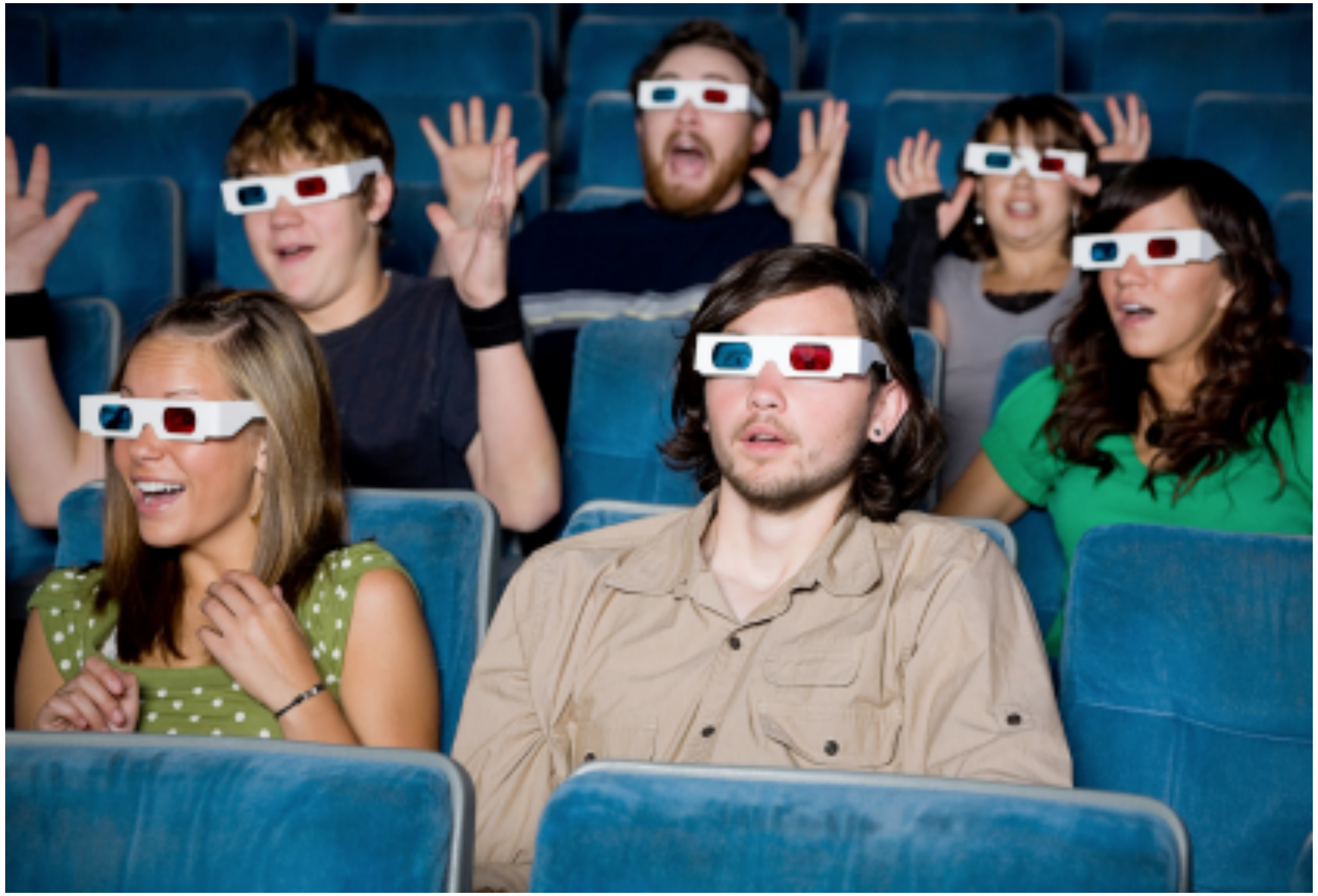
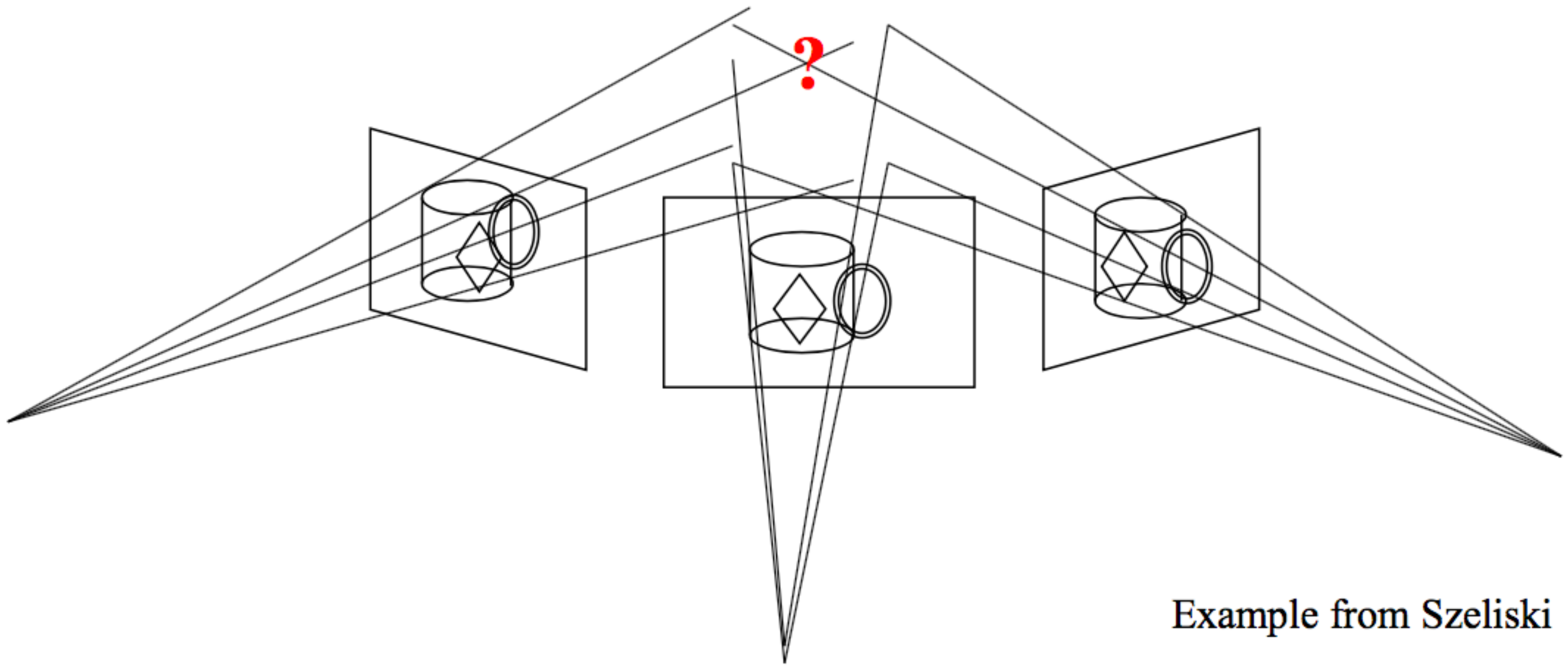
Stereo

Gary Overett (Slides adapted from CMU 16-720 2014)

