COMPUTER VISION AND PHOTOGRAMMETRY

REVIEW AND CLOSURE

REVIEW

- Keypoint detection
- Feature Descriptor & Matching
- 2D-3D Geometry
- Homographies
- Epipolar Geometry
- DLS Linear Methods
- RANSAC
- Bundle Adjustment
- Auto-Calibration
- Stereo *
- Meshing
- Texturing
- SLAM

LIST OF QUESTIONS TO CLARIFY TODAY

- How to know what are inliers for P?
- Auto-calibration?
- Automatic Correct E?
- Dense?
- Align P2,P3… from pairs?

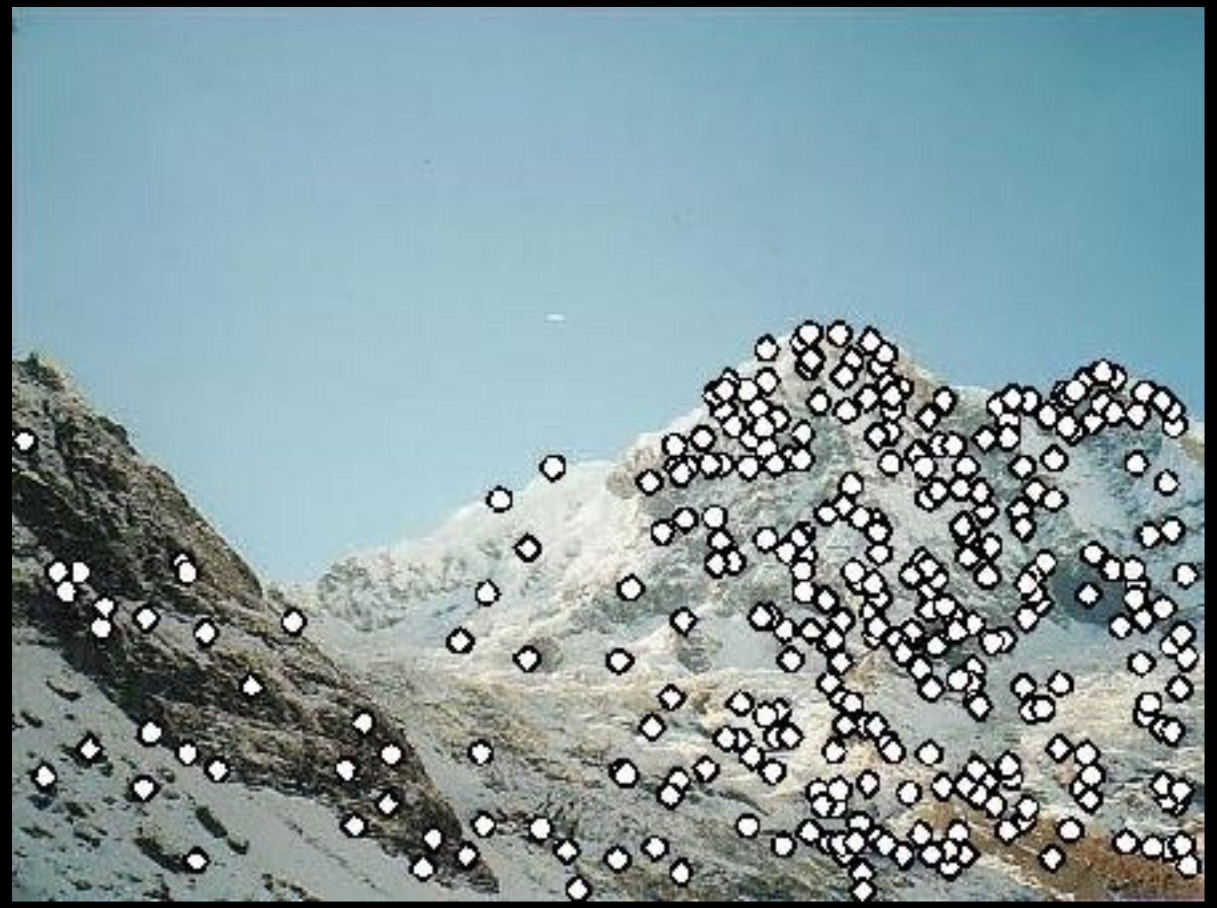
APPLICATIONS OF PHOTOGRAMMETRY

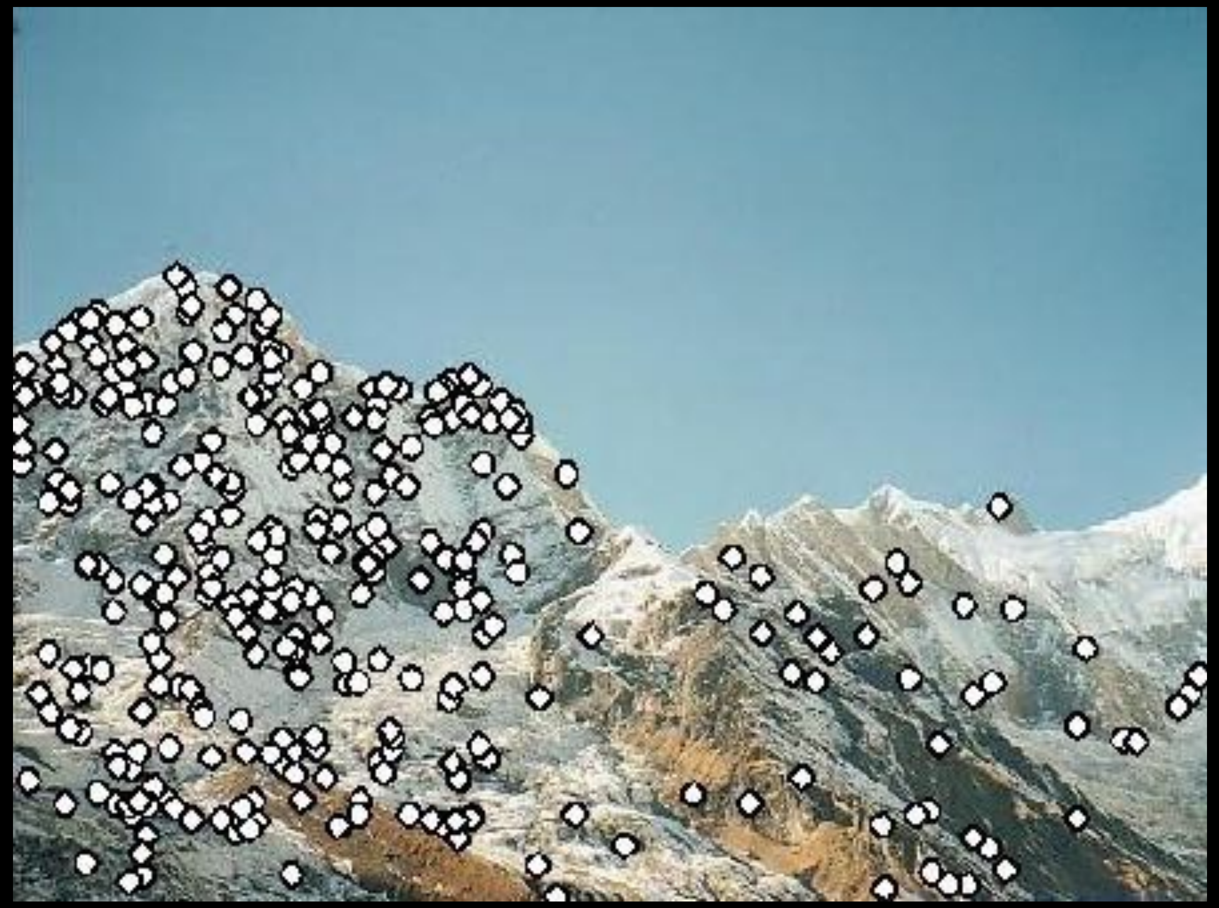
- Find cool examples and commercial applications
- 3D scanning
- Site Mapping using Drones
- Compositing in VFX

