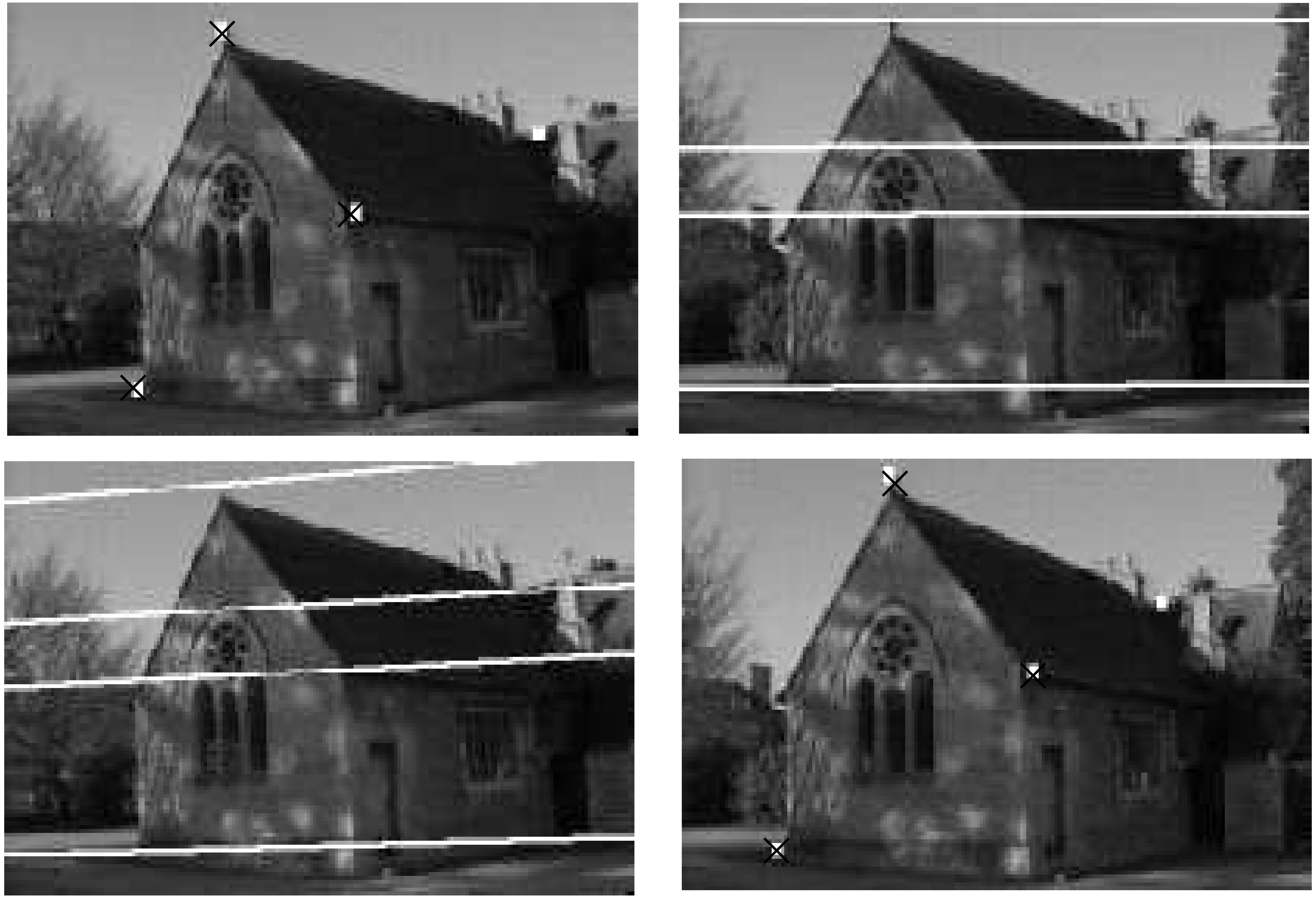
A point in image 1 gives a **line** in image 2

Epipolar Lines



[R. Cipolla]

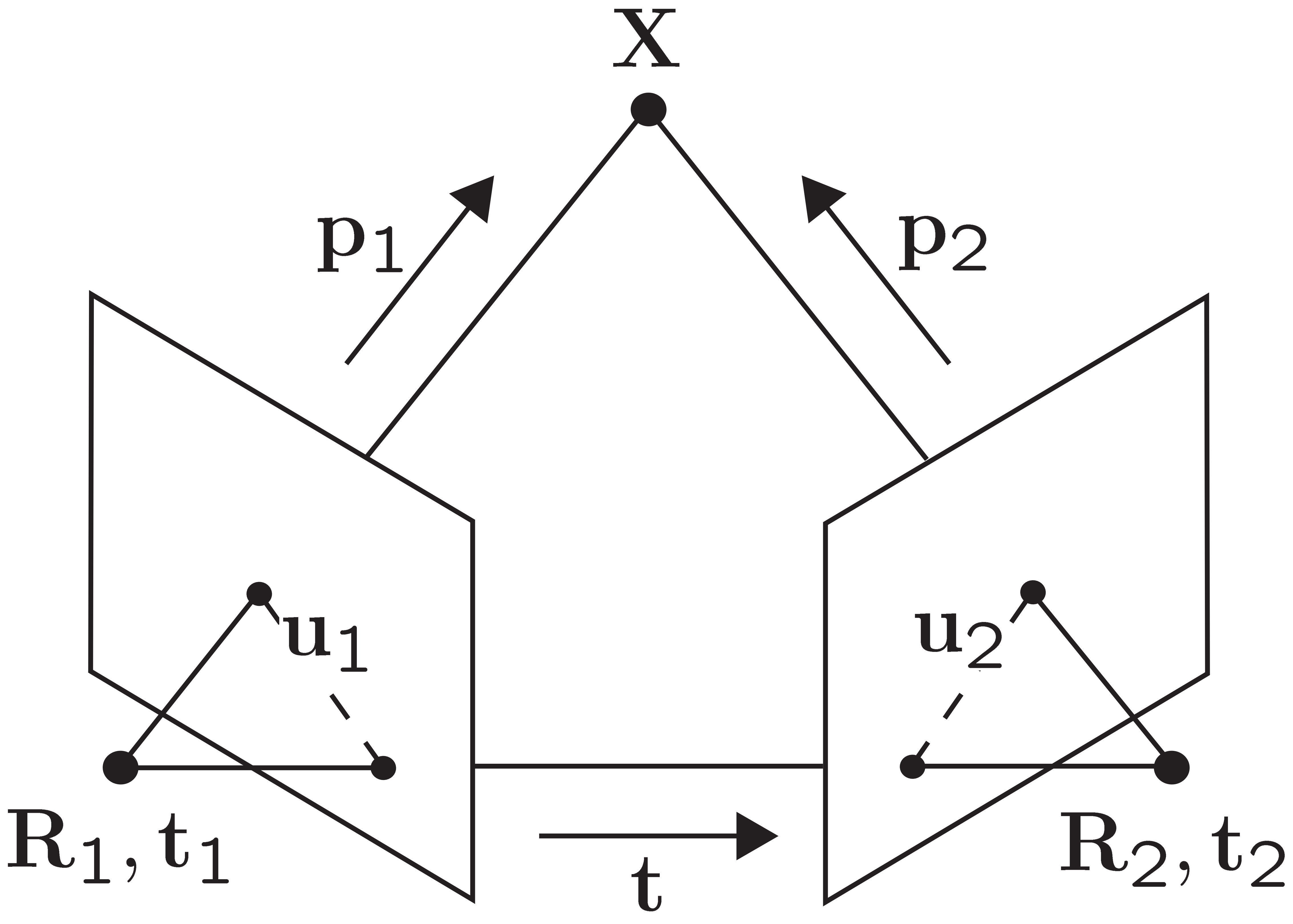
Epipolar Lines



[R. Cipolla]

Epipolar Geometry

- For rays to intersect at a point (X), the two rays and the camera translation must lie in the same plane



14.1

[Szeliski 7.2, 11.1]