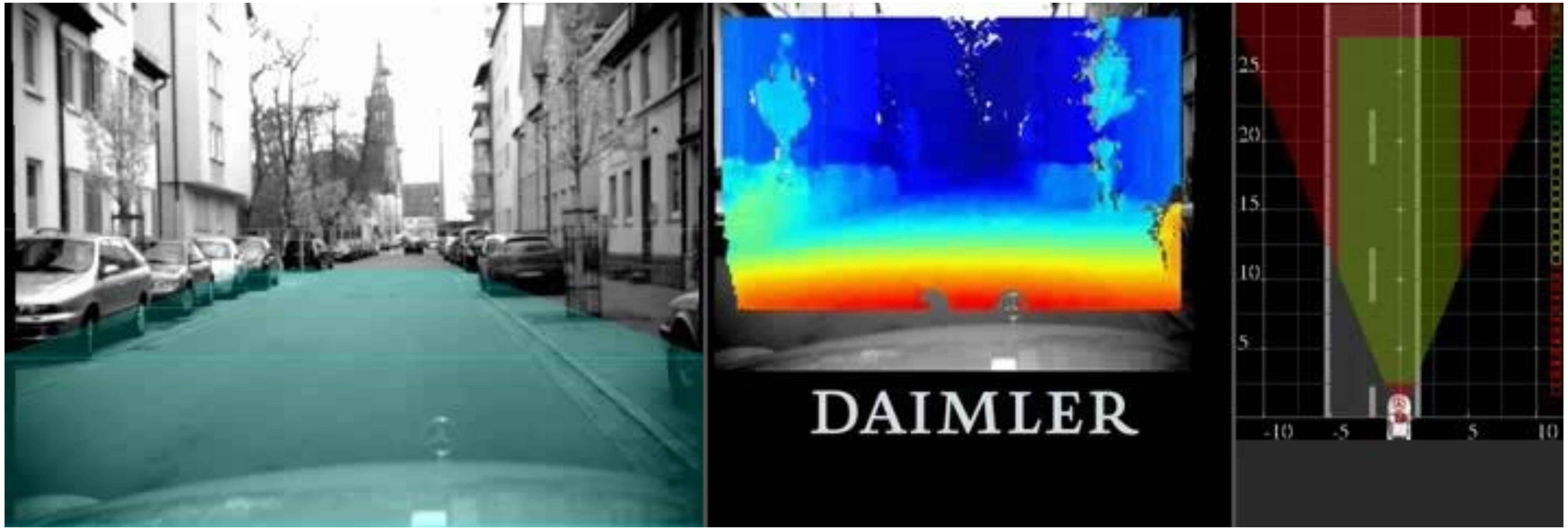


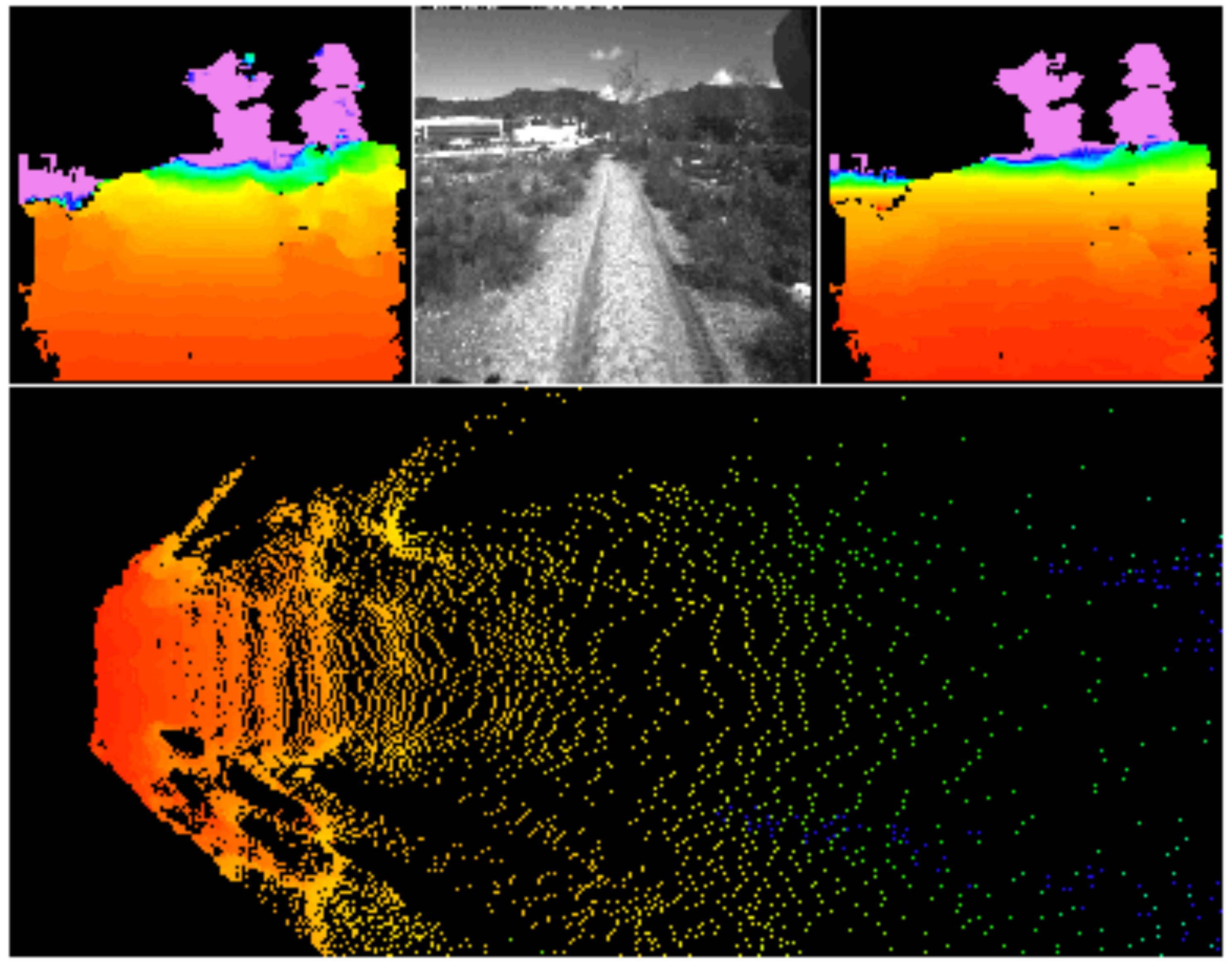
Dense Stereo



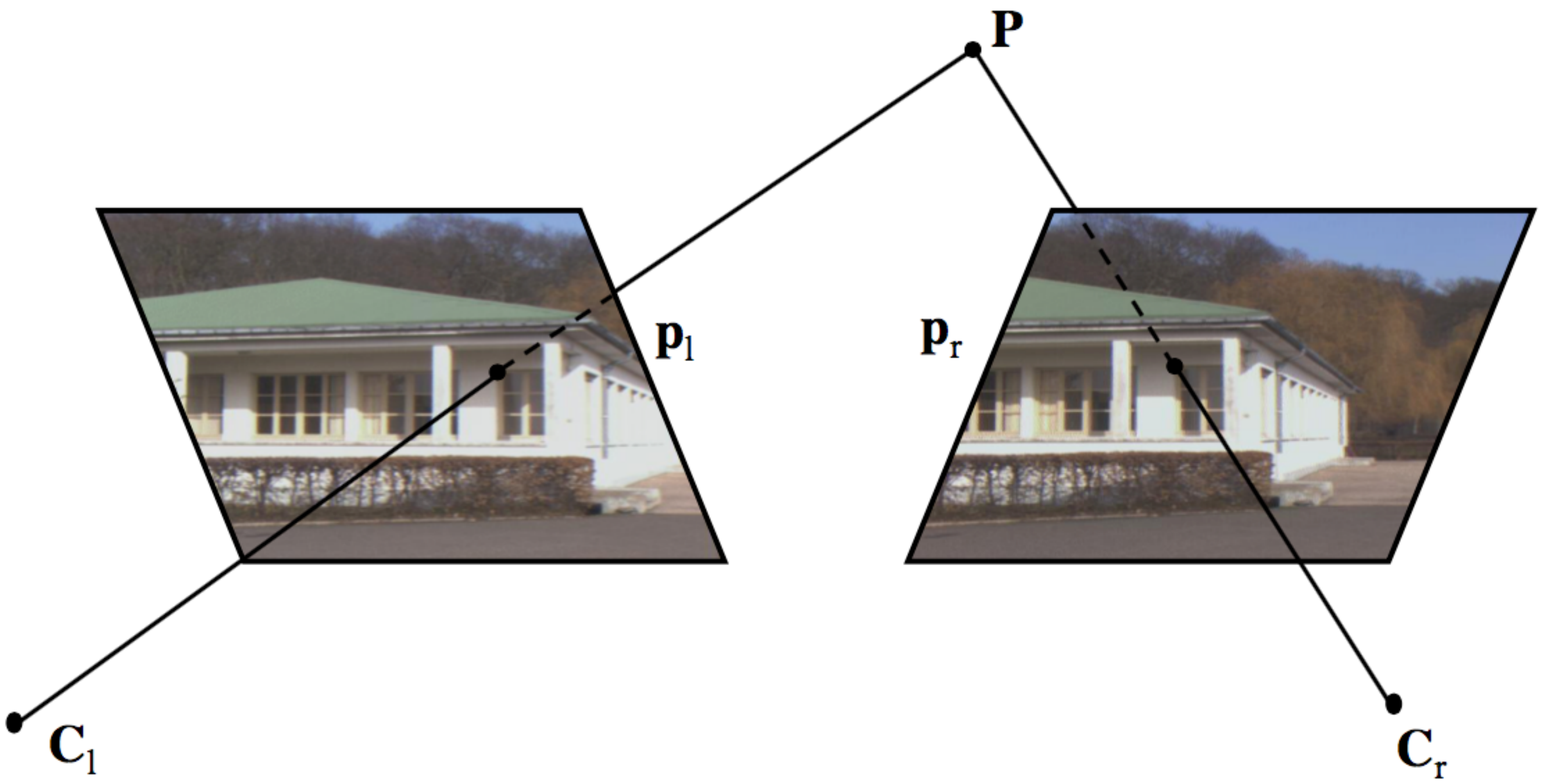
- Reconstruct the 3D position of the points corresponding to (all the) pixels in a set of images.
- Key assumption : We know the relative position, orientation, K, of all the cameras.
- Assume first 2 views only.

Examples



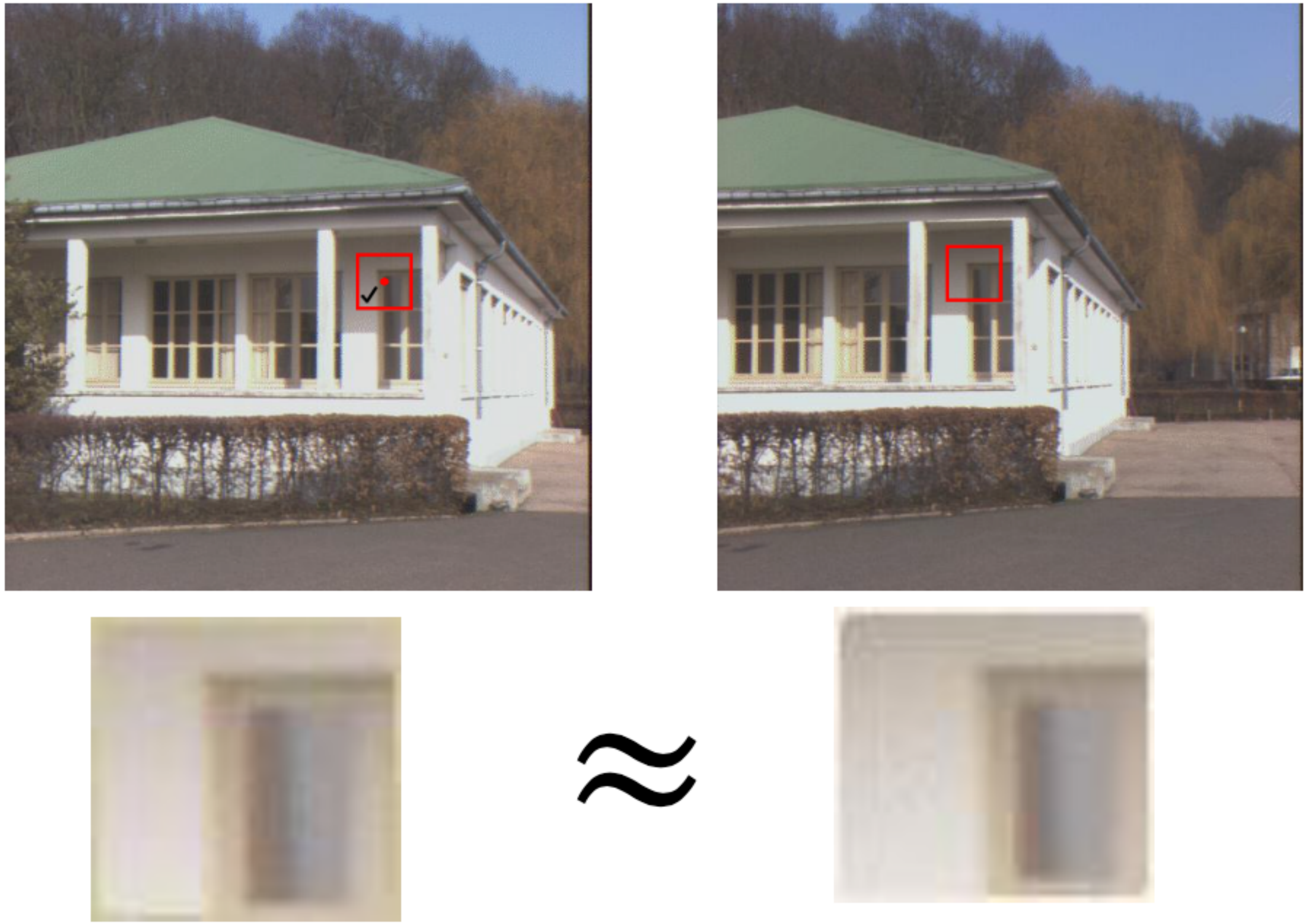


Stereo Vision = Correspondences + Reconstruction



- *Correspondence*: Given a point $\mathbf{p}_l$ in one image, find the corresponding point in the other image.
- *Reconstruction*: Given a correspondence ($\mathbf{p}_l$,$\mathbf{p}_r$), compute the 3-D coordinates of the corresponding point in space, **P**.