FEATURE DETECTION AND MATCHING

CHARACTERISTICS OF GOOD FEATURES

- Repeatability
  - The same feature can be found in several images despite geometric and photometric transformations
- Saliency
  - Each feature is distinctive
- Compactness and efficiency
  - Many fewer features than image pixels
- Locality
  - A feature occupies a relatively small area of the image; robust to clutter and occlusion





KEY POINTS

MANY EXISTING DETECTORS AVAILABLE

- Hessian & Harris [Beaudet '78], [Harris '88]
- Laplacian, DoG [Lindeberg '98], [Lowe 1999]
- Harris-/Hessian-Laplace [Mikolajczyk & Schmid '01]
- Harris-/Hessian-Affine [Mikolajczyk & Schmid '04]
- EBR and IBR [Tuytelaars & Van Gool '04]
- MSER [Matas '02]
- Salient Regions [Kadir & Brady '01]
- Others…

K. Grauman, B. Leibe

INVARIANCE AND COVARIANCE

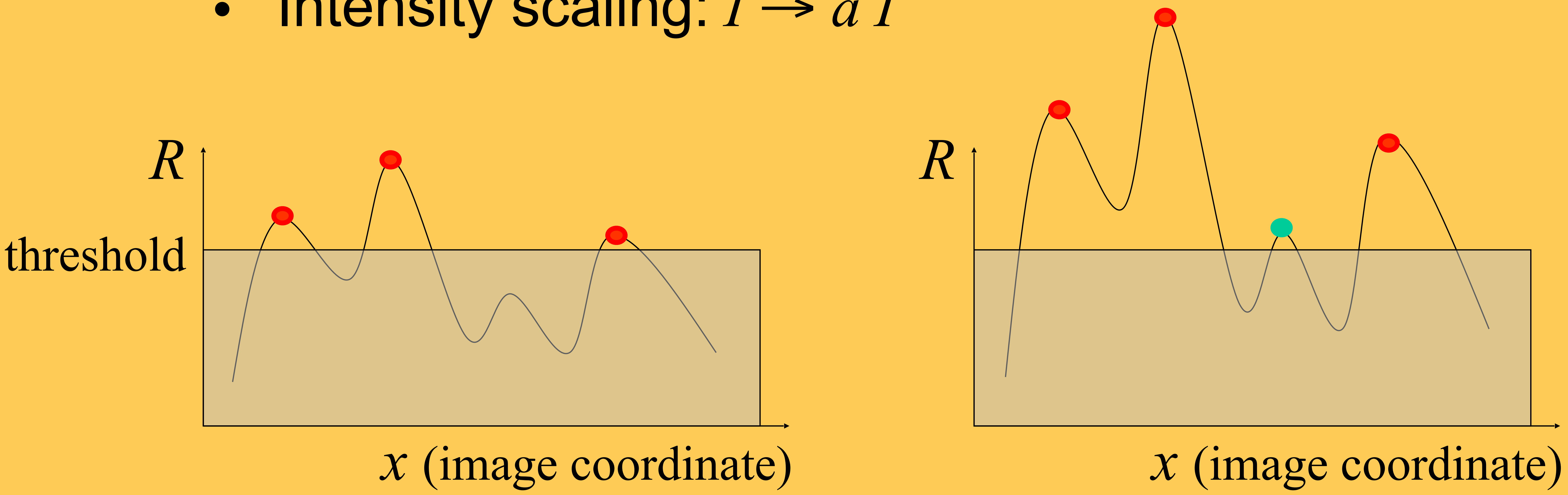
- We want corner locations to be *invariant* to photometric transformations and *covariant* to geometric transformations
  - **Invariance:** image is transformed and corner locations do not change
  - **Covariance:** if we have two transformed versions of the same image, features should be detected in corresponding locations

AFFINE INTENSITY CHANGE

$$I \rightarrow a\, I + b$$

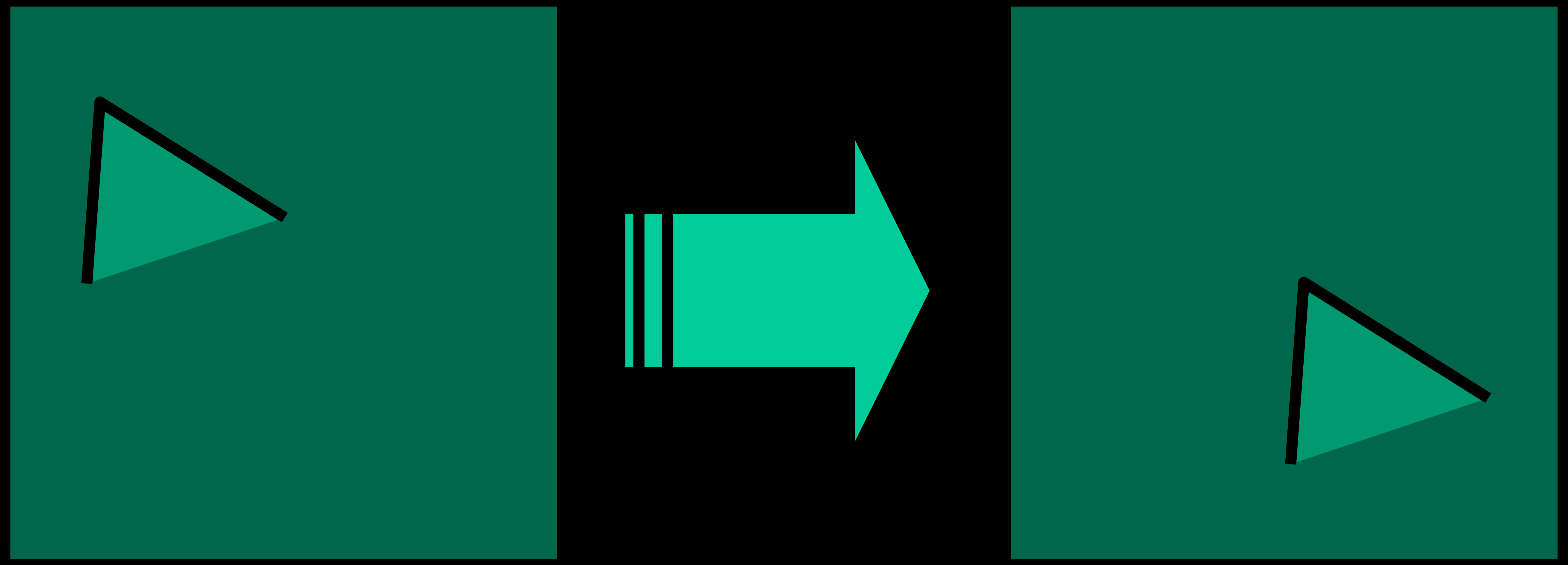
- Only derivatives are used => invariance to intensity shift $I \rightarrow I + b$