Epipolar Geometry

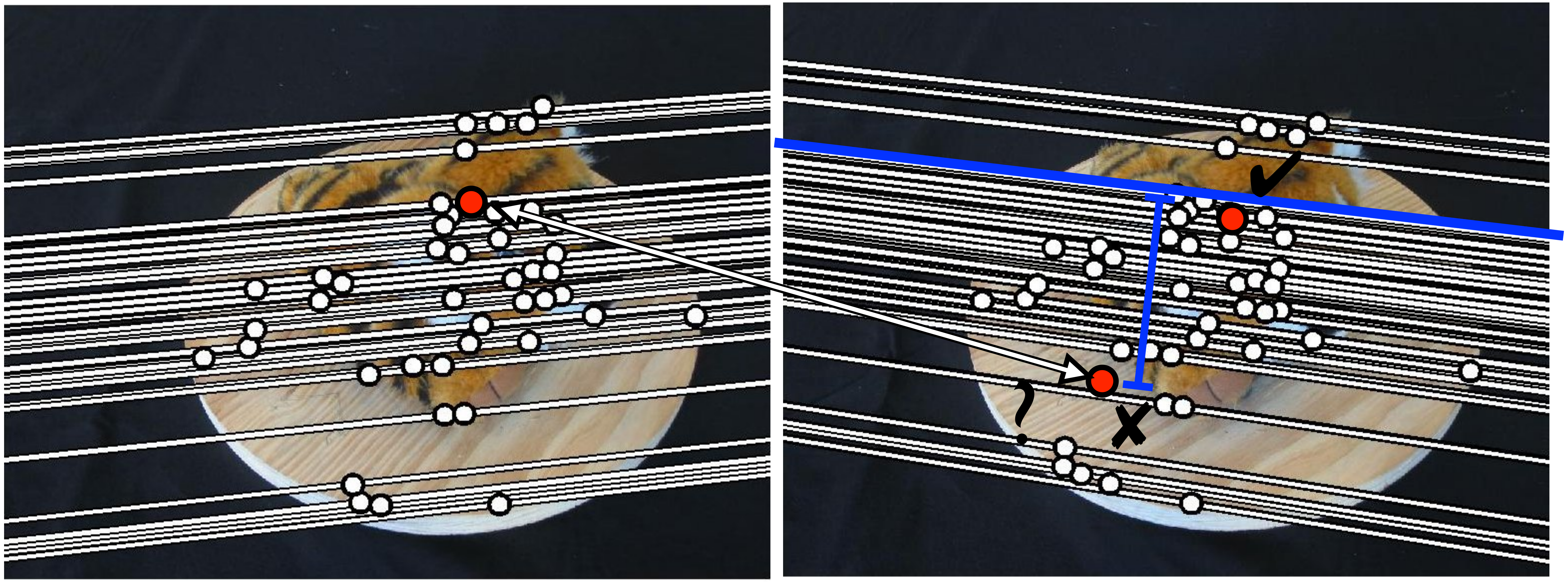
- Example: 2-view matching in 3D

Epipolar Geometry

- Raw SIFT matches

Epipolar Geometry

- Epipolar lines

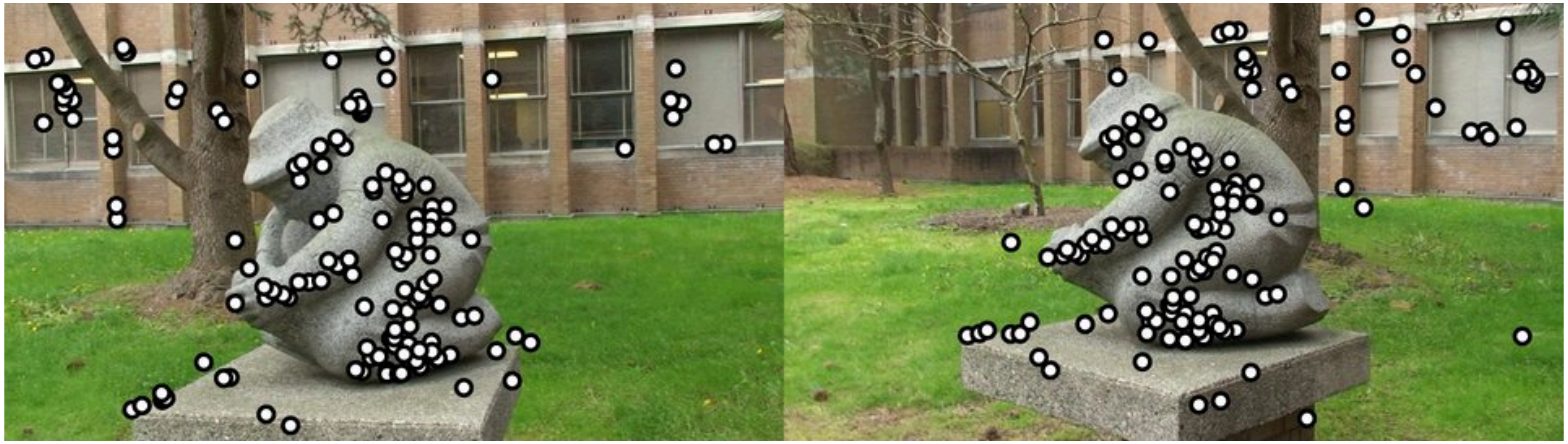


Can use RANSAC to find inliers with small distance from epipolar line

Epipolar Geometry

- Consistent matches

RANSAC for Epipolar Geometry



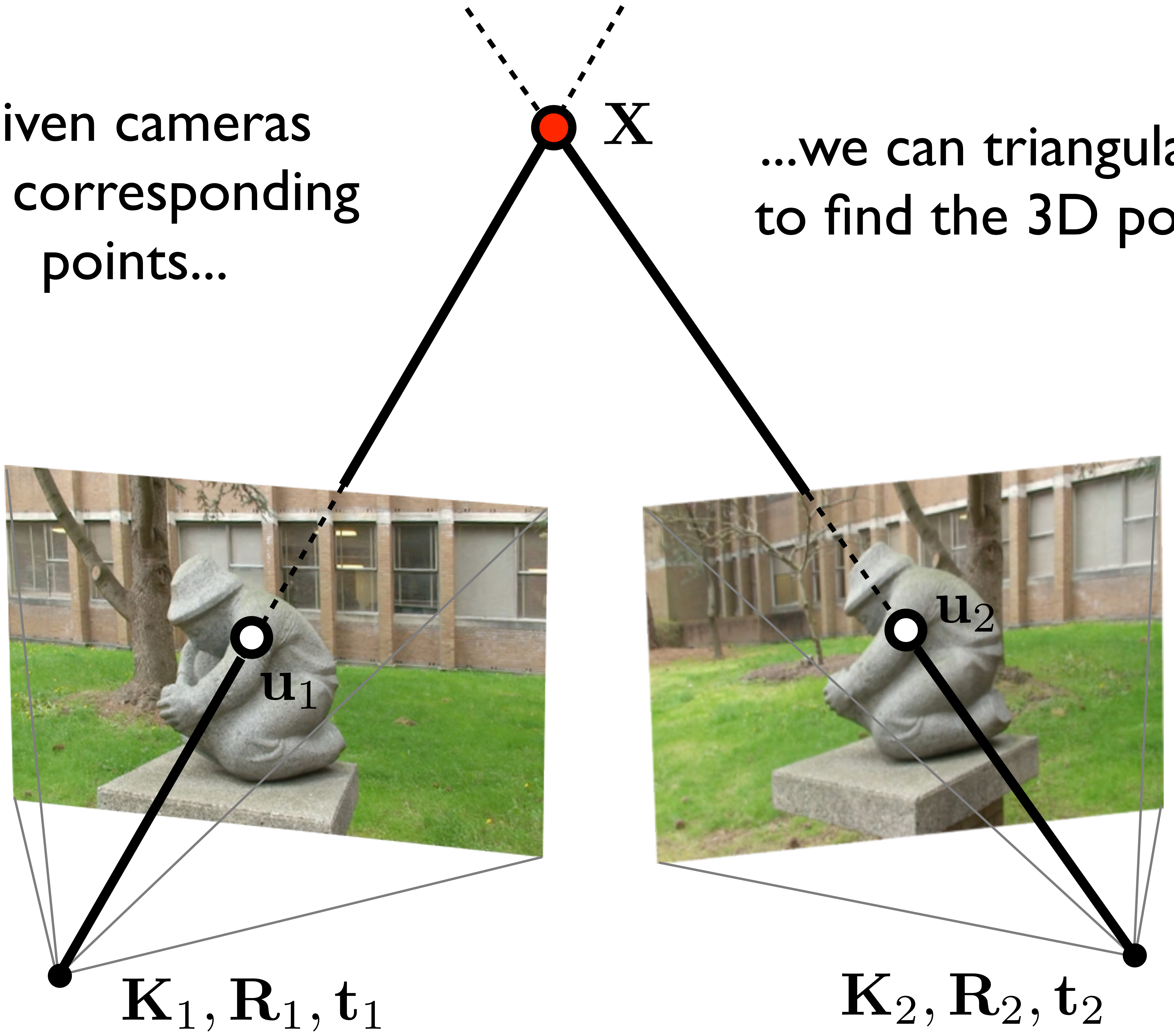
Raw feature matches (after ratio test filtering)

Solve for camera geometry and RANSAC inliers

Triangulation

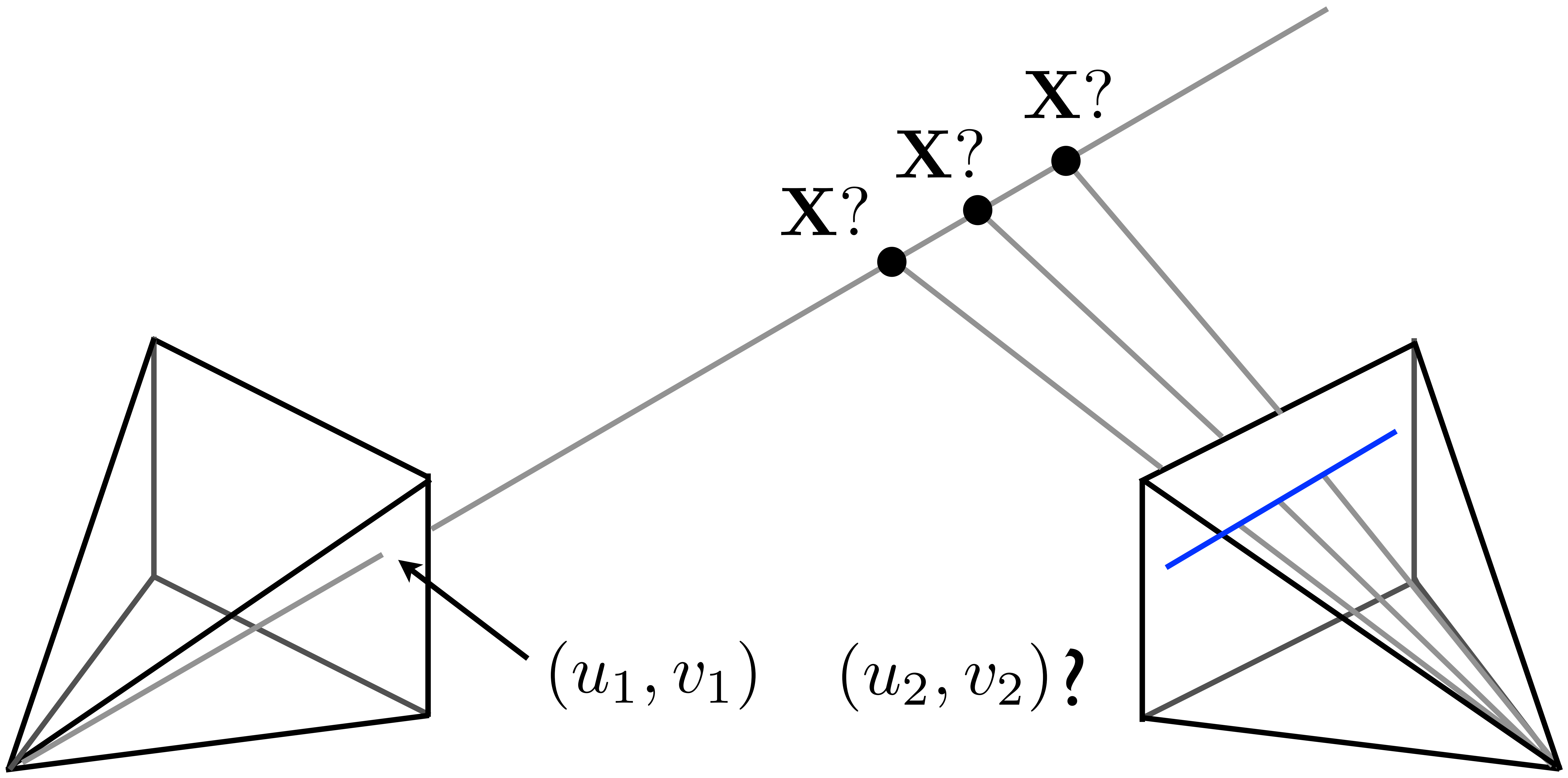
Given cameras and corresponding points...