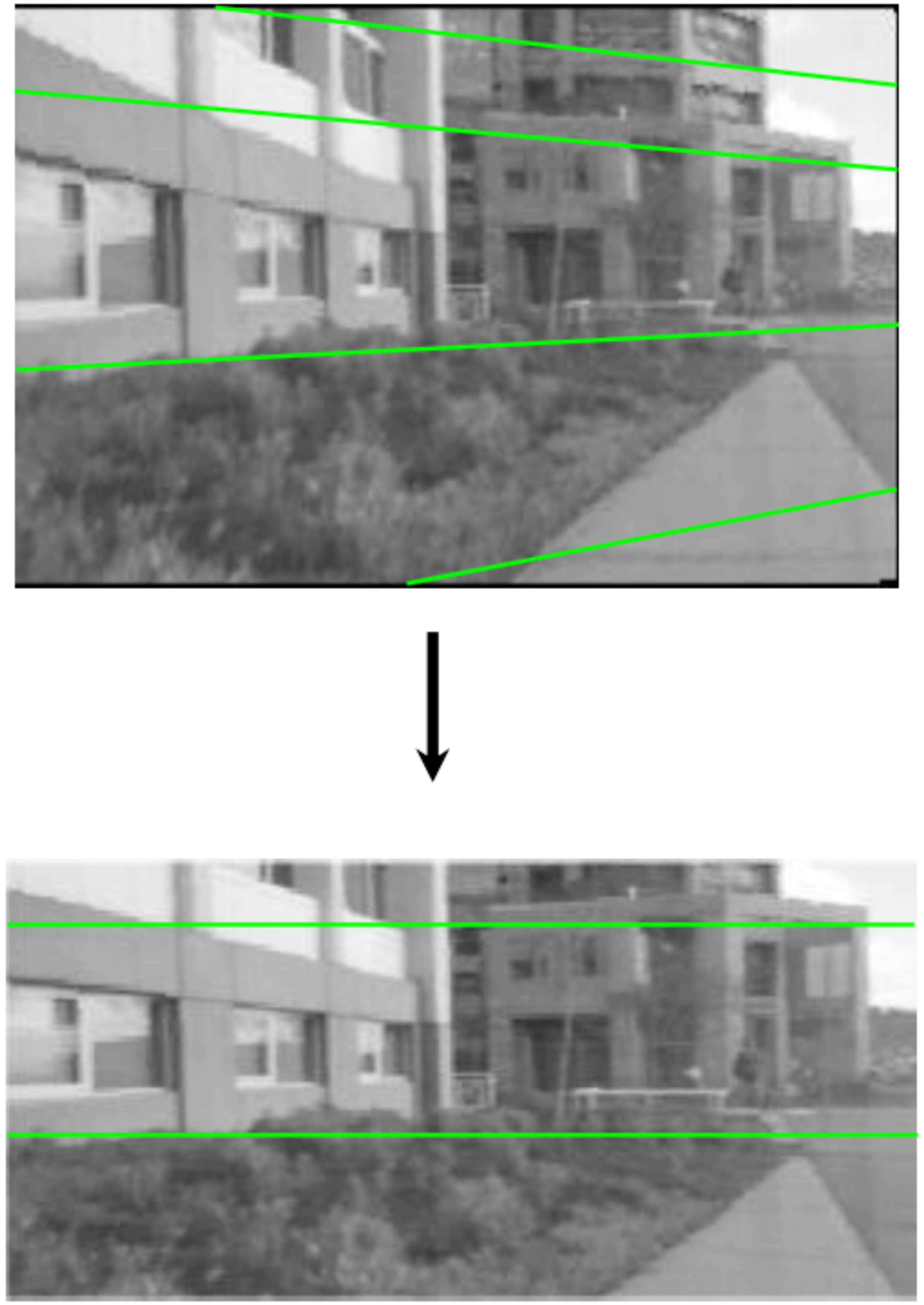
Finding Correspondences



$W(\mathbf{p}_l) \approx W(\mathbf{p}_r)$

Rectification

- Searching along epipolar lines at arbitrary orientation is intuitively expensive.
- Given the epipolar geometry of the stereo pair, there exists in general a transformation that maps the images into a pair of images with the epipolar lines parallel to the rows of the image. This transformation is called rectification.
- Images are almost always rectified before searching for correspondences in order to simplify the search.



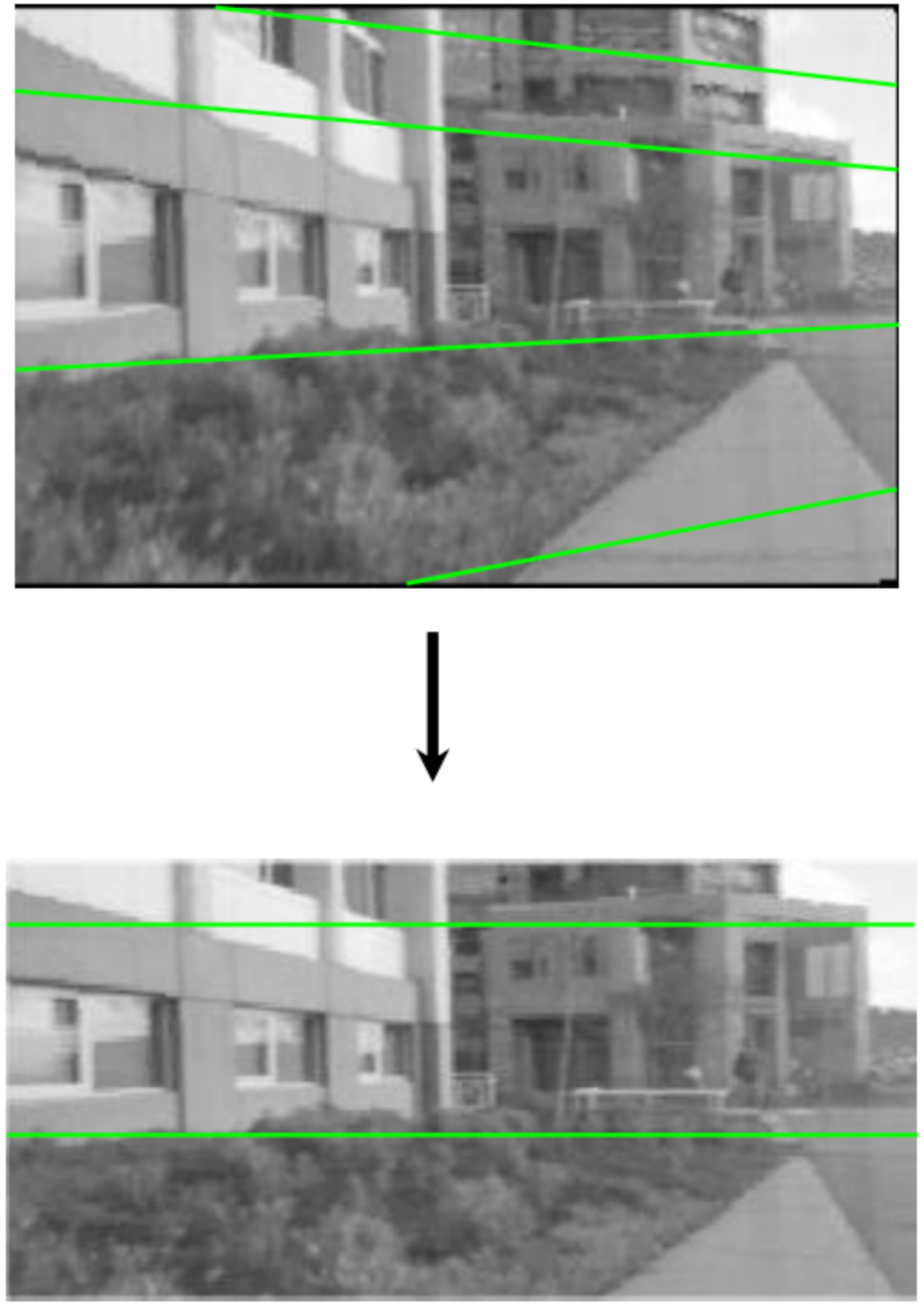
Rectification

- Searching along epipolar lines at arbitrary orientation is intuitively expensive.

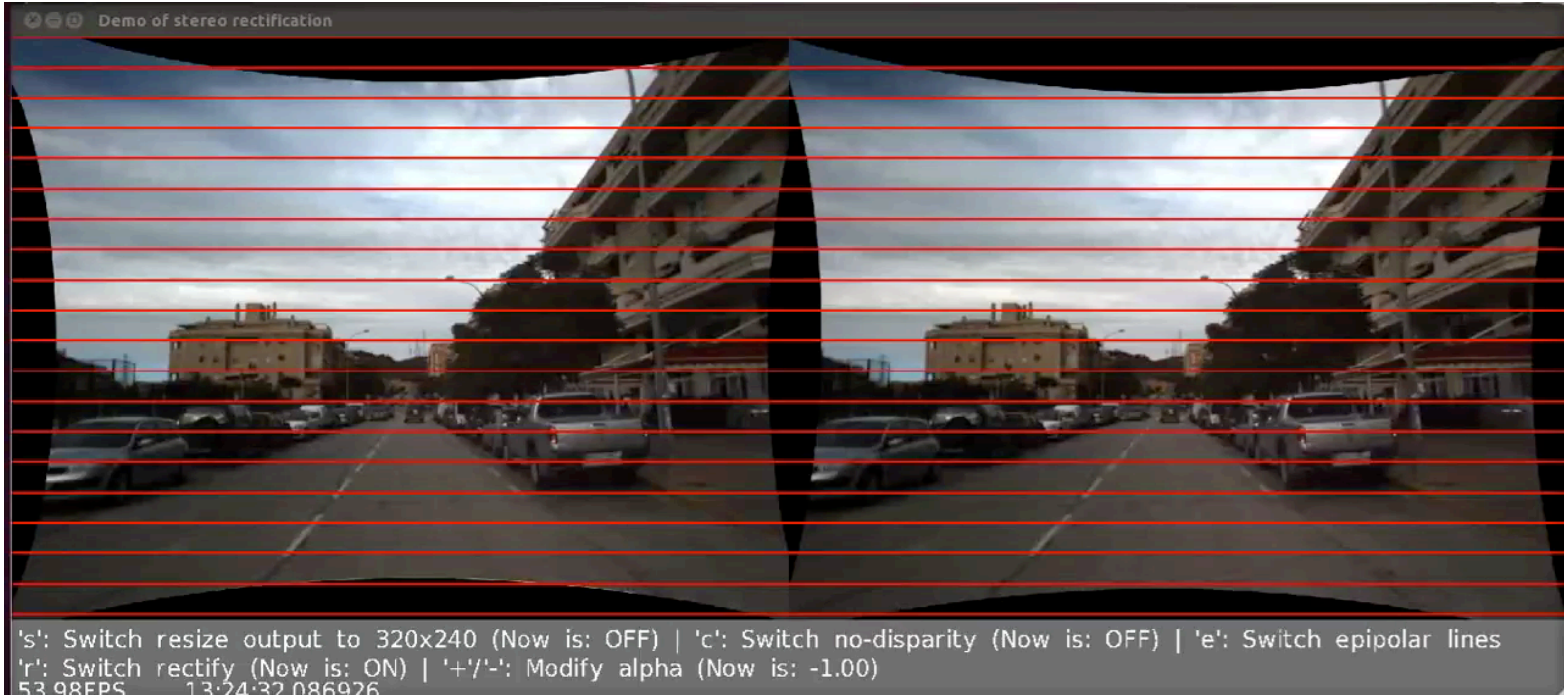
- Given the epipolar geometry of the stereo pair, there exists in general a transformation that maps the images into a pair of images with the epipolar lines parallel to the rows of the image. This transformation is called rectification.

- Images are almost always rectified before searching for correspondences in order to simplify the search.