- Intensity scaling: $I \rightarrow a\, I$



Partially invariant to affine intensity change

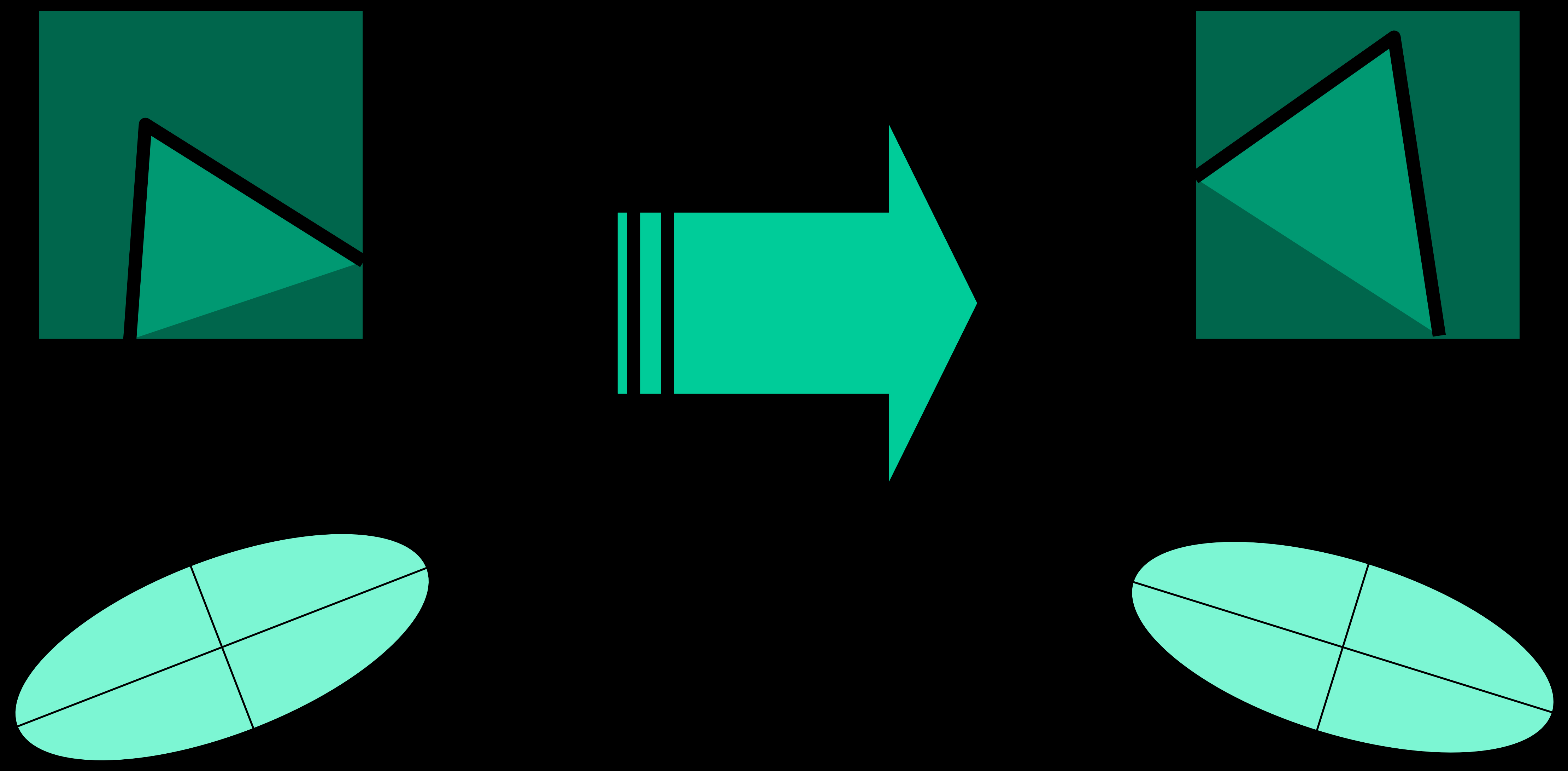
IMAGE TRANSLATION



- Derivatives and window function are shift-invariant

Corner location is covariant w.r.t. translation

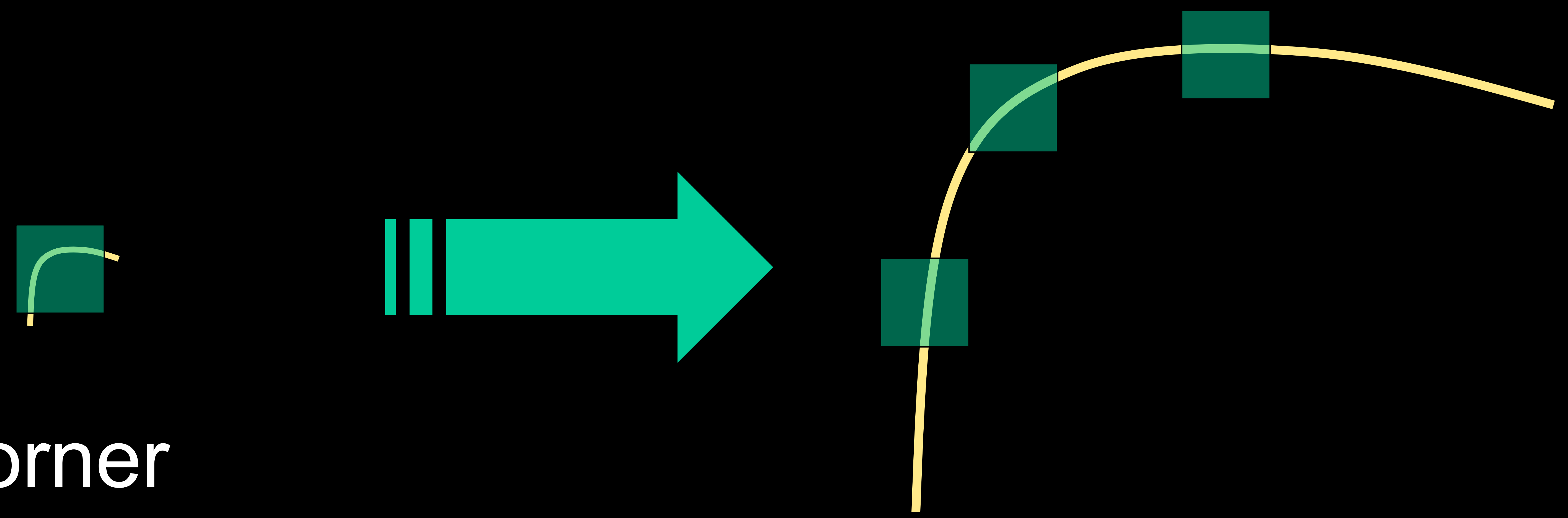
IMAGE ROTATION



Second moment ellipse rotates but its shape (i.e. eigenvalues) remains the same

Corner location is covariant w.r.t. rotation

SCALING

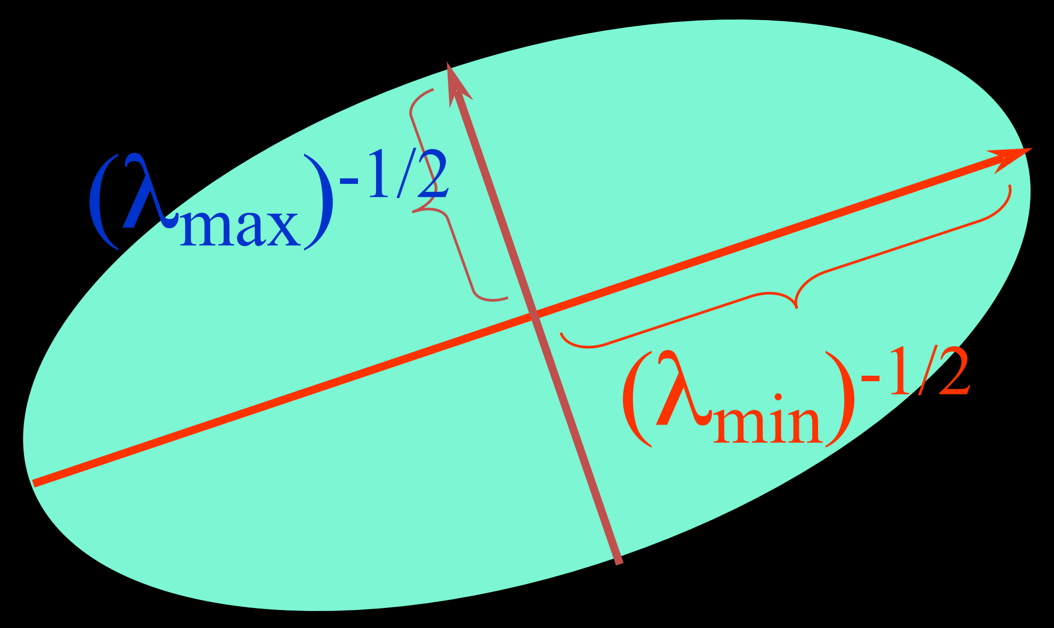


All points will be classified as edges

Corner location is not covariant to scaling!