...we can triangulate to find the 3D point



Epipolar Geometry

- A point in one view may lie on a line in the 2nd



Position in image 2 depends on the **depth** of the 3D point

2-view Stereo

- Camera motion only, points constrained to epipolar lines

1D Search

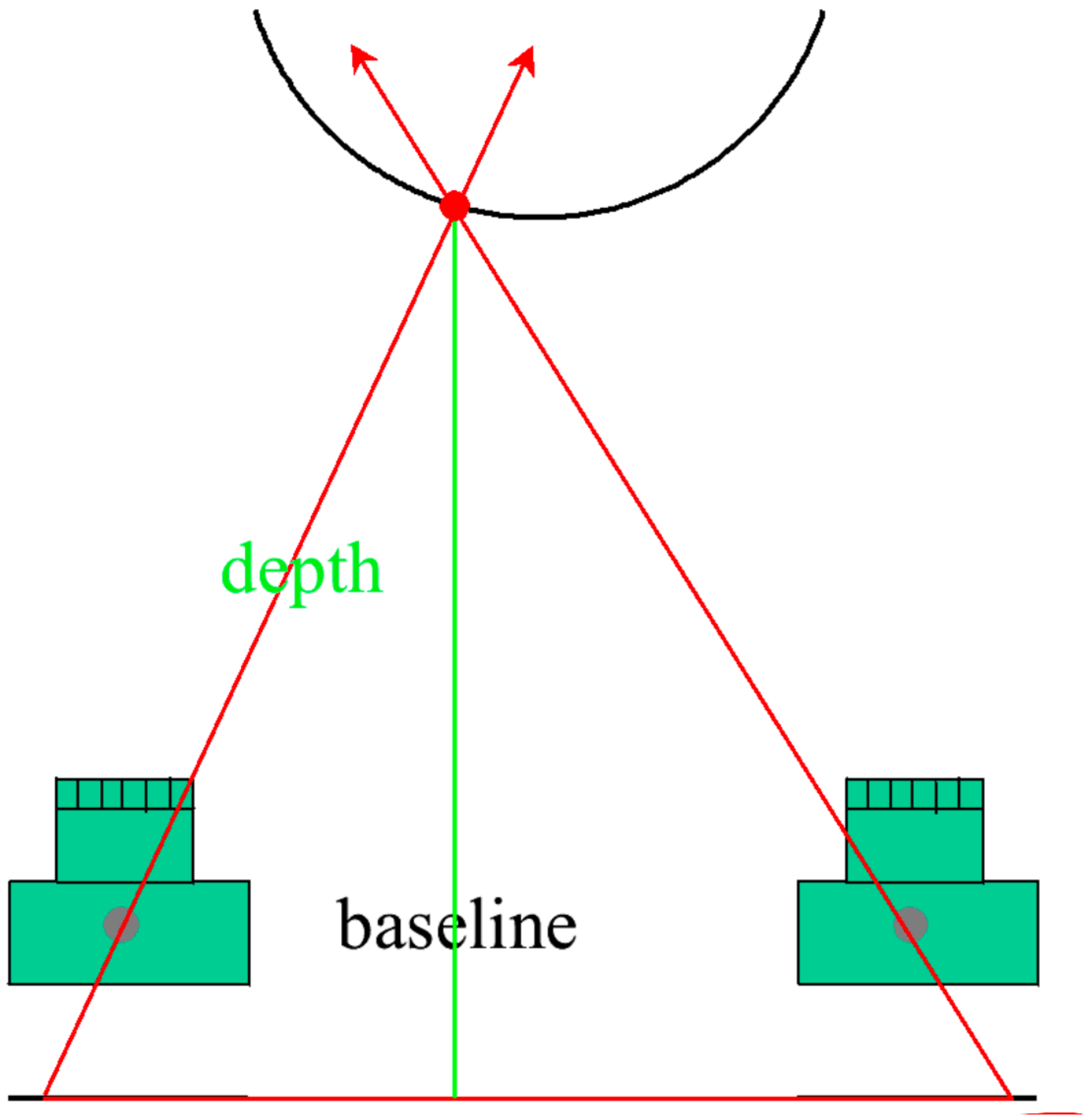
Stereo Camera Configuration

- Humans and many stereo cameras have parallel optical axes

[J. Elson]

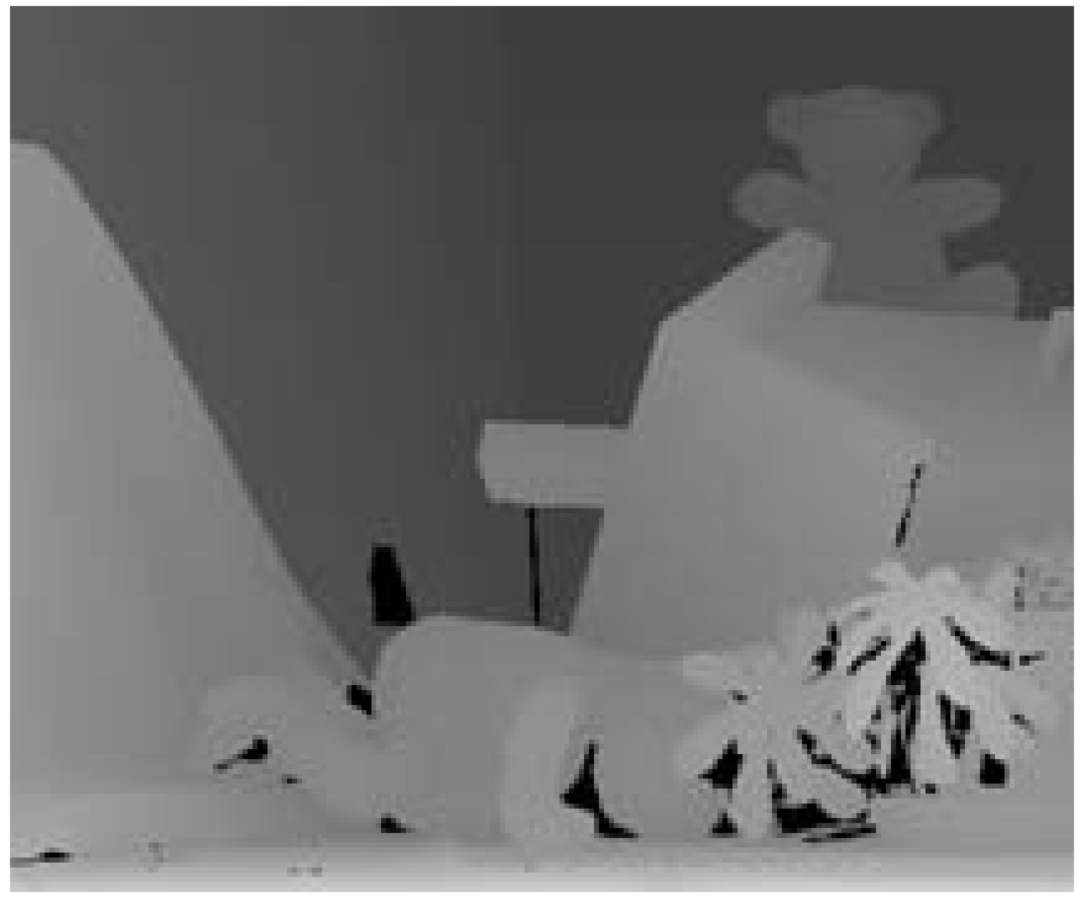
Axis Aligned Stereo

- A common stereo configuration has camera optical axes aligned, with cameras related by a translation in the x direction



Stereo Matching

- In a standard stereo setup, where cameras are related by translation in the x direction, epipolar lines are horizontal

- Stereo algorithms search along scanlines for matches
- Distance along the scanline (difference in x coordinate) for a corresponding feature is called **disparity**

[D. Scharstein]

Disparity and Depth: R

[D. Scharstein]

Disparity and Depth: L

[D. Scharstein]