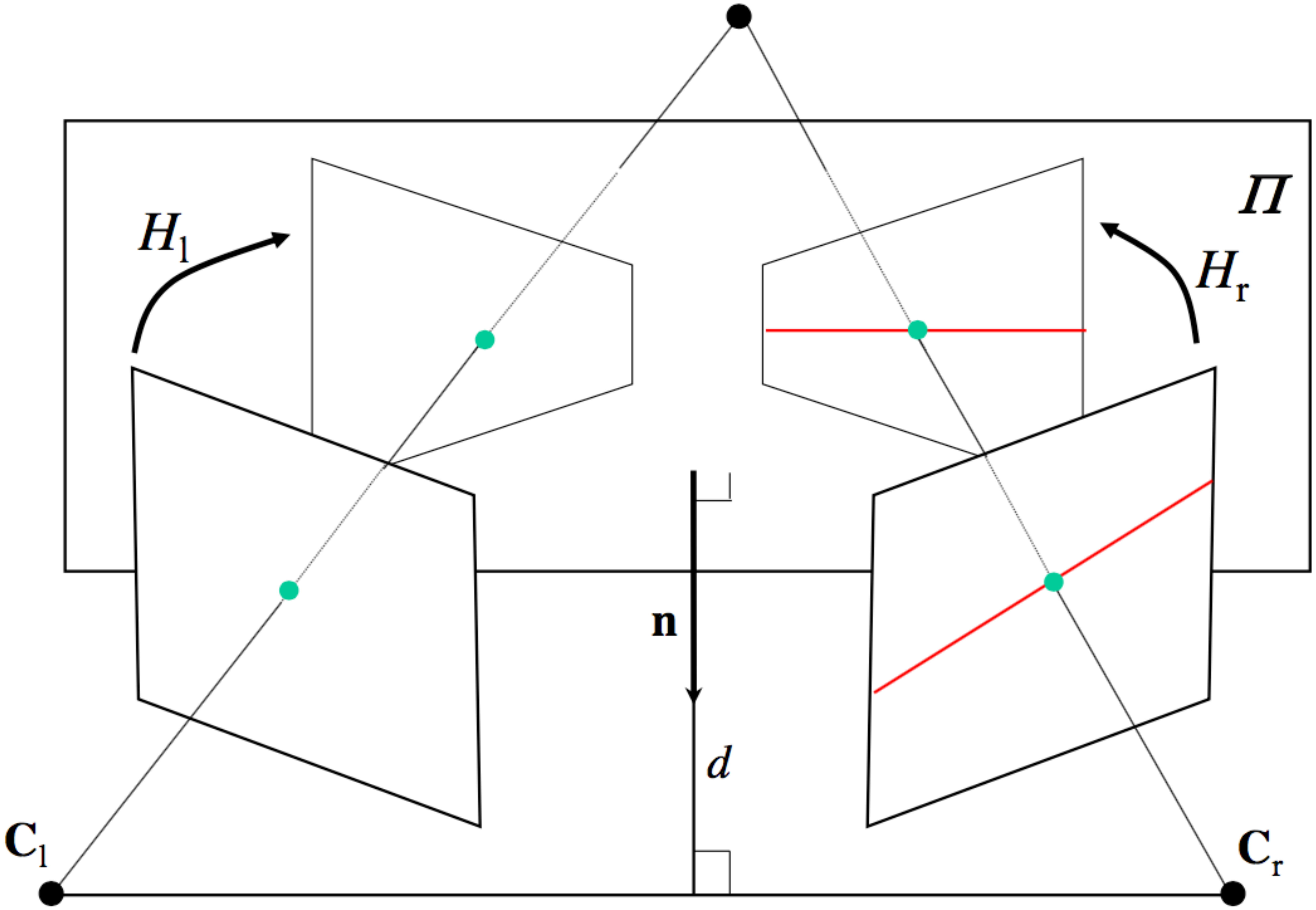
- When is this transformation not possible? (Hint: epipoles)

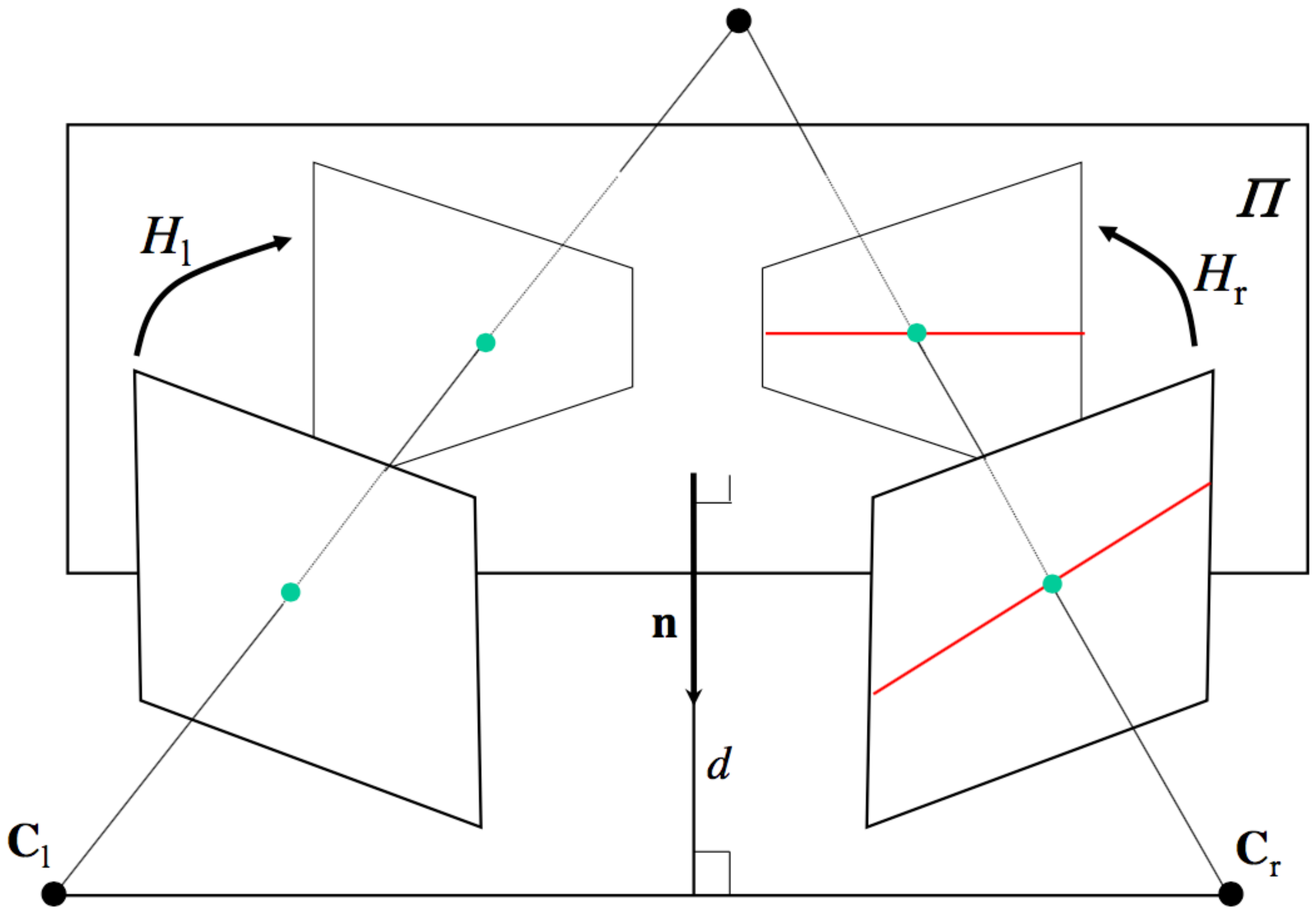


Rectification



Rectification Plane



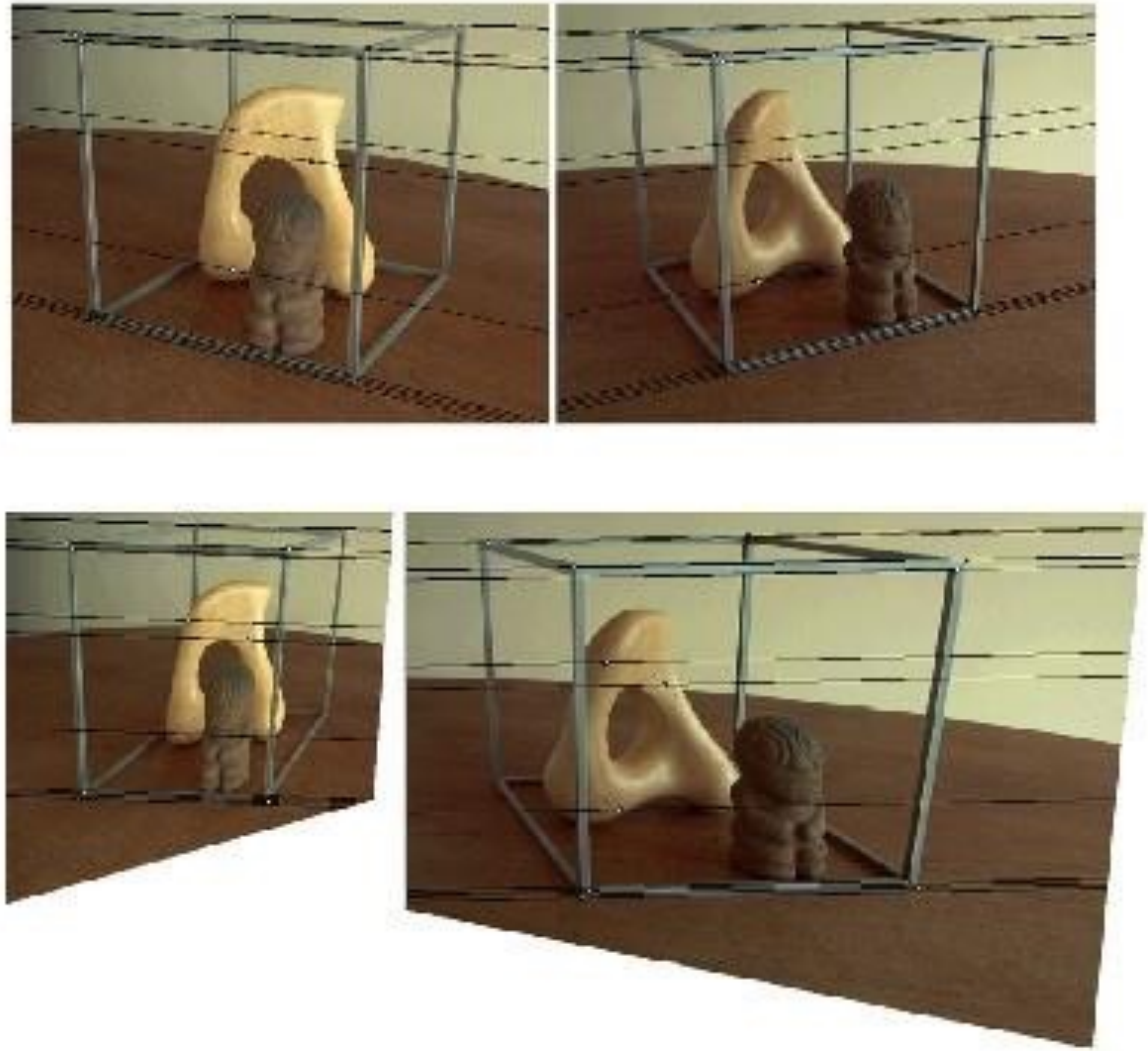


We know that, given a plane **P** in space, there exists two homographies $\mathbf{H}_l$ and $\mathbf{H}_r$ that map each image plane onto **P**. That is, if $\mathbf{p}_l$ is a point in the left image, then the corresponding point in **P** is **Hp** (in homogeneous coordinates). If we map both images to a common plane **P** such that P is parallel to the line $\mathbf{C}_l\mathbf{C}_r$, then the pair of virtual (rectified) images is such that the epipolar lines are parallel. With proper choice of the coordinate system, the epipolar lines are parallel to the rows of the image.