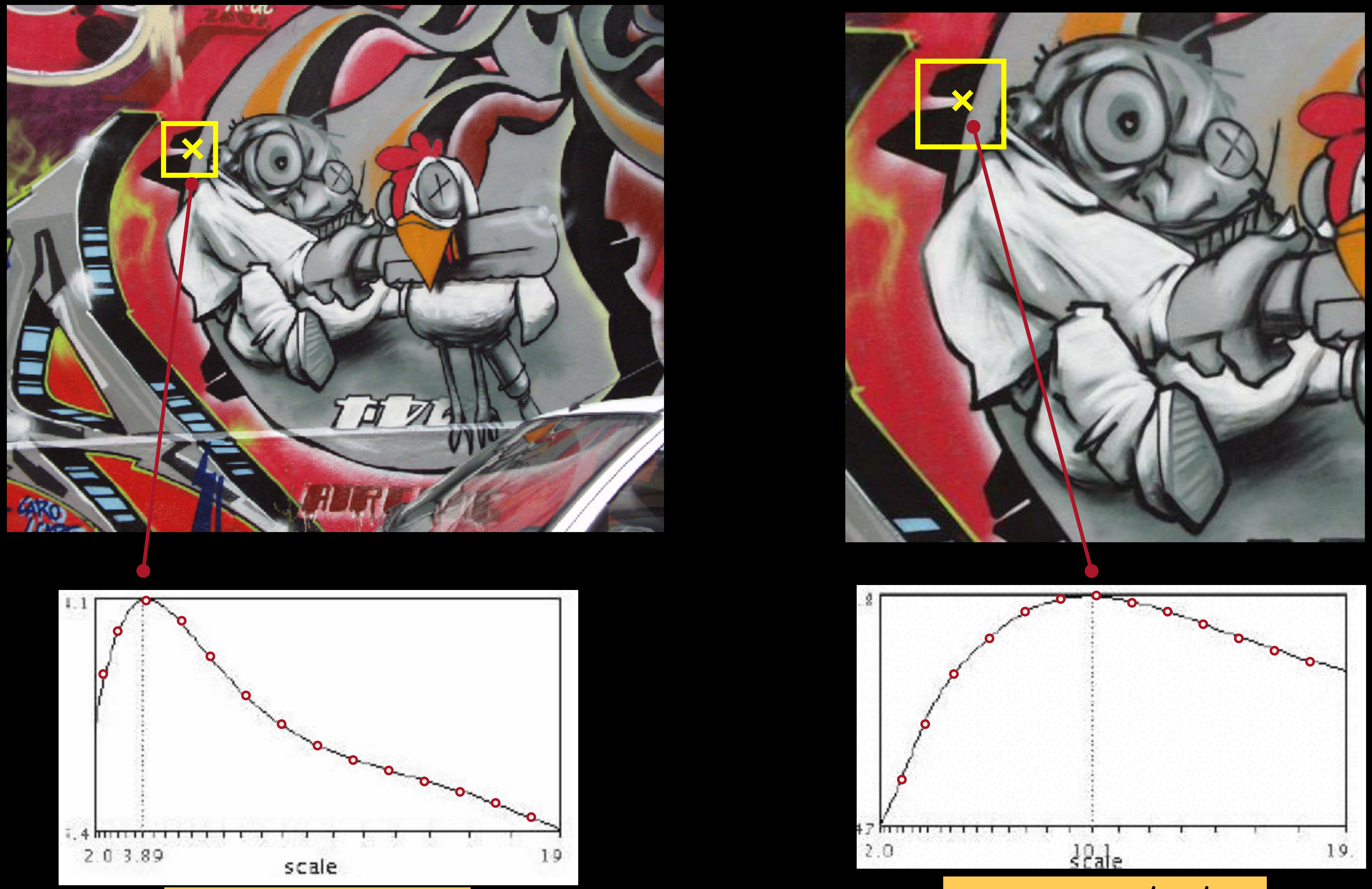
REVIEW: HARRIS CORNER DETECTOR

- Approximate distinctiveness by local auto-correlation.
- Approximate local auto-correlation by second moment matrix
- Quantify distinctiveness (or cornerness) as function of the eigenvalues of the second moment matrix.
- But we don’t actually need to compute the eigenvalues by using the determinant and trace of the second moment matrix.



AUTOMATIC SCALE SELECTION

- Function responses for increasing scale (scale signature)