Disparity and Depth: R+L

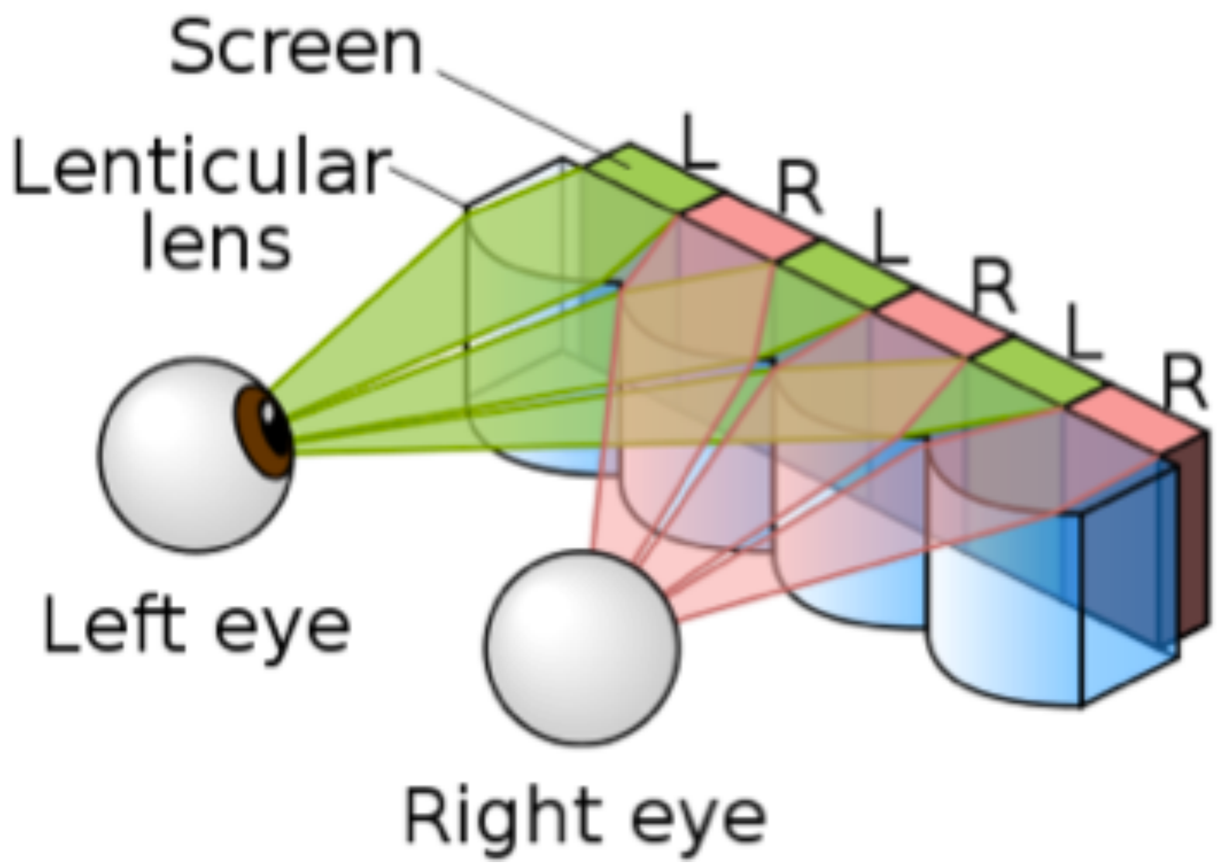
[D. Scharstein]

Anaglyph

- Stereo pair with images encoded in different color channels

Stereo Displays

- Field sequential (shutter) glasses transmit alternate left/right image at 120Hz



Lenticular lenses send different images directly to each eye, without the need for glasses

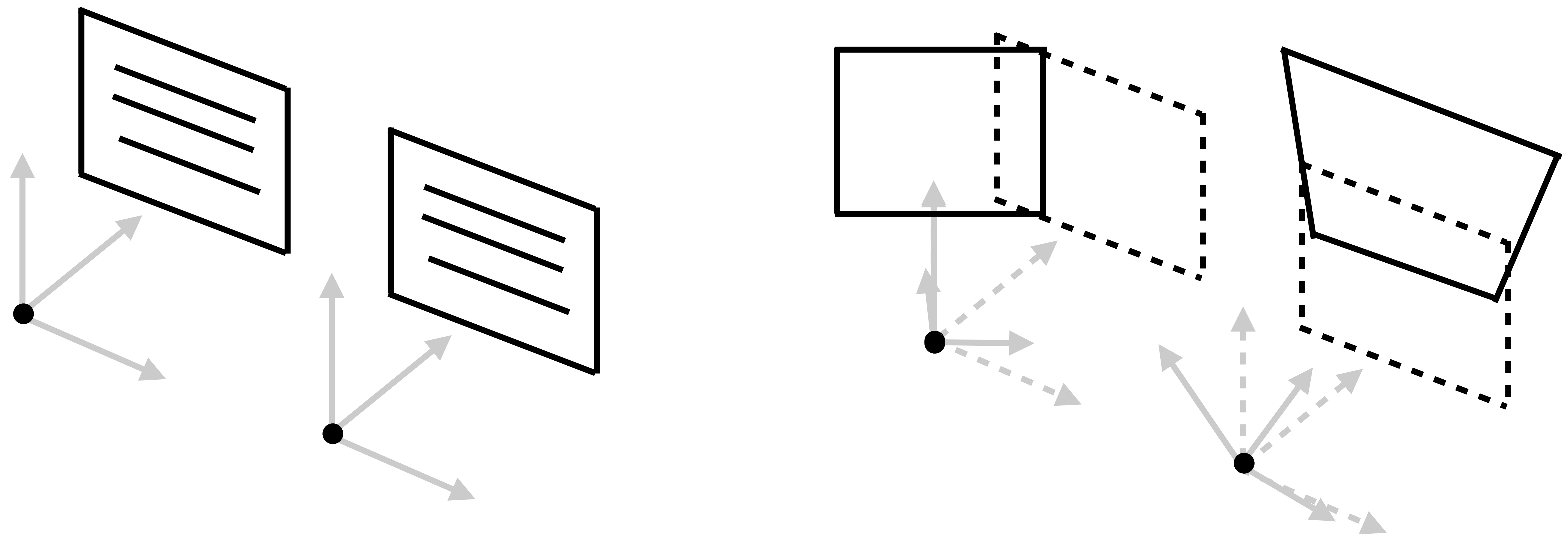
Stereo Displays

- VR headsets send L/R images directly to each eye

[Google Cardboard]

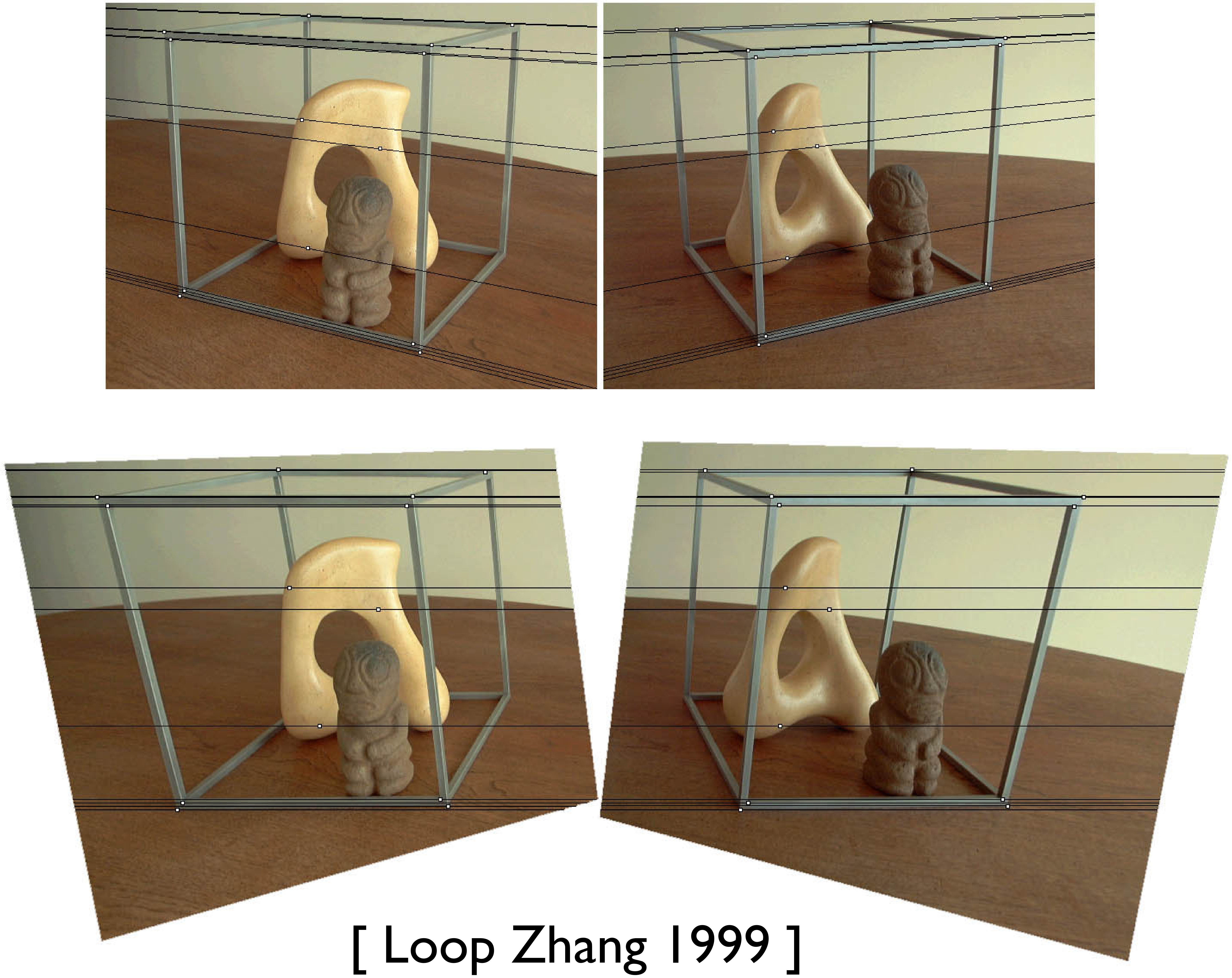
Stereo Rectification

- If the optical axes are not aligned, we can rotate the images (homography) until they are perpendicular to the baseline



Stereo Rectification

- Transform (rotate) images so that epipolar lines are horizontal



[Loop Zhang 1999]