The algorithm for rectification is then:

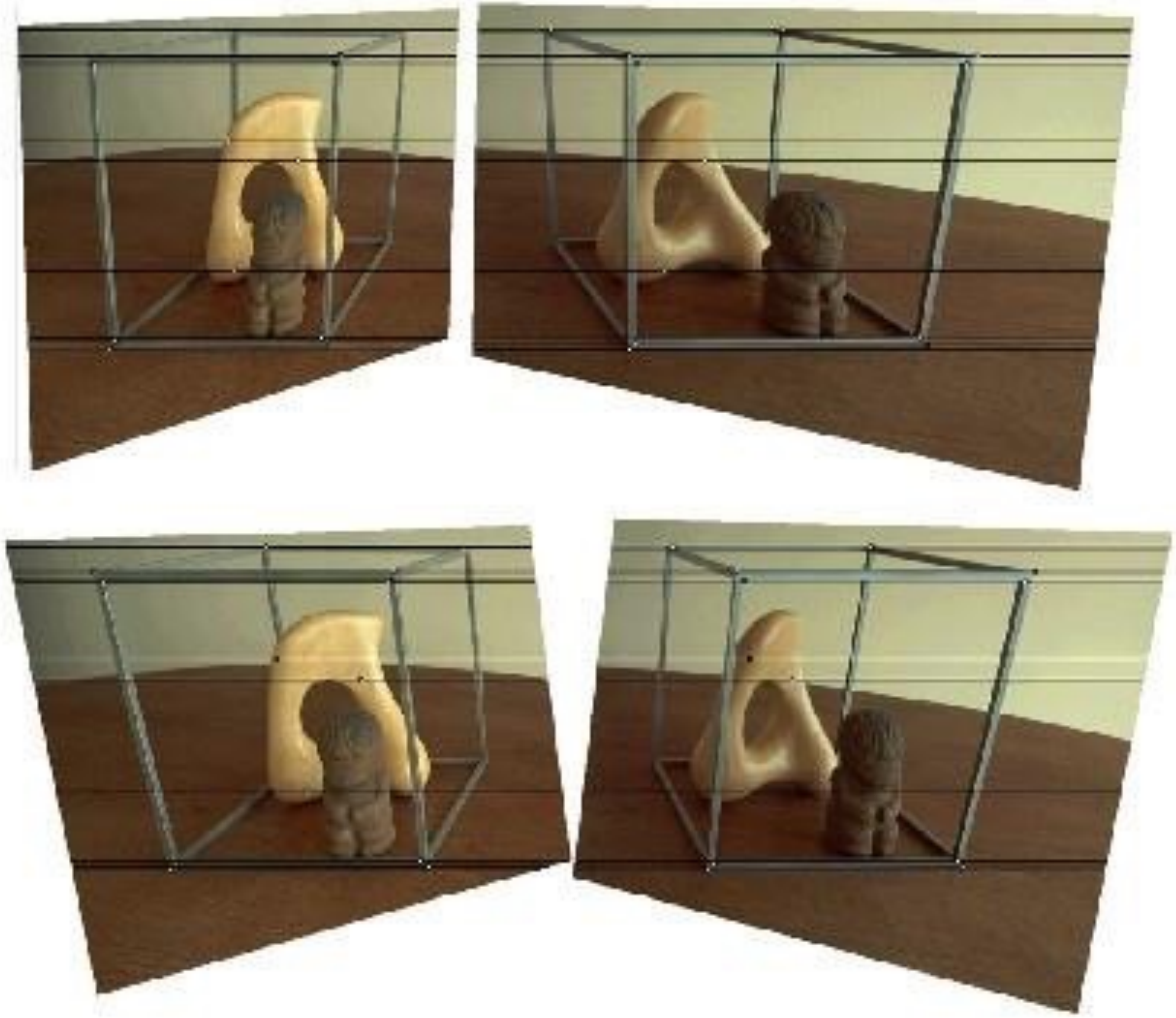
- Select a plane **P** parallel to $\mathbf{C}_r\mathbf{C}_l$
- Define the left and right image coordinate systems on **P**
- Construct the rectification matrices $\mathbf{H}_l$ and $\mathbf{H}_r$ from **P** and the virtual image's coordinate systems.

Example (Seitz/Szeliski)



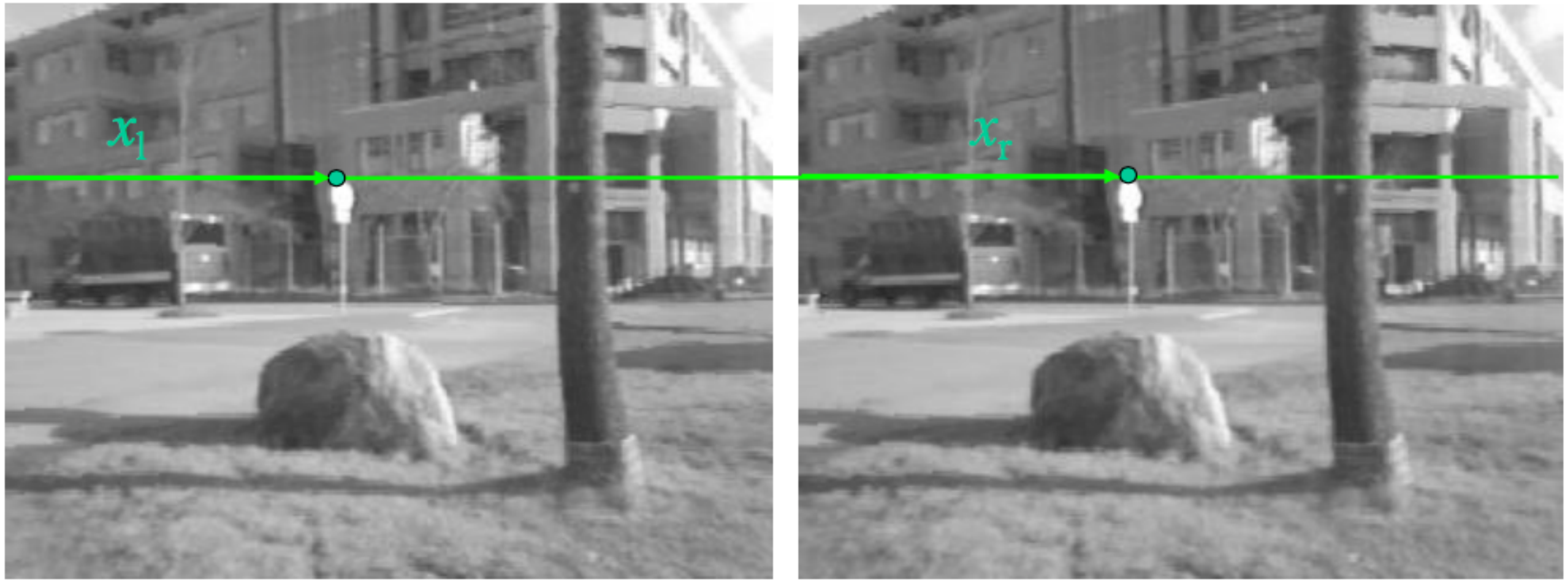
Destination images have significantly different sizes

Example (Seitz/Szeliski)



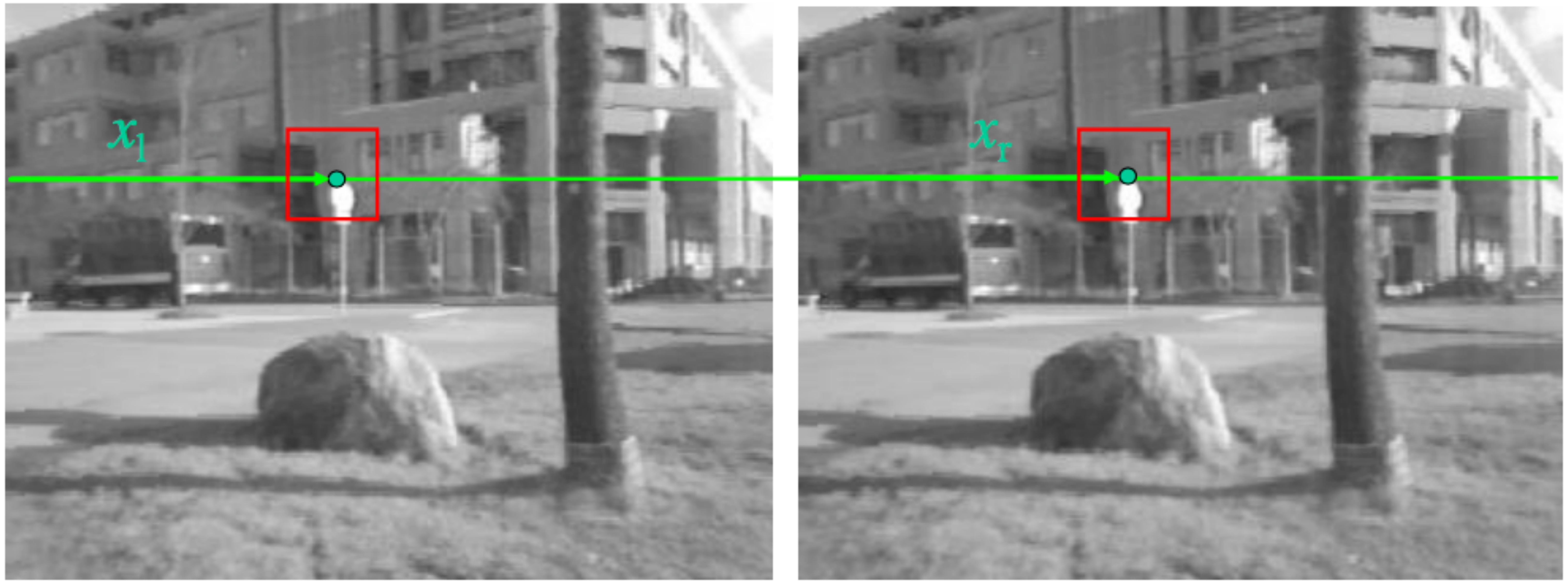
Better: Destination images have similar sizes

Now that images are lined up
Disparity!



$$d = x_l - x_r$$

Disparity



$$d = x_l - x_r$$

$$\underset{d}{Min} \sum_{x,y \in W} \psi(I_l(x,y), I_r(x-d,y))$$

Disparity Image

Fun: Cross your eyes and look at the image to get a 3D picture

Matching Functions

SSD:

$$\psi(I_l(x,y), I_r(x+d,y)) = (I_l(x,y) - I_r(x-d,y))^2$$

SAD:

$$\psi(I_l(x,y), I_r(x+d,y)) = \left|I_l(x,y) - I_r(x-d,y)\right|$$

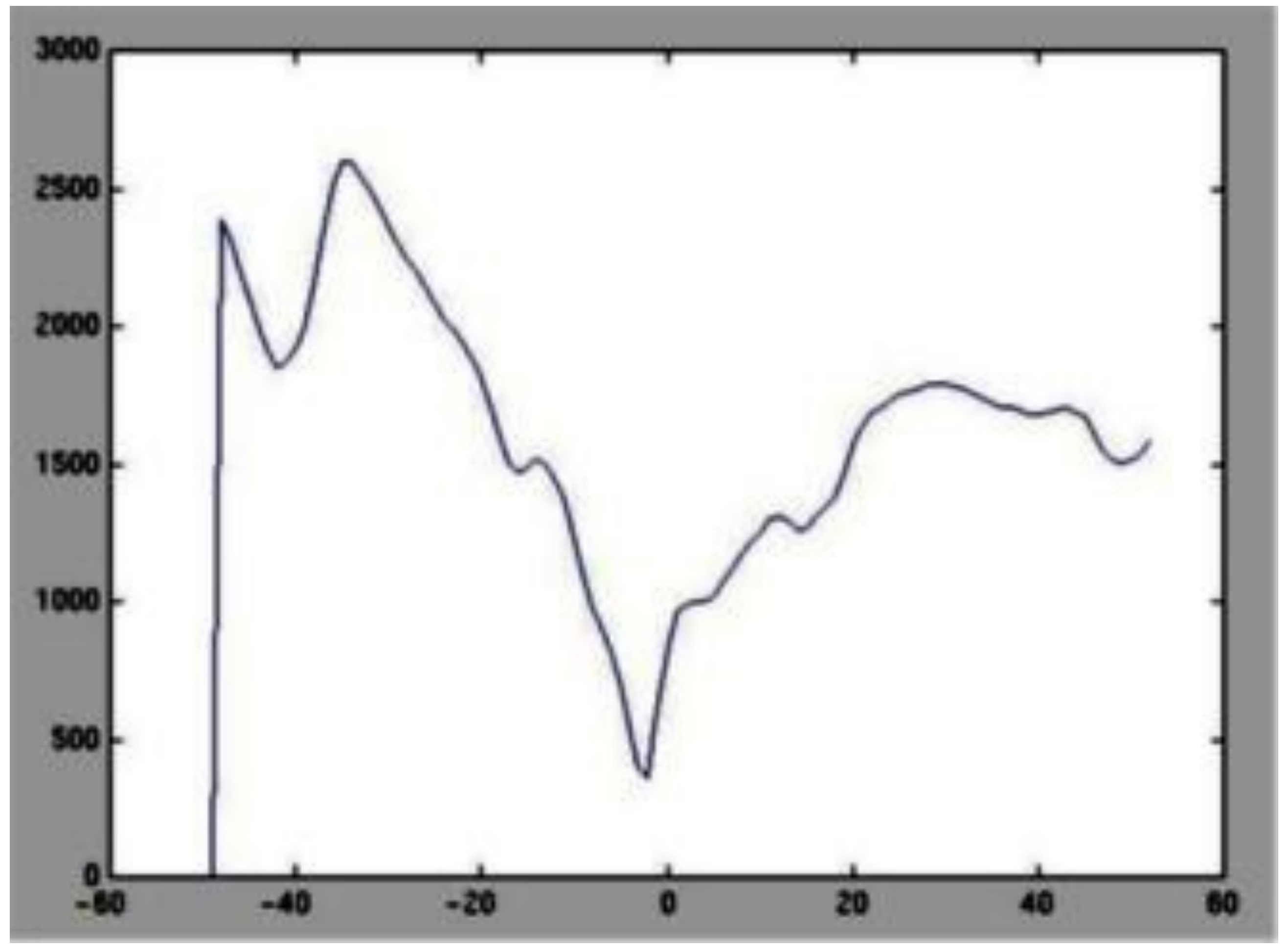
Correlation:

$$\psi(I_l(x,y), I_r(x+d,y)) = I_l(x,y).I_r(x-d,y)$$

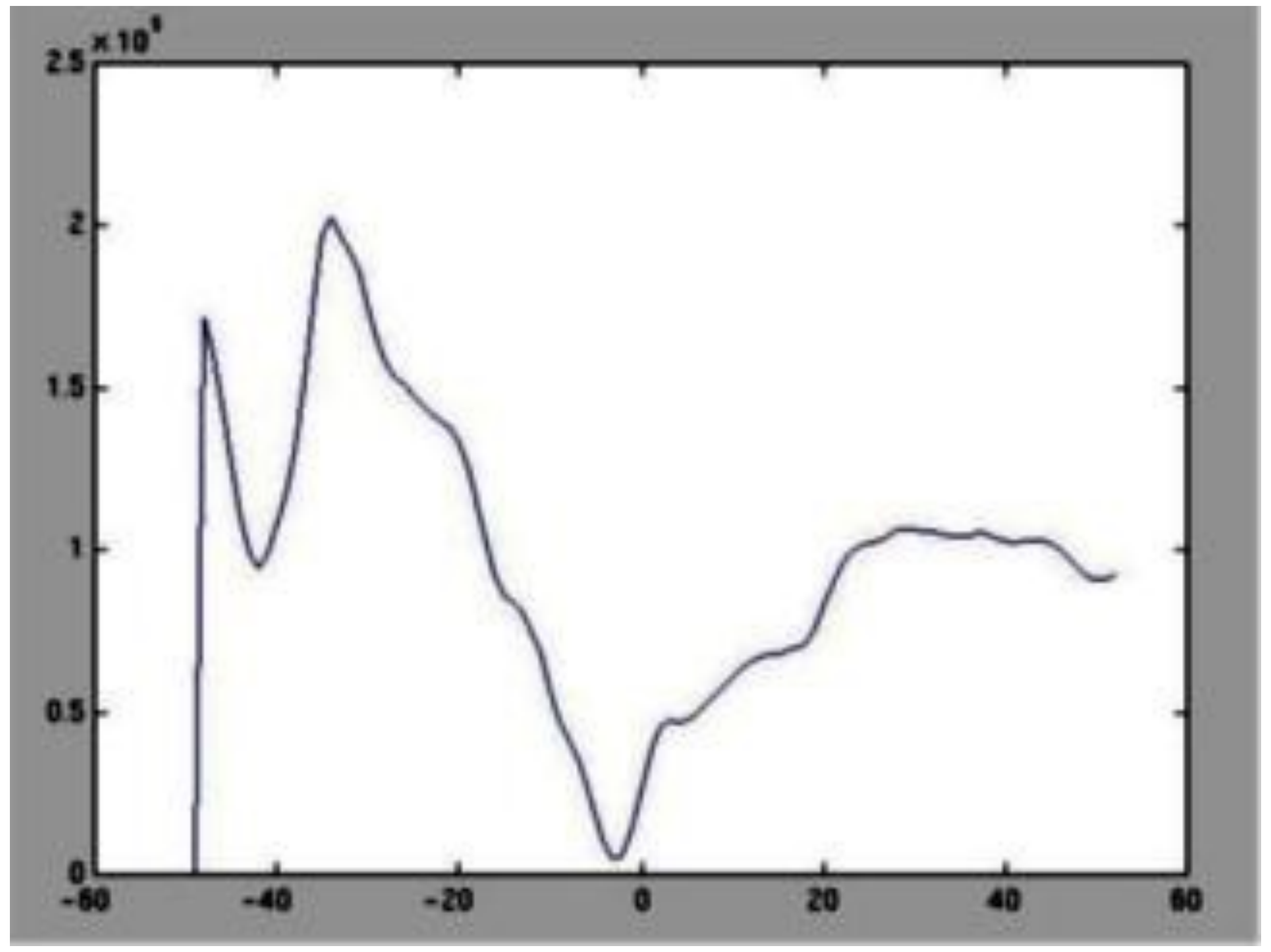
Normalized Correlation:

$$\psi(I_l(x,y), I_r(x+d,y)) = \frac{I_l(x,y).I_r(x-d,y) - \bar{I}_l \bar{I}_r}{\sigma_l \sigma_r(d)}$$