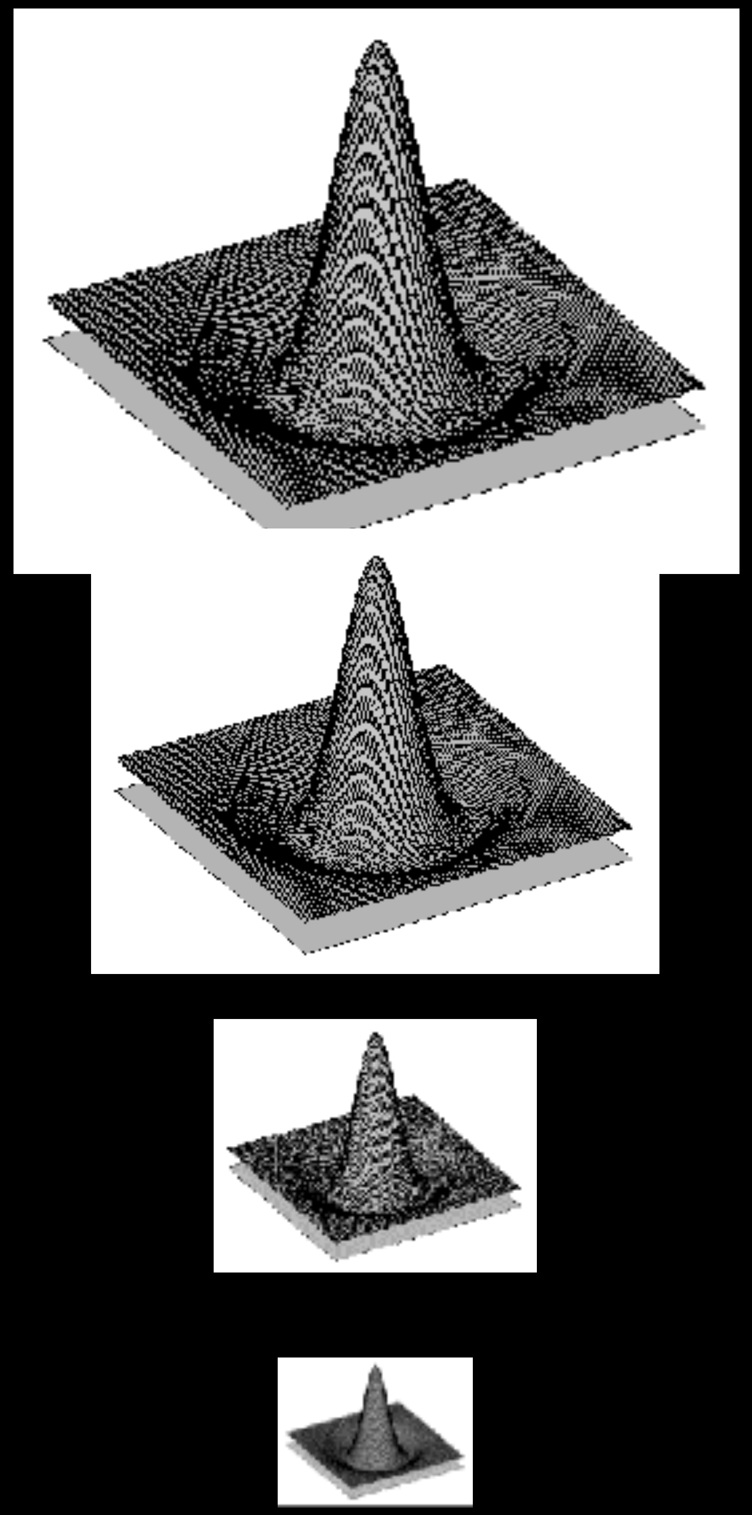


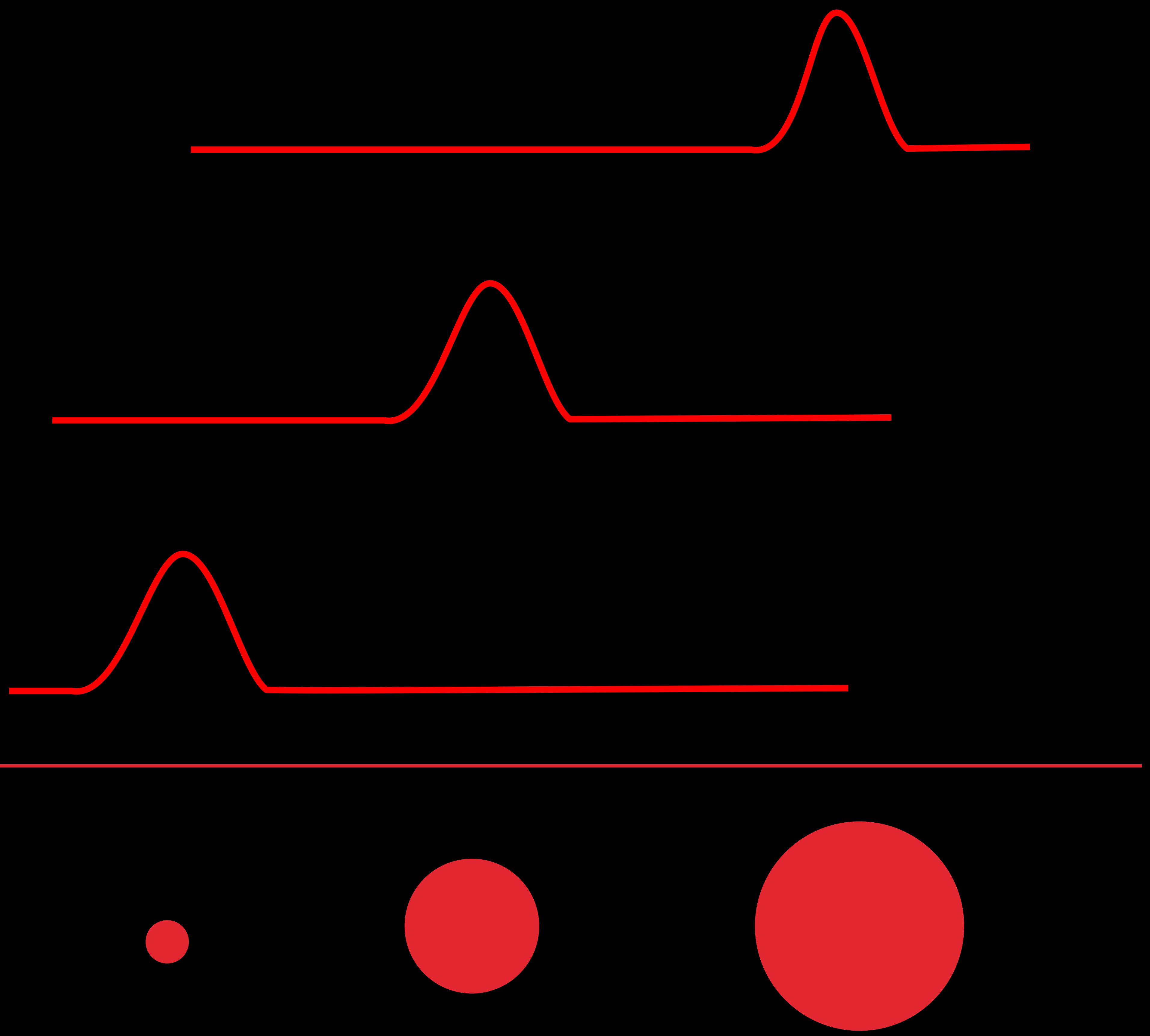
$f(I_{i_1 \dots i_m}(x,\sigma))$ $f(I_{i_1 \dots i_m}(x',\sigma'))$

K. Grauman, B. Leibe

WHAT IS A USEFUL SIGNATURE FUNCTION?

- Laplacian-of-Gaussian = "blob" detector





K. Grauman, B. Leibe

MAXIMALLY STABLE EXTREMAL REGIONS [MATAS '02]

- Based on Watershed segmentation algorithm
- Select regions that stay stable over a large parameter range

TEXT

CONCLUSION KEYPOINTS

- We want corner locations to be *invariant* to photometric transformations and *covariant* to geometric transformations
- Scale is important
- Corners and Blobs are most common, but there are others
- You can combine them, use several key point detectors.

DESCRIPTORS

DESCRIPTORS

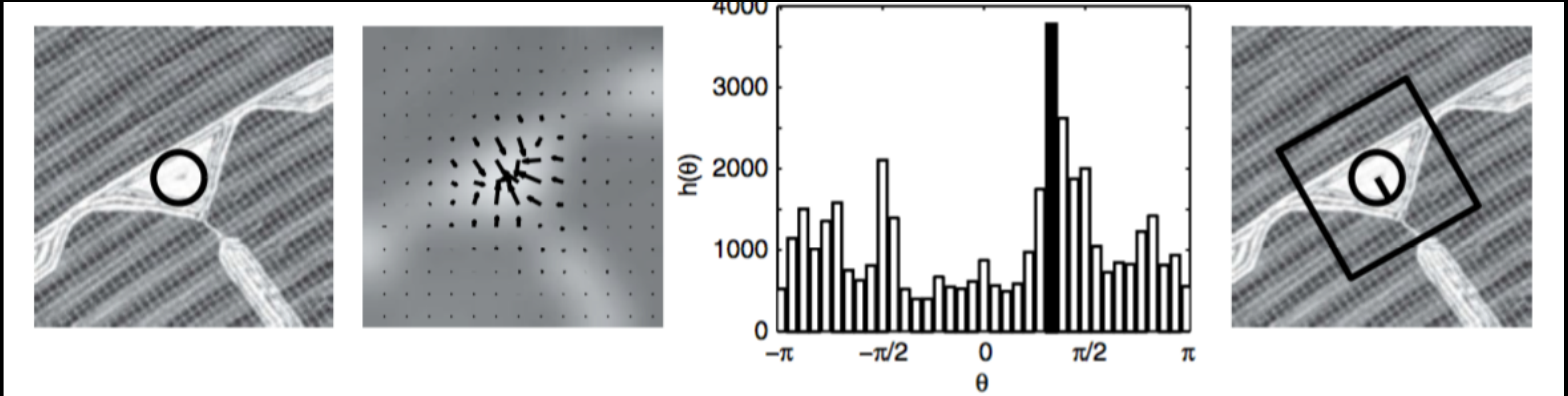
- Stereo: Patches around keypoints
  - After rectification, it's only translations, so it's fine
- What about changes in orientation, scale, affine transformations?
  - Rectification for those things… patches will still vary significantly.
  - Descriptors ***invariant*** but keeping ***discriminability***