Point Grey Research **Digiclops**

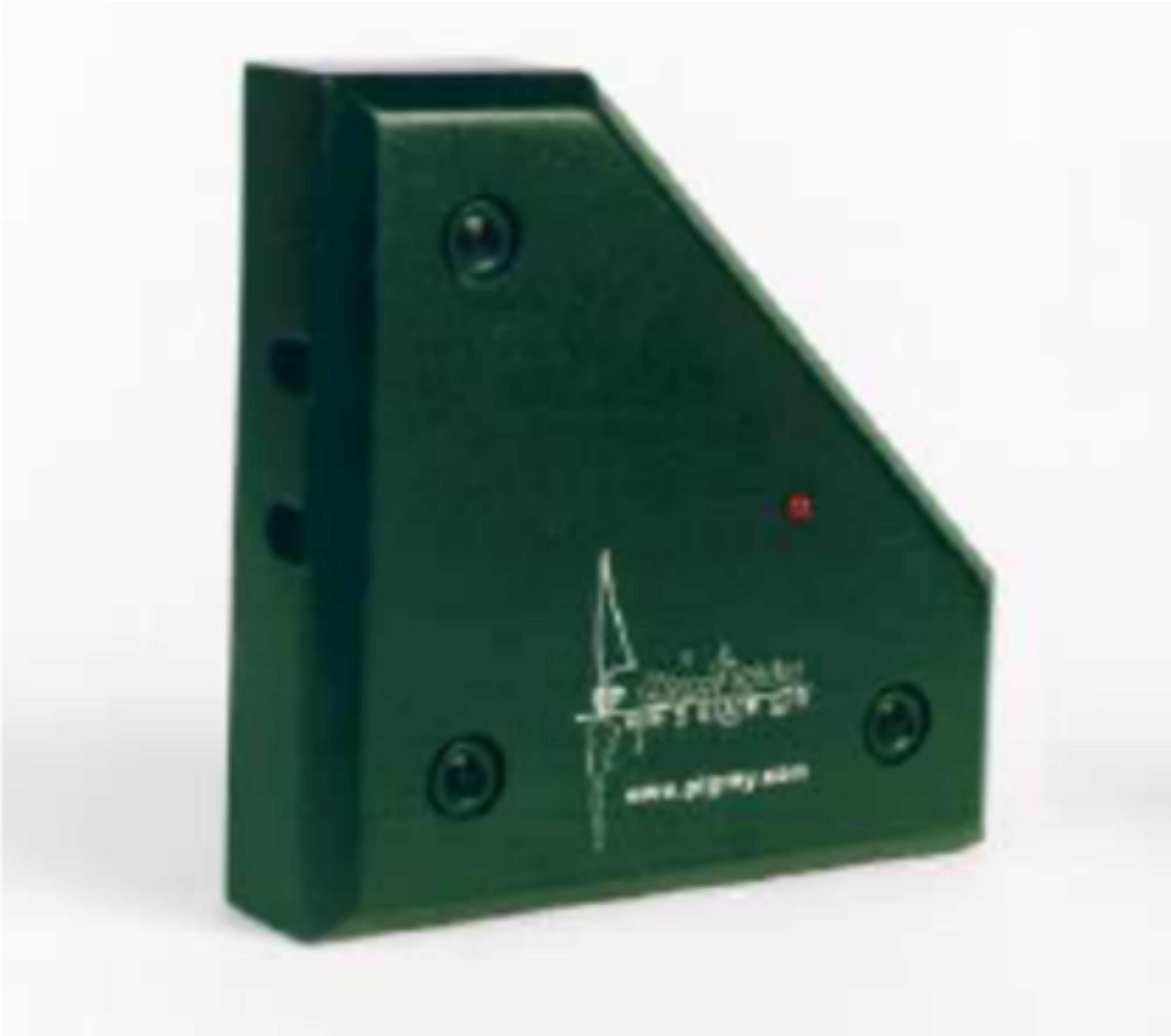


Image credit: Point Grey Research

Stereo Matching

- In a standard stereo setup, where cameras are related by translation in the x direction, epipolar lines are horizontal

- Stereo algorithms search along scanlines for matches
- Distance along the scanline (difference in x coordinate) for a corresponding feature is called **disparity**

[D. Scharstein]

Matching along a Scanline

$\mathbf{w}_L$ and $\mathbf{w}_R$ are corresponding $m \times m$ windows of pixels

Define a distance function between image patches, e.g.,

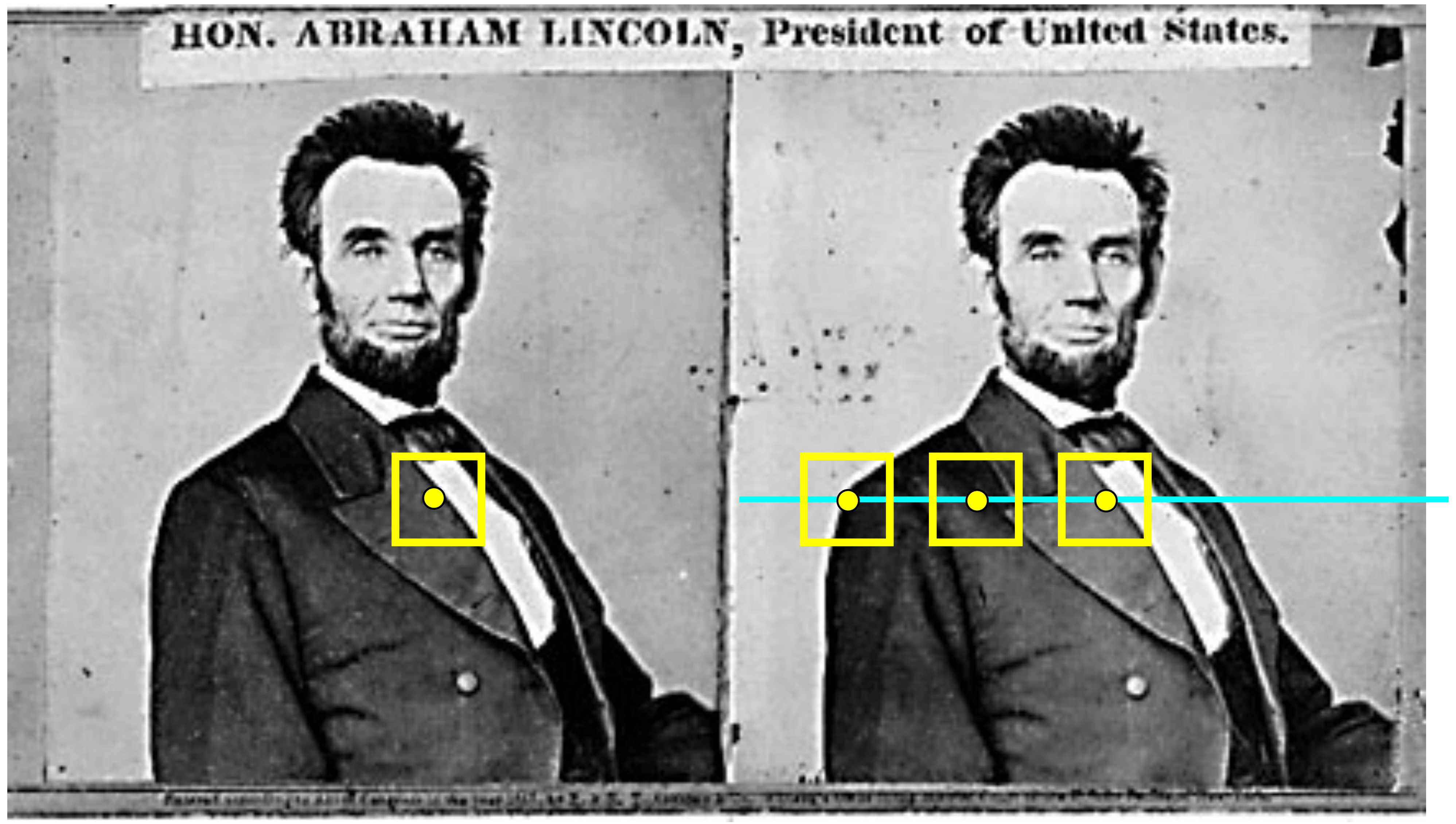
$$\text{SSD} = ||\mathbf{w}_L - \mathbf{w}_R(d)||^2$$

$$\text{correlation} = \mathbf{w}_L \cdot \mathbf{w}_R(d) = \cos\theta$$

Matching along a Scanline

Best match is at minimum of SSD function along a scanline

(simple) **Stereo** Algorithm



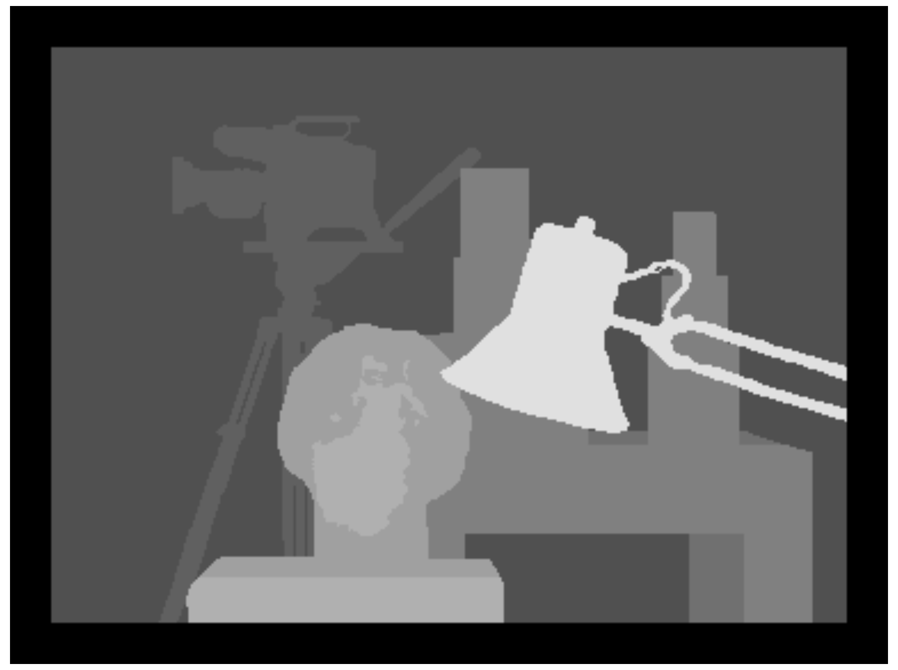
1. Rectify images
   (make epipolar lines horizontal)
2. For each pixel in image 1
   a. Search along epipolar line in image 2
   b. Find best match and record offset = disparity
   c. Compute depth from disparity