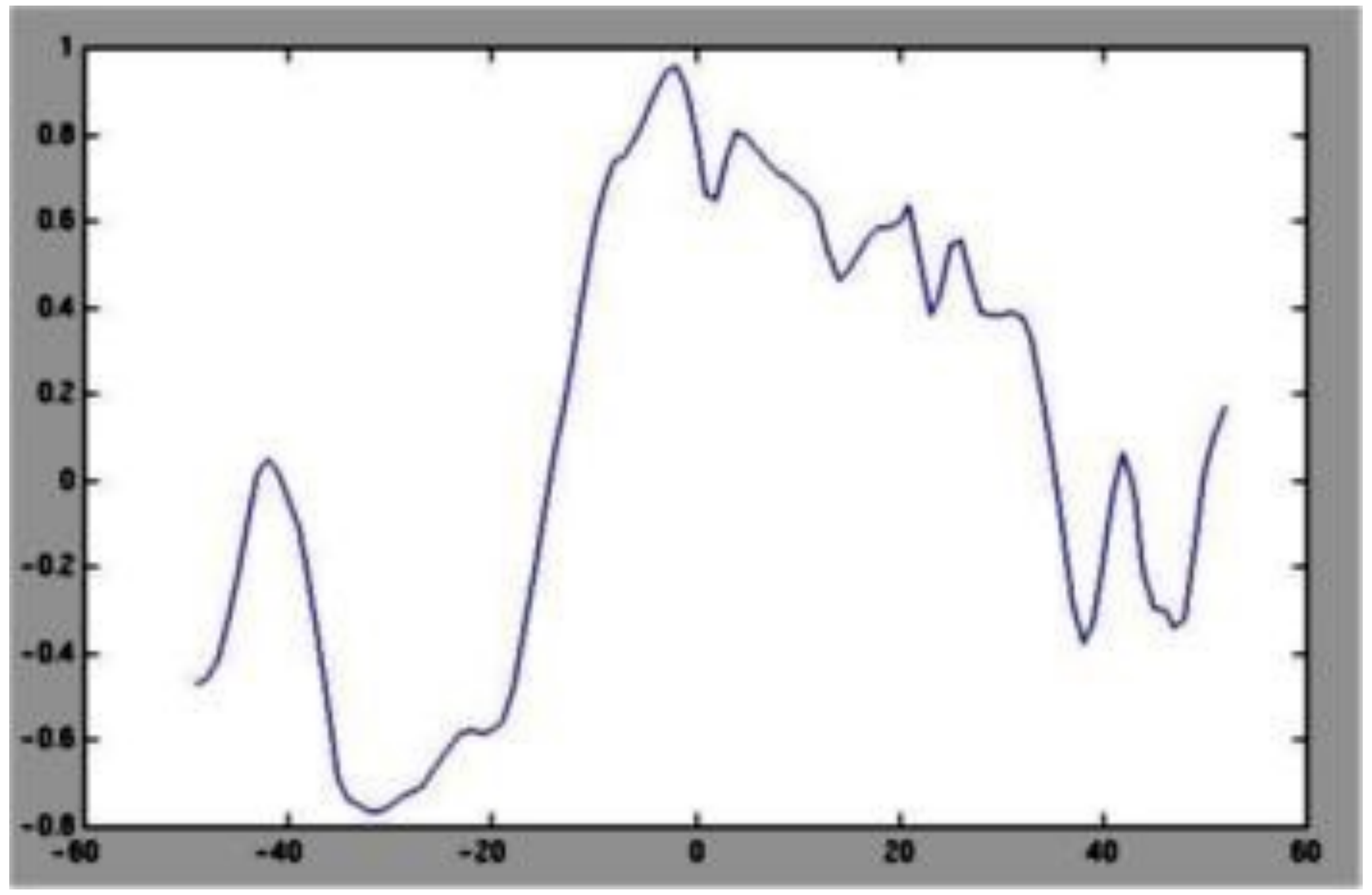
Comparison of Matching Functions



SAD

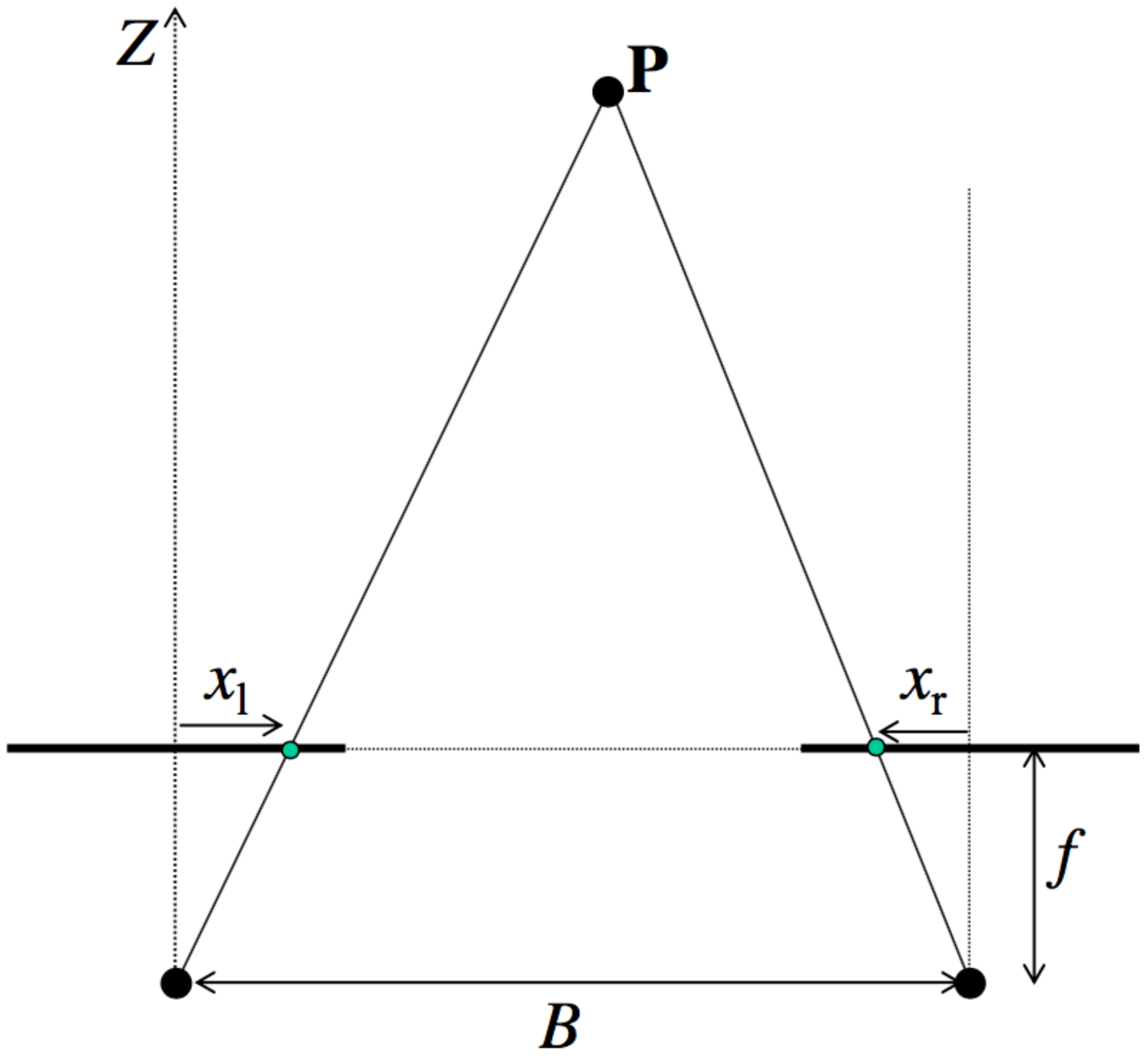


SSD



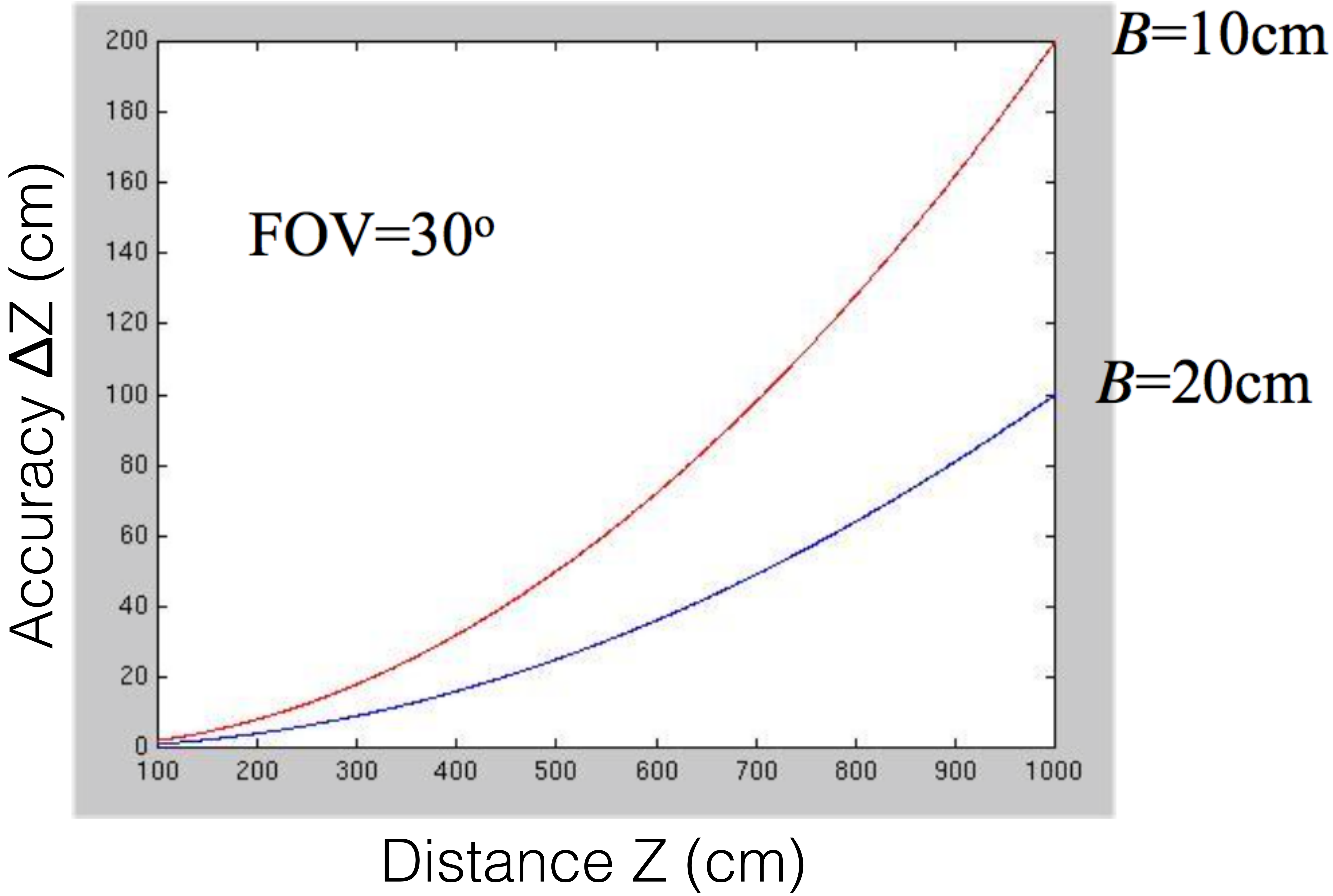
NCC

Geometry of Disparity



$$Z = \frac{Bf}{d}$$

Accuracy as Function of Depth

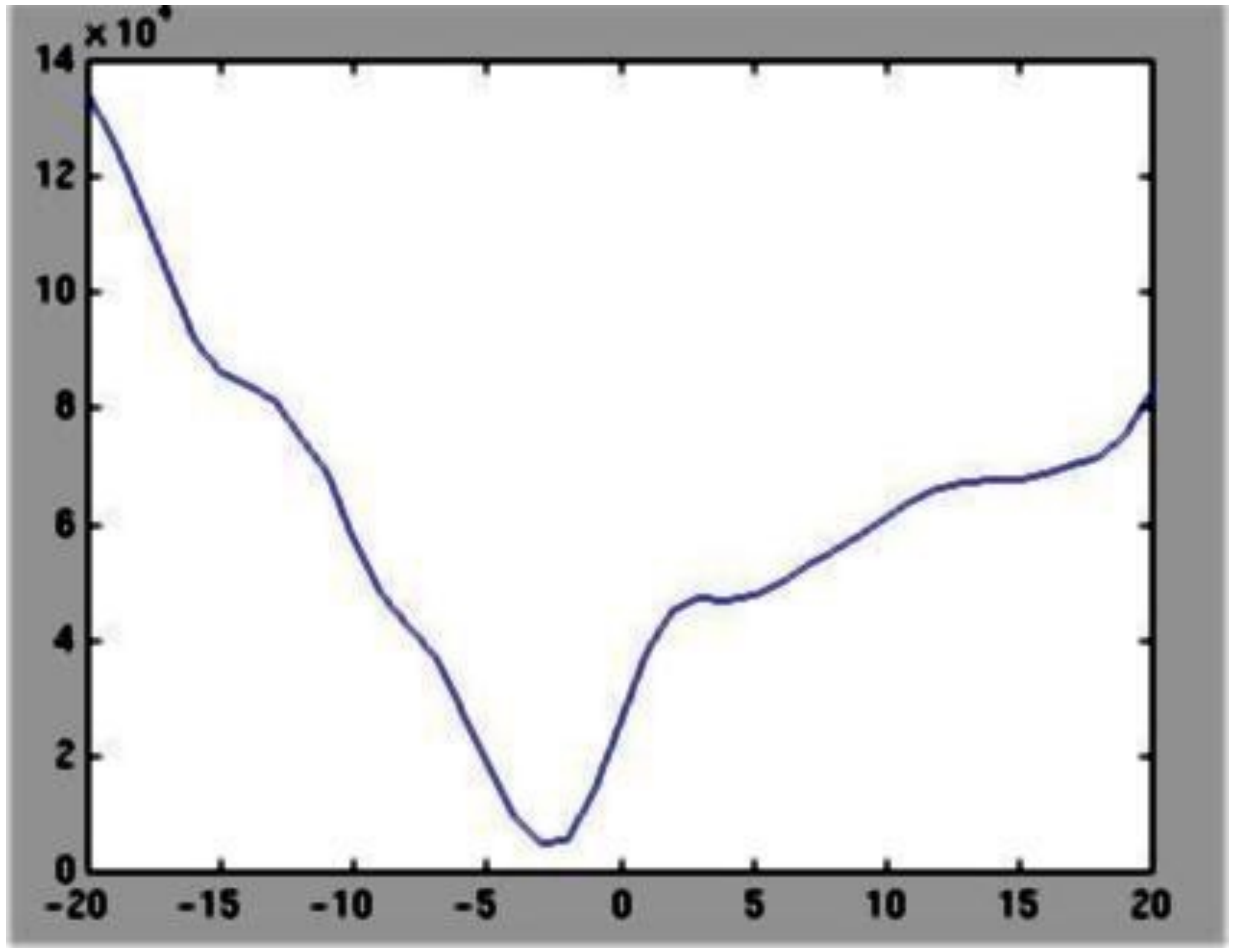


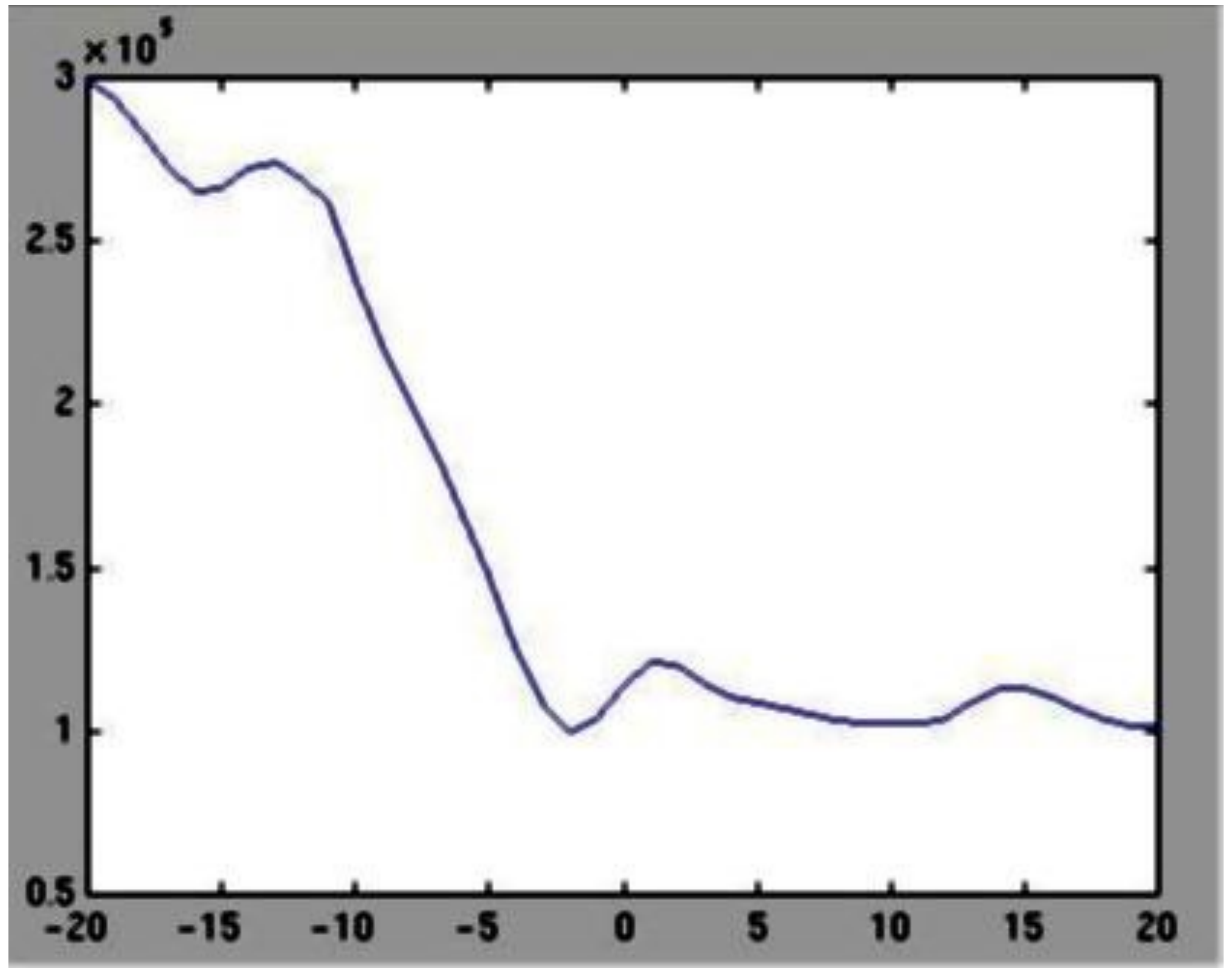
$$\Delta Z \approx \frac{Z^2}{Bf}$$

Think about it…

- Why don't we always just go for the largest possible baseline in computer vision?

Lighting/reflectance issues





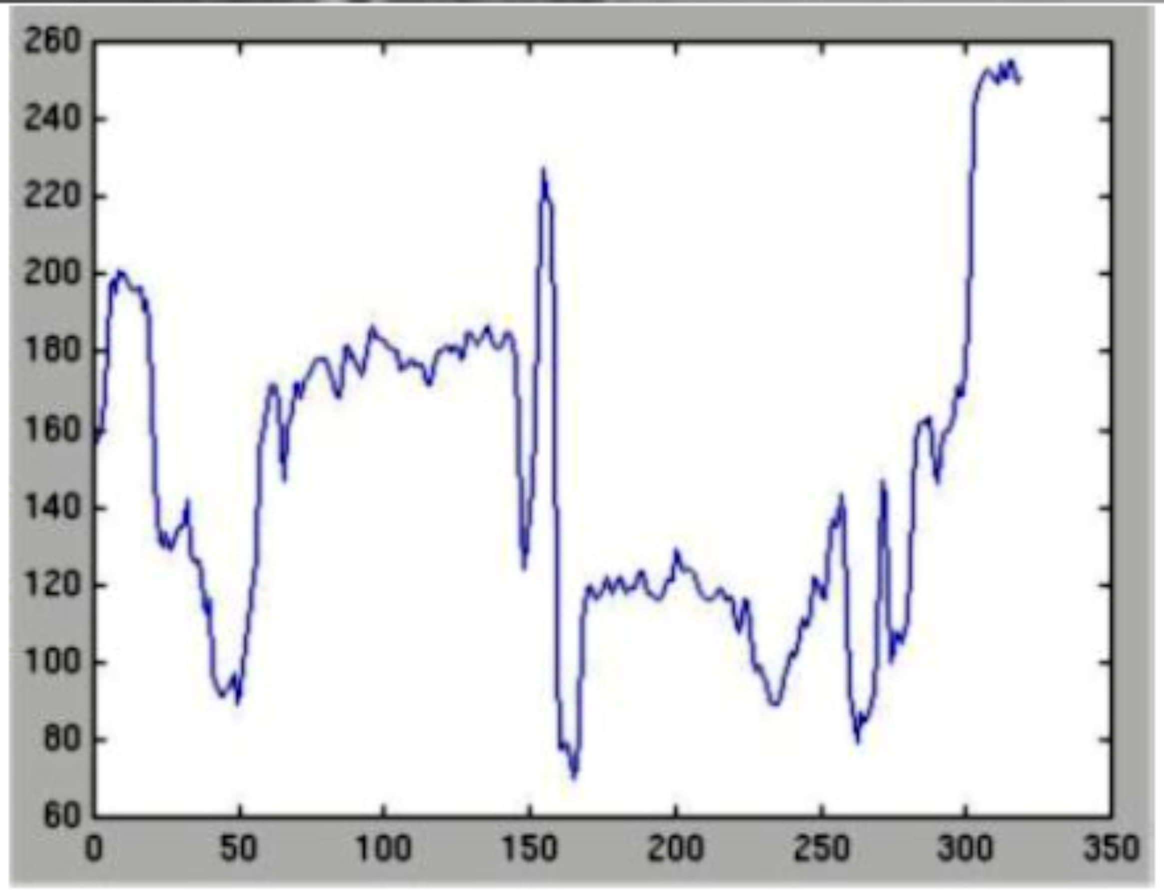
Lighting conditions/Photometric Variation

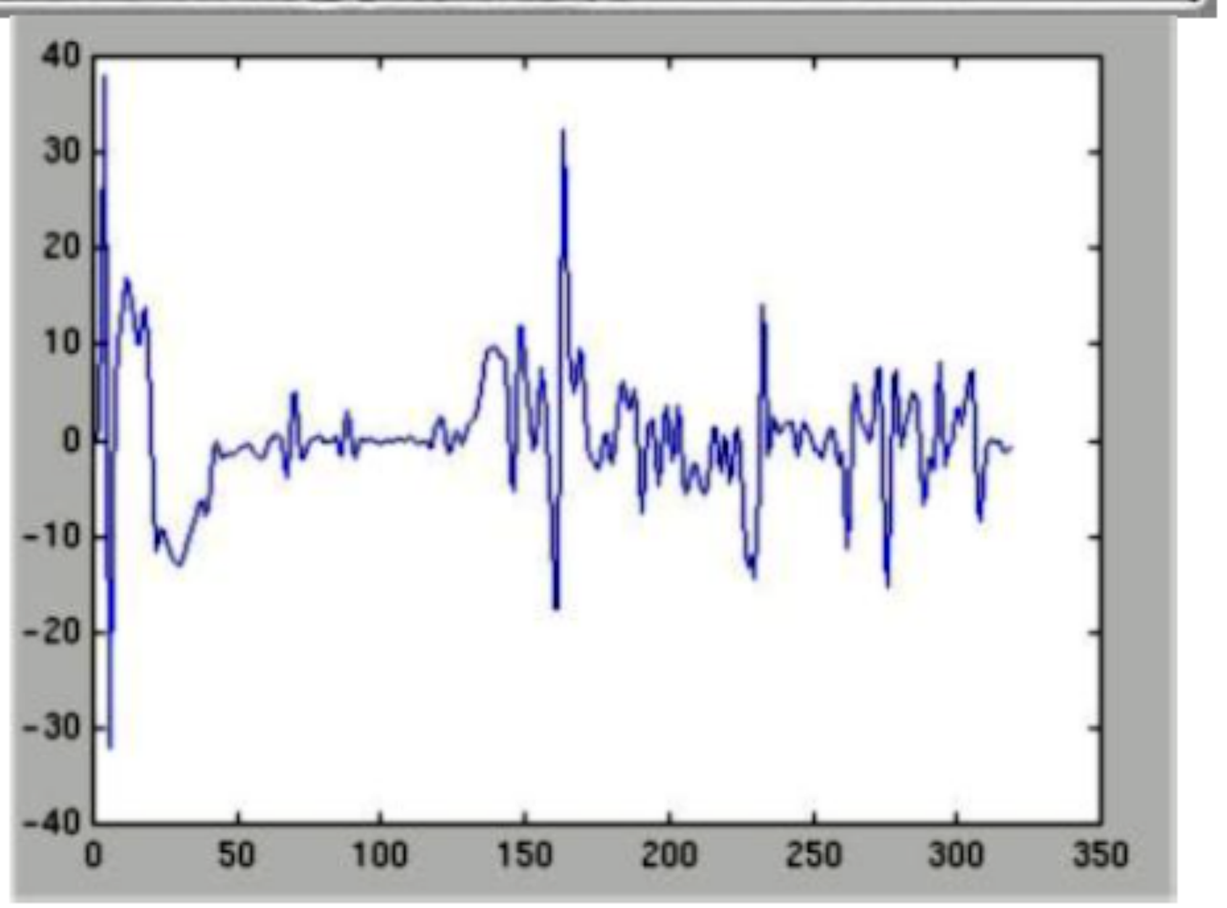
$W(\mathbf{P}_l)$

$W(\mathbf{P}_r)$

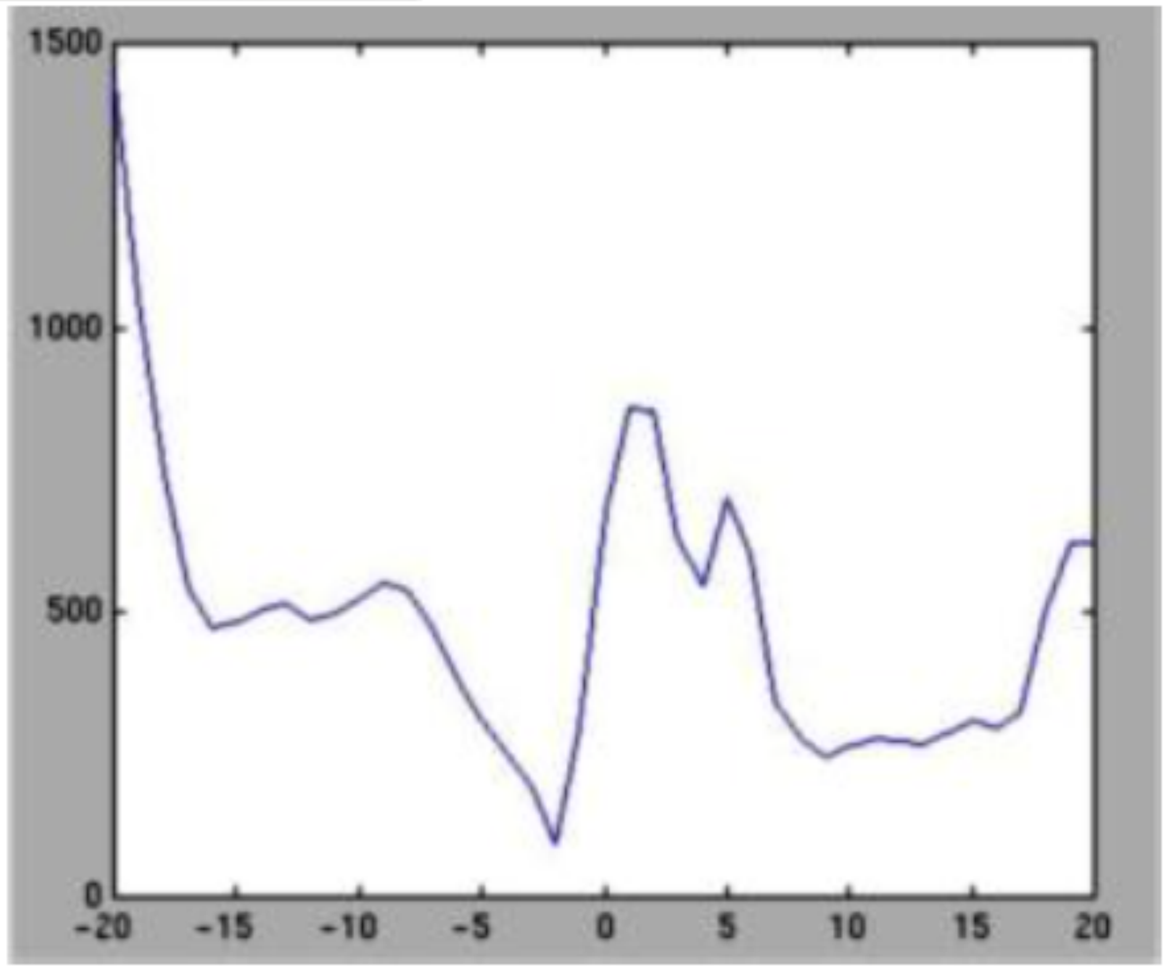
Original

$$LoG = \nabla^2 G_\sigma * I$$





Laplacian of Gaussian reduces the effect of illumination/photometric variations



LoG SSD

Multi-scale Matching