IMAGE REPRESENTATIONS

- Templates
  - Intensity, gradients, etc.
  - SDD, NCC
- Histograms
  - Color, texture, SIFT descriptors, etc.
  - Orientation, scale, affine transformations

SIFT VECTOR FORMATION

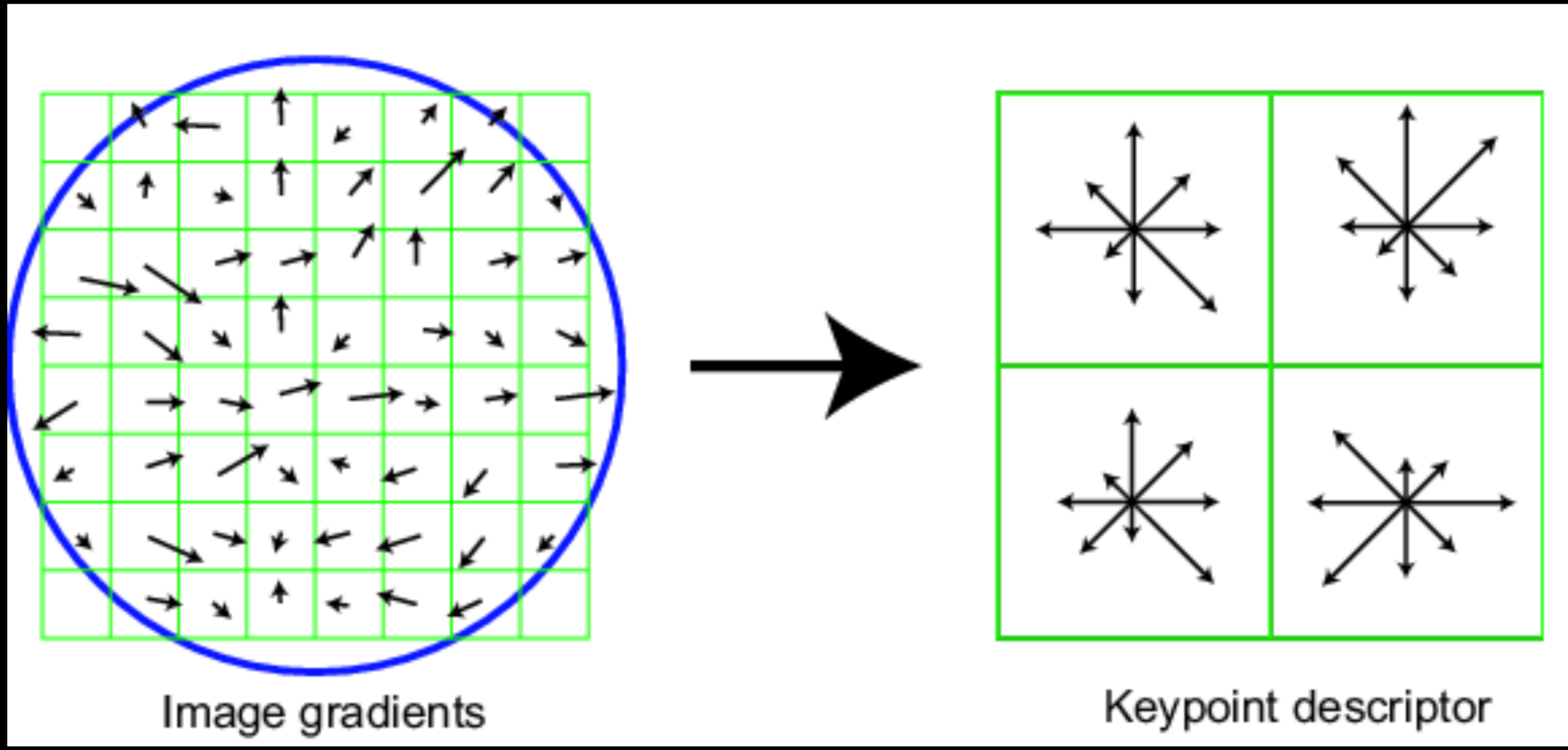
- Computed on rotated and scaled version of window according to computed orientation & scale



- 36 bins. Take maximum. If more than one peak, two keypoints are set with different orientations.

SIFT VECTOR FORMATION

- 16 by 16 window
- 4x4 array of gradient orientation histogram weighted by magnitude
- 8 orientations x 4x4 array = 128 dimensions
- Motivation: some sensitivity to spatial layout, but not too much.



showing only 2x2 here but is 4x4

OTHER DESCRIPTORS

- SURF
- PCA-SIFT
- GLOH
- Matching Shapes
- Self-similarity descriptor

CONCLUSION DESCRIPTORS

- Saliency (distinctiveness) and Repeatability
- Orientation is important
- Based on Histograms
- SIFT and SURF are most popular, but there are many others

MATCHING

FEATURE MATCHING

- Ratio distance = SSD(f1, f2) / SSD(f1, f2′)
  - f2 is best SSD match to f1 in Image2
  - f2′ is 2nd best SSD match to f1 in Image2

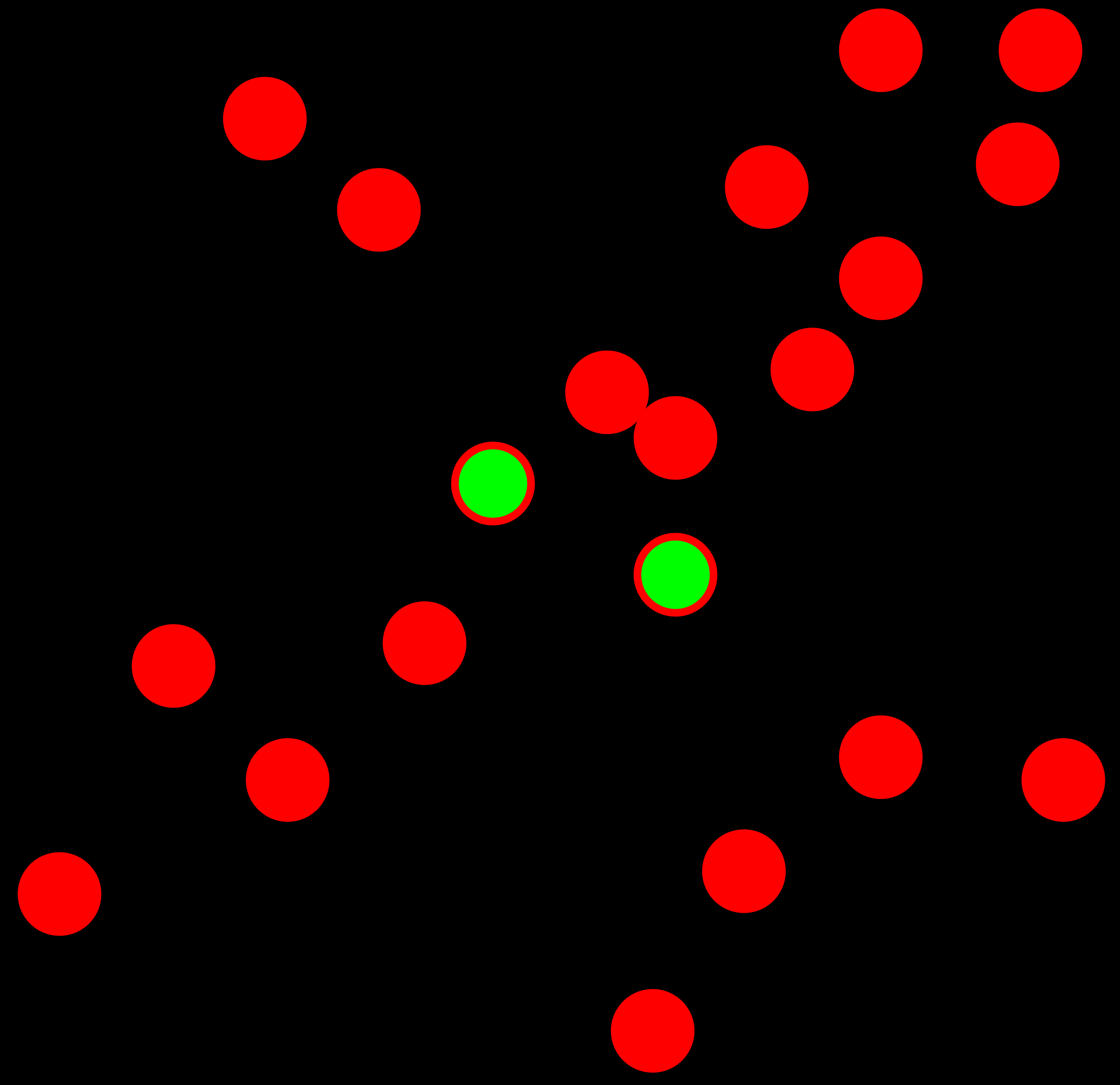
The value is in the range [0, 1.0]

NEAREST NEIGHBOR DISTANCE RATIO

- NN1/NN2
- where NN1 is the distance to the first nearest neighbor and NN2 is the distance to the second nearest neighbor
- Sorting by this ration puts matches in order of confidence

RANSAC

LINE FITTING EXAMPLE



Algorithm:

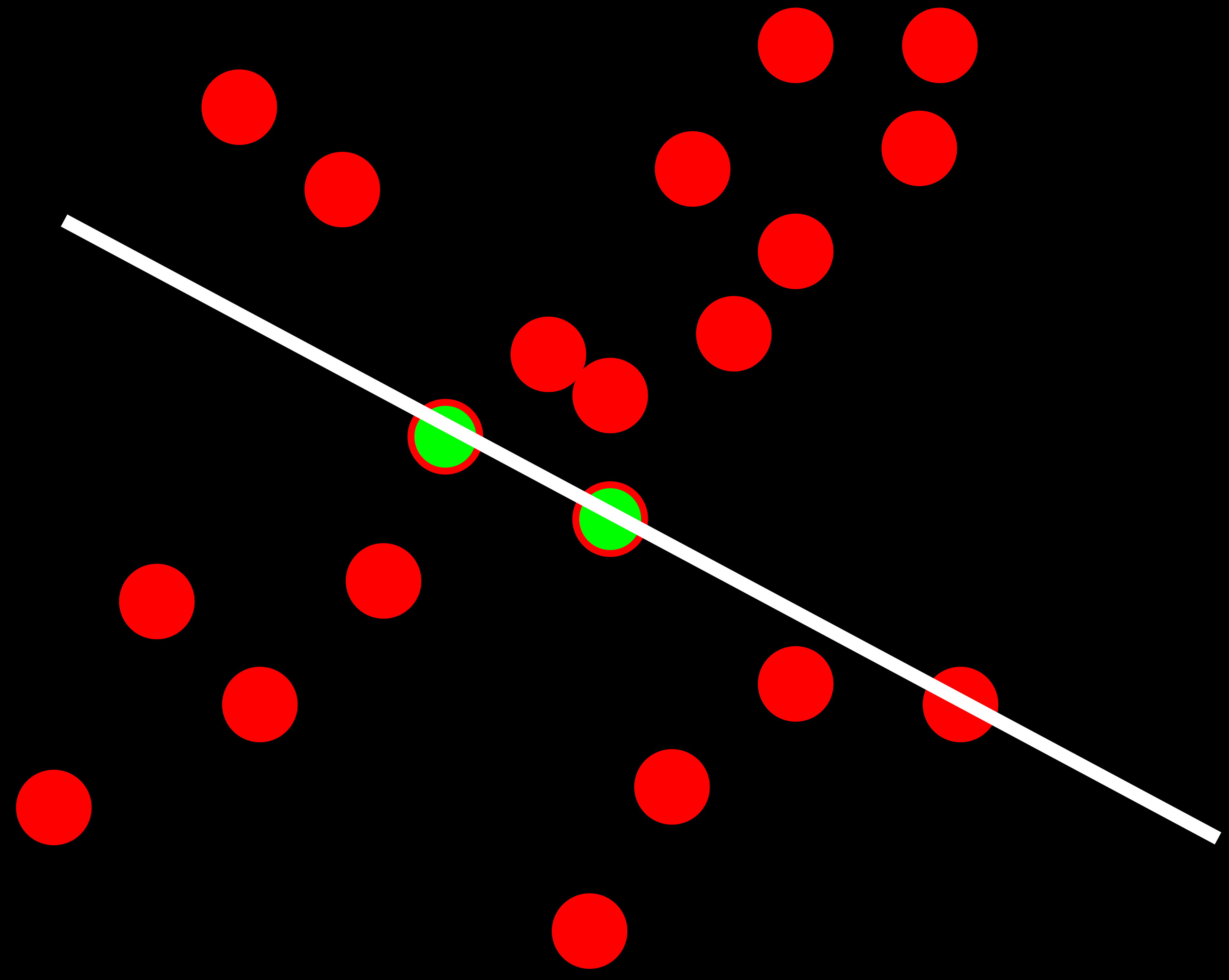
1. **Sample** (randomly) the number of points required to fit the model (#=2)
2. **Solve** for model parameters using samples
3. **Score** by the fraction of inliers within a preset threshold of the model

Repeat 1-3 until the best model is found with high confidence

Illustration by Savarese

RANSAC

LINE FITTING EXAMPLE



Algorithm:

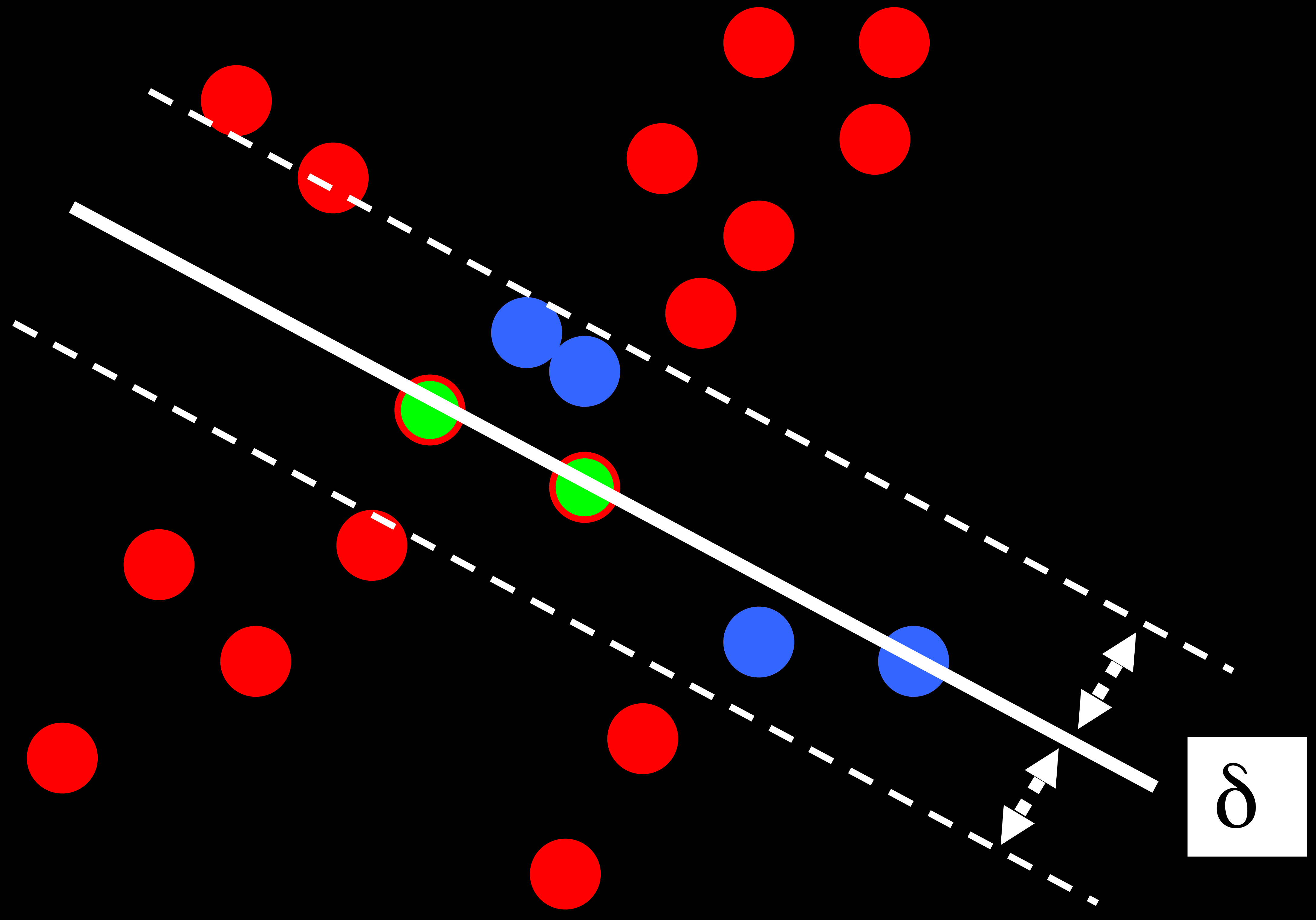
1. **Sample** (randomly) the number of points required to fit the model (#=2)
2. **Solve** for model parameters using samples
3. **Score** by the fraction of inliers within a preset threshold of the model

Repeat 1-3 until the best model is found with high confidence

RANSAC

LINE FITTING EXAMPLE

$N_I = 6$



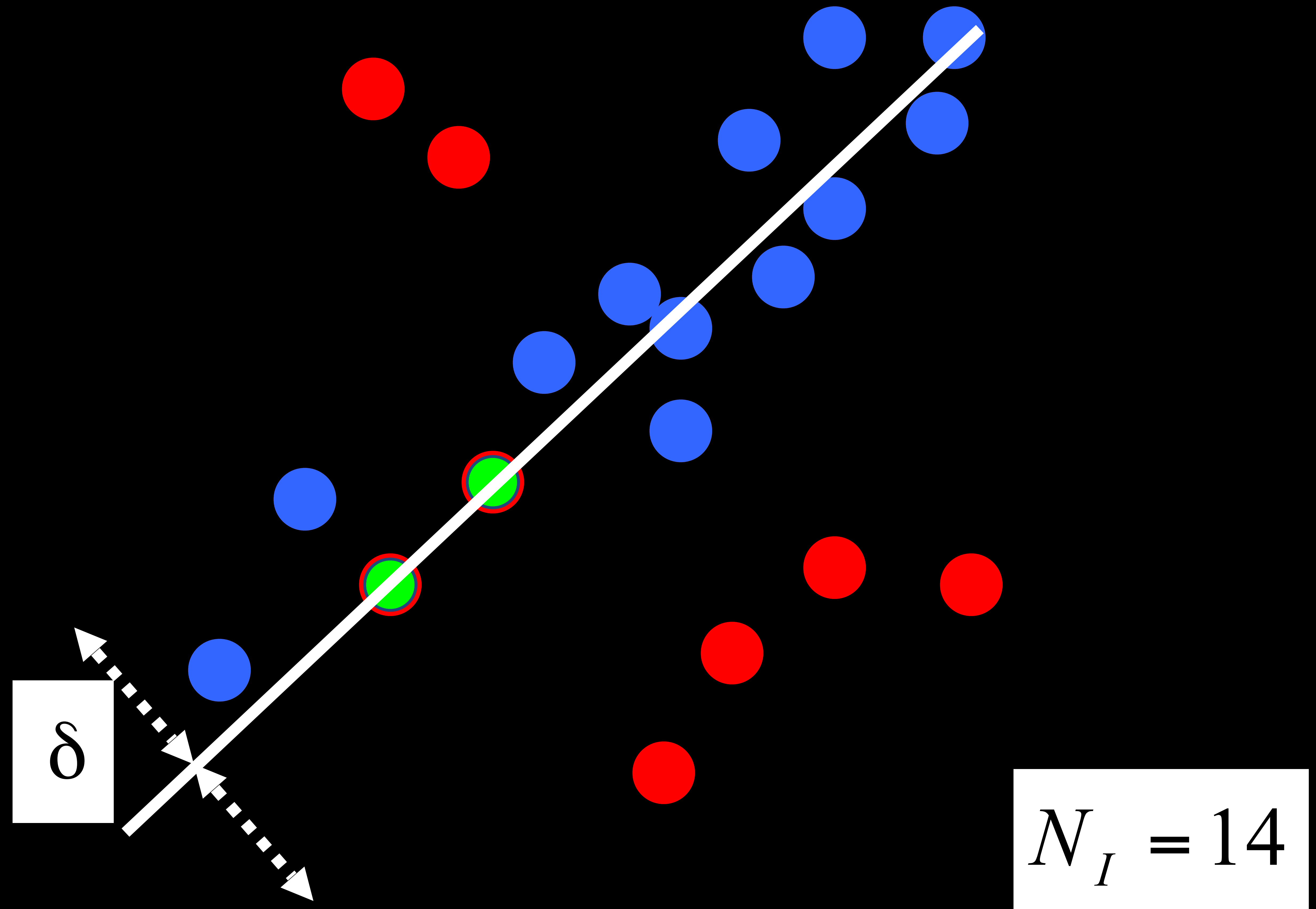
Algorithm:

1. **Sample** (randomly) the number of points required to fit the model (#=2)
2. **Solve** for model parameters using samples
3. **Score** by the fraction of inliers within a preset threshold of the model

Repeat 1-3 until the best model is found with high confidence

RANSAC

LINE FITTING EXAMPLE



Algorithm:

1. **Sample** (randomly) the number of points required to fit the model (#=2)
2. **Solve** for model parameters using samples
3. **Score** by the fraction of inliers within a preset threshold of the model

Repeat 1-3 until the best model is found with high confidence

CONCLUSION MATCHING

- Descriptors are design so L2 norm works a good metric
- Repeated features will happen, need a robust method for matching
- Confidence of matching based on uniqueness
- RANSAC, validates using consensus over a model

EPIPOLAR GEOMETRY
CAMERA MATRICES
TRIANGULATION

HOMOGENEOUS COORDINATES

Homogeneous representation of lines

$$ax + by + c = 0 \qquad (a,b,c)^{\mathsf{T}}$$

$$(ka)x + (kb)y + kc = 0, \forall k \neq 0 \qquad (a,b,c)^{\mathsf{T}} \sim k(a,b,c)^{\mathsf{T}}$$

equivalence class of vectors, any vector is representative

Set of all equivalence classes in $\mathbf{R}^3 - (0,0,0)^{\mathsf{T}}$ forms $\mathbf{P}^2$

Homogeneous representation of points

$\mathrm{x} = (x, y)^{\mathsf{T}}$ on $\mathrm{l} = (a,b,c)^{\mathsf{T}}$ if and only if $ax + by + c = 0$

$$(x,y,1)(a,b,c)^{\mathsf{T}} = (x,y,1)\mathrm{l} = 0 \qquad (x, y,1)^{\mathsf{T}} \sim k(x, y,1)^{\mathsf{T}}, \forall k \neq 0$$

The point x lies on the line l if and only if $\mathrm{x}^{\mathsf{T}}\mathrm{l}=\mathrm{l}^{\mathsf{T}}\mathrm{x}=0$

Homogeneous coordinates $(x_1, x_