$$Z = f\frac{\Delta X}{disp}$$

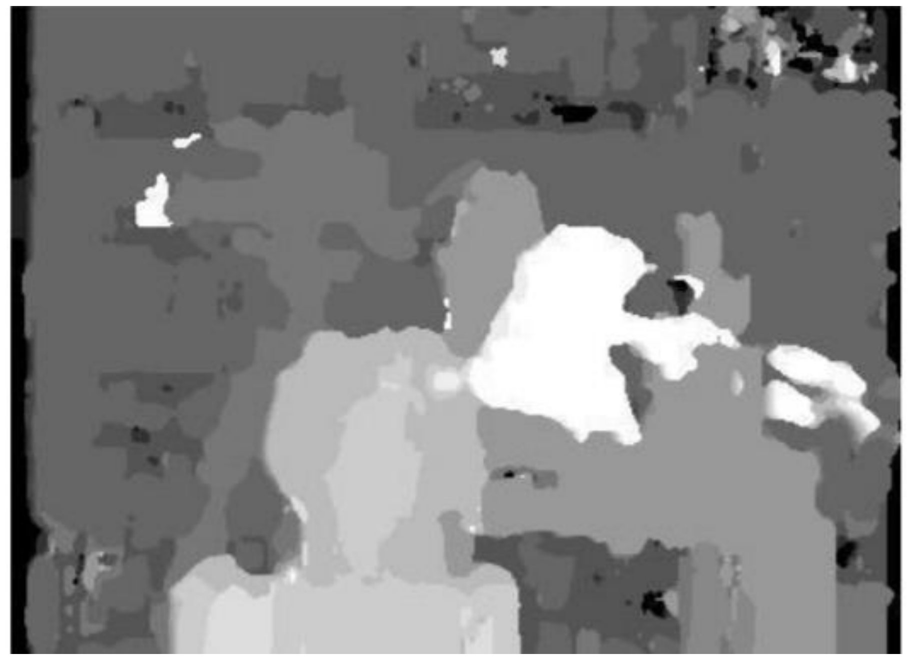
Slide Credit: Ioannis (Yannis) Gkioulekas (CMU)

Effect of Window Size

- Larger windows → smoothed result



W=3



W=11

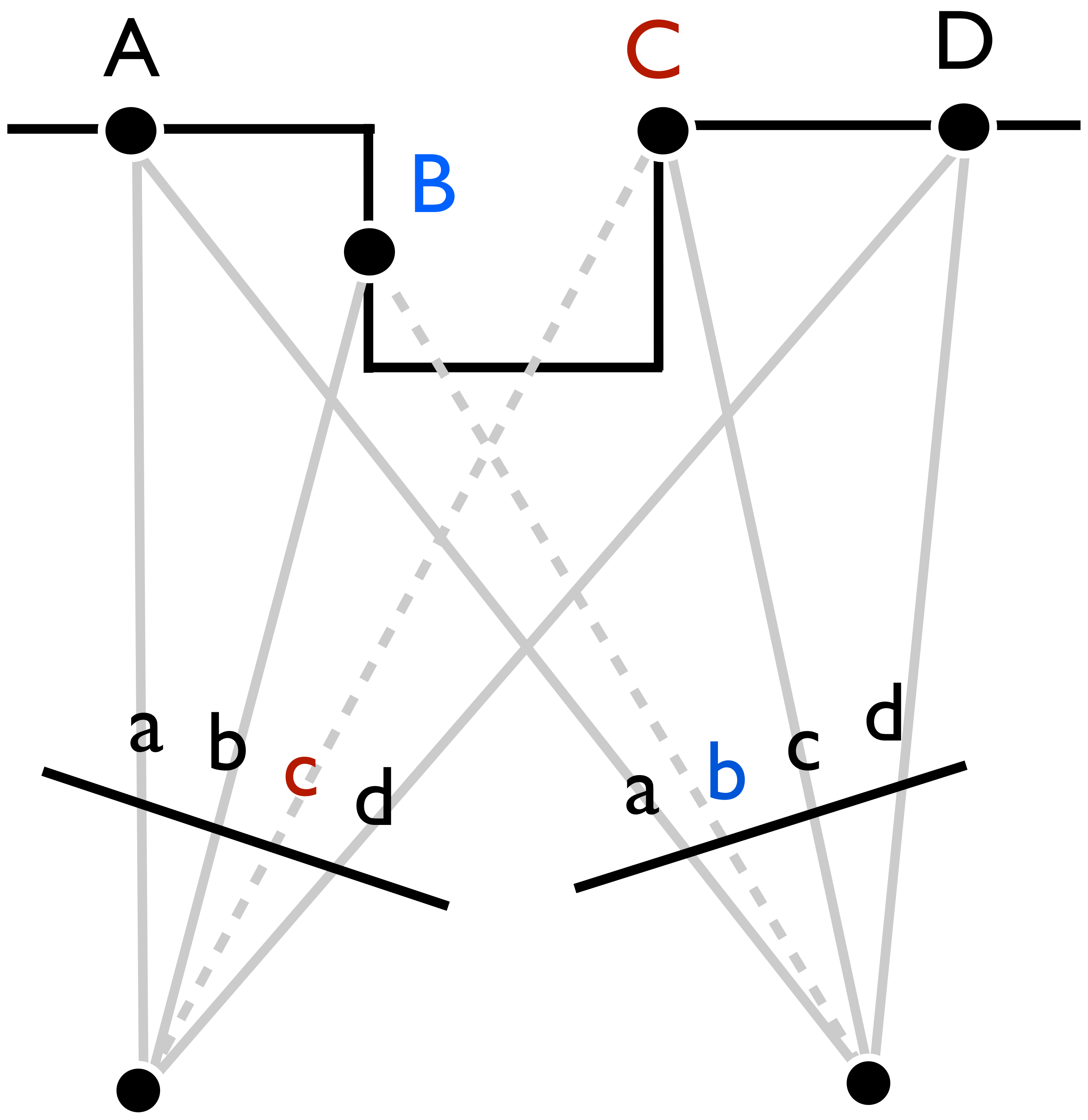
W=25

Smaller window
+ More detail
- More noise

Larger window
+ Smoother disparity maps
- Less detail
- Fails near boundaries

Occlusions

- Sometimes a point in image 1 does not appear in image 2, or vice-versa (this is called an **occlusion**)



- Occlusions cause gaps in the stereo reconstruction
- + Matching is difficult nearby as aggregation windows often overlap the occluded region

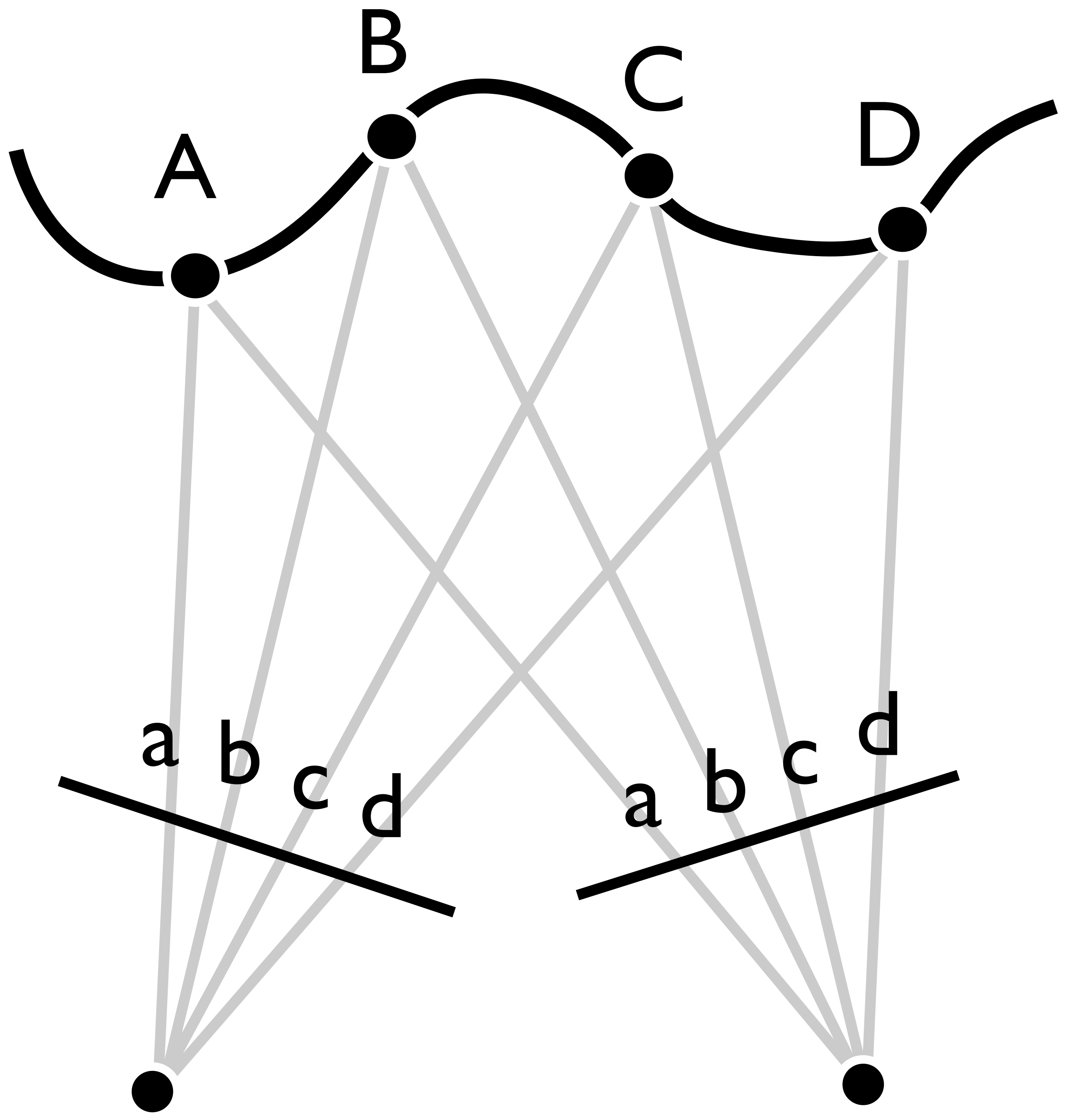
Edge Aware Stereo

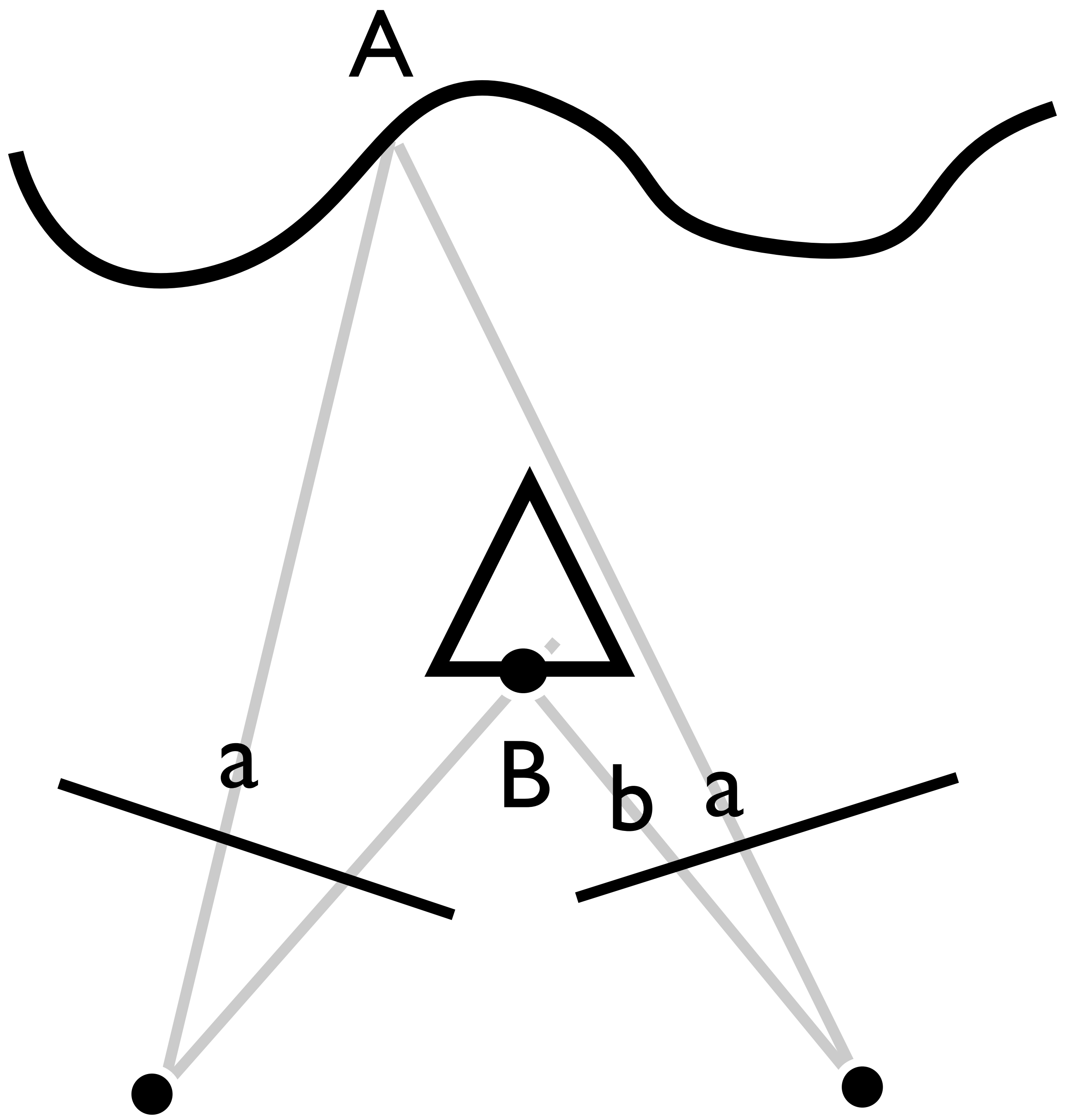
- Occlusions and depth discontinuities cause problems for stereo matching, as aggregation windows overlap multiple depths

- Segmentation-based stereo approaches aim to solve this by trying to guess the depth edges (e.g., joint segmentation and depth estimation [Taguchi et al 2008])

Ordering Constraint

- If point B is to the right of point A in image 1, the same is *usually* true in image 2

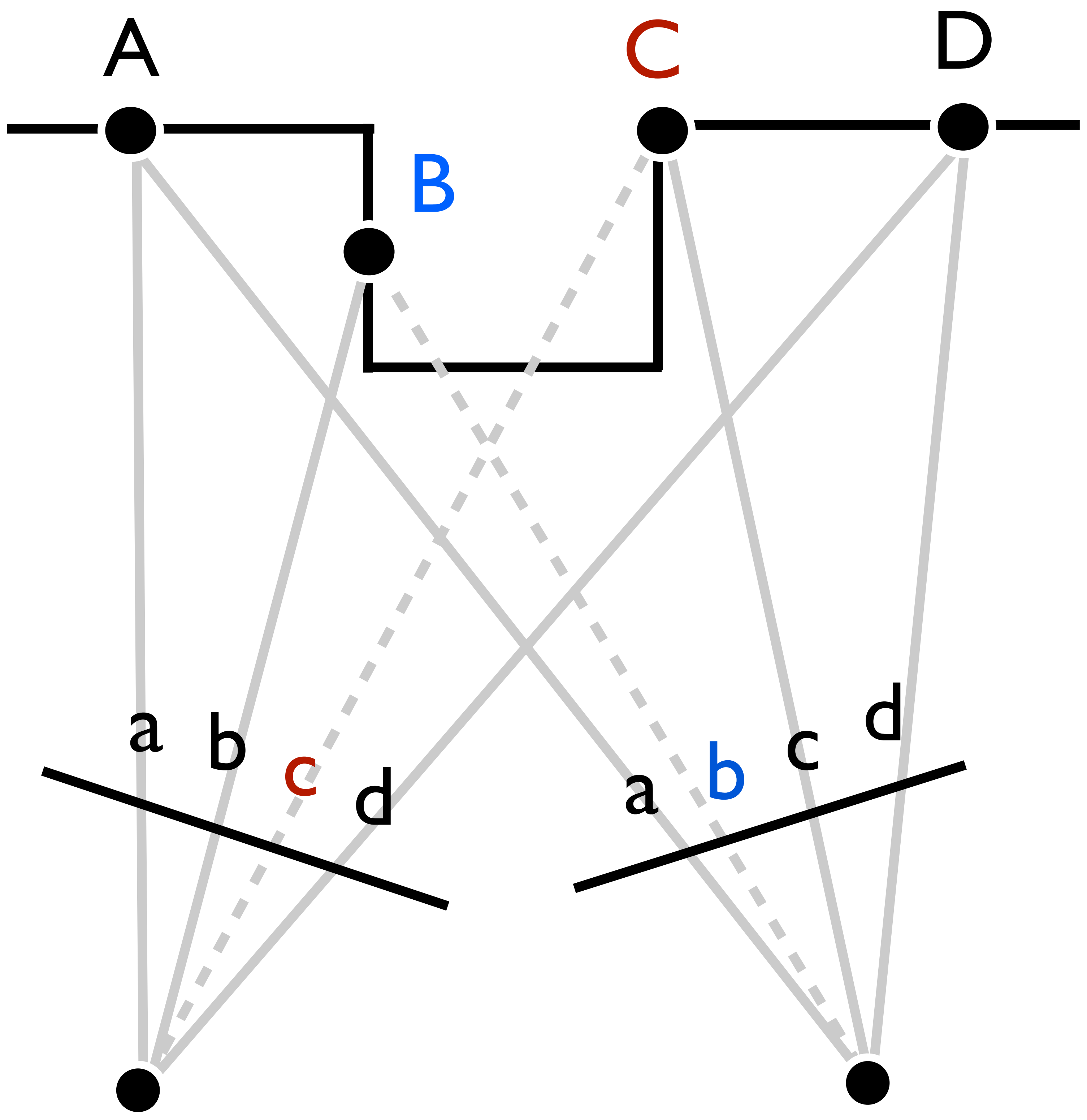




Not always, e.g., if an object is wholly within the ray triangle generated by A

Occlusions + Ordering

- Note that the ordering constraint is still maintained in the presence of occlusions (unless there is an object off surface as in the previous slide)



Stereo Cost Functions

- Energy function for stereo matching based on disparity d(x,y)
- Sum of data and smoothness terms

$$E(d) = E_d(d) + \lambda E_s(d)$$

- Data term is cost of pixel x,y allocated disparity d (e.g., SSD)

$$E_d(d) = \sum_{(x,y)} C(x, y, d(x, y))$$

- Smoothness cost penalises disparity changes with robust ρ(.)

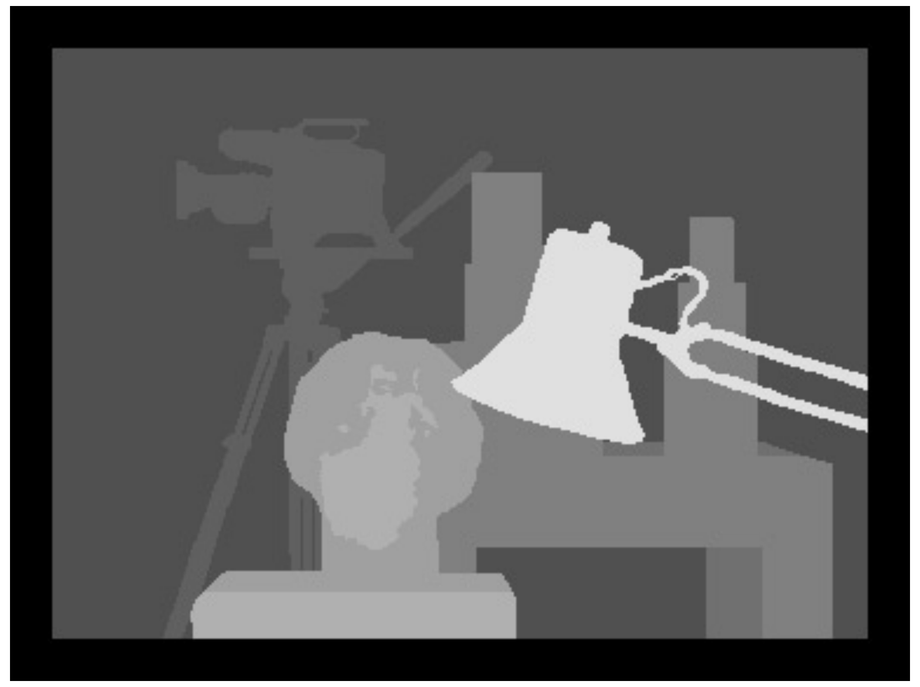
$$E_s(d) = \sum_{(x,y)} \rho(d(x, y) - d(x + 1, y)) + \rho(d(x, y) - d(x, y + 1))$$

- This is a Markov Random Field (MRF), which can be solved using techniques such as Graph Cuts

[Szeliski B5]

Stereo Comparison

- Global vs Scanline vs Local optimization



Ground truth

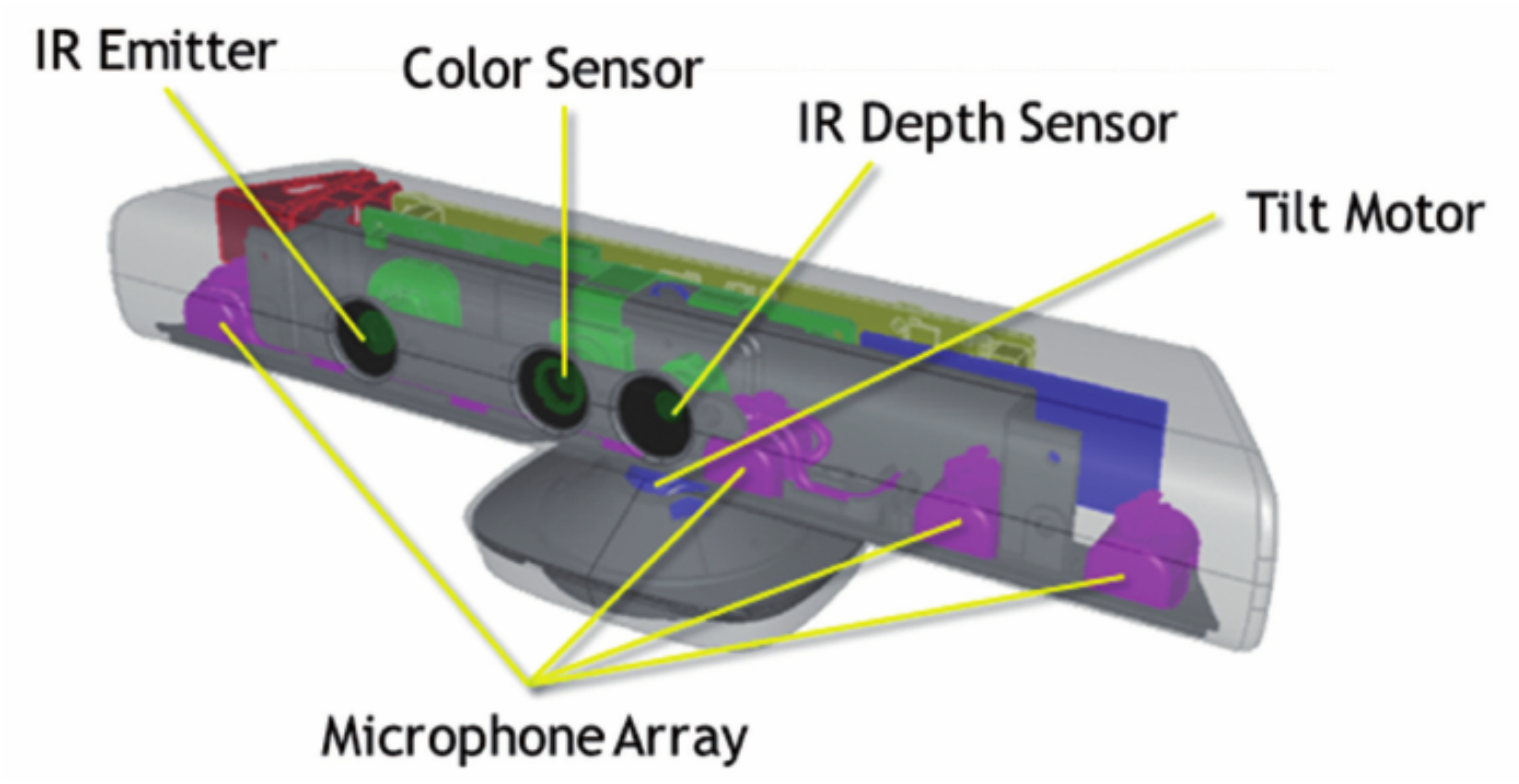
Graph Cuts
[Kolmogorov Zabih 2001]

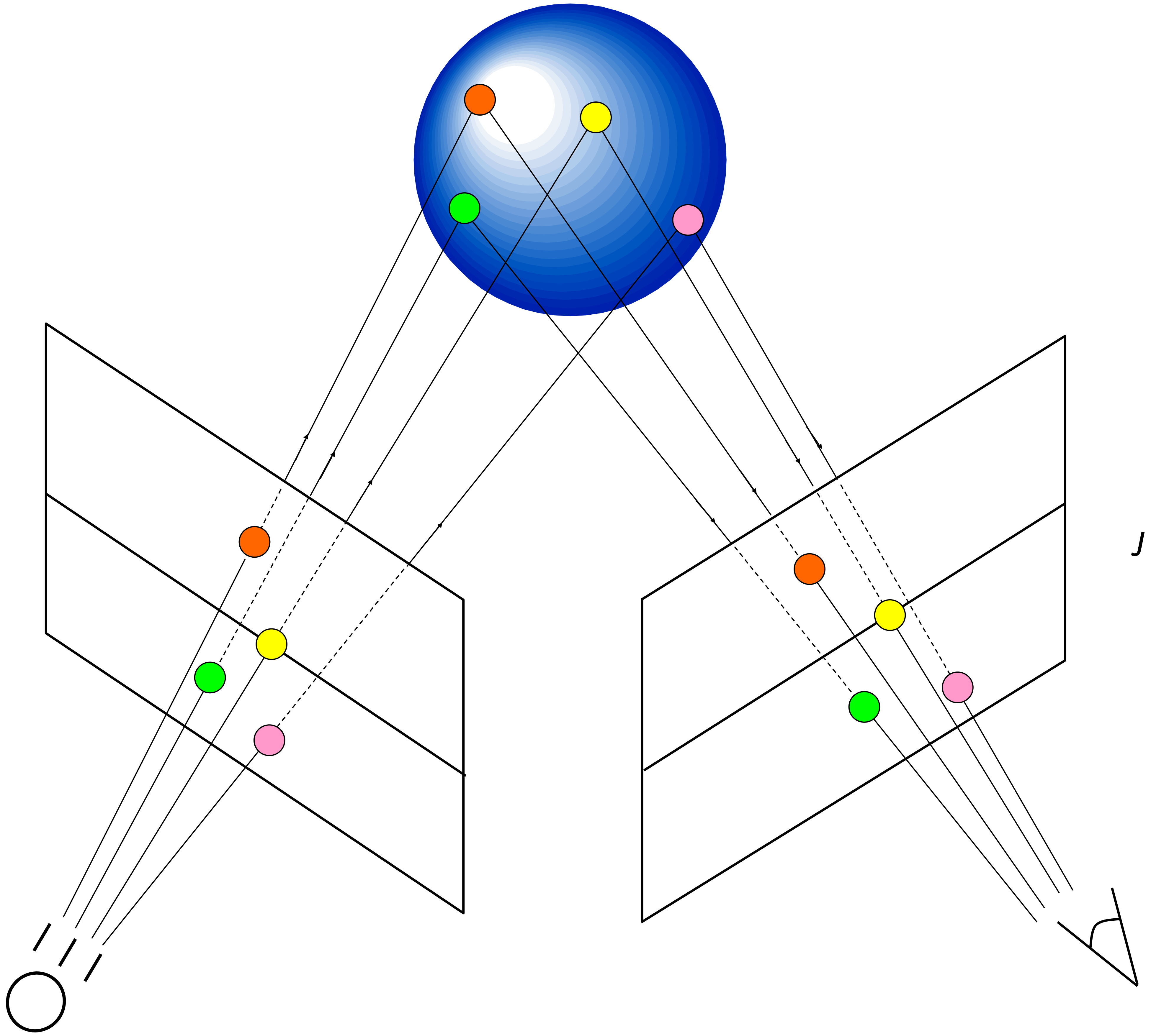
Dynamic Programming

SSD 21px aggregation

[Scharstein Szeliski 2002]

Application: Microsoft **Kinect**





Projector
(NIR dot pattern)

Camera

Stereo Vision Summary

With two eyes, we acquire images of the world from slightly different viewpoints

We perceive **depth** based on **differences in the relative position of points** in the left image and in the right image

Stereo algorithms work by finding **matches** between points along corresponding lines in a second image, known as epipolar lines.

A point in one image projects to an **epipolar line** in a second image

In an axis-aligned / rectified stereo setup, matches are found along horizontal scanlines

Menu for Today

Topics:

- 3D Correspondence, **Epipolar** Geometry
- **Stereo** Vision
- **Quiz 4**

Readings:

- **Today's** Lecture: Szeliski 12.1, 12.3-12.4, 9.3

Reminders:

- **Assignment 4**: RANSAC and Panoramas due **November 9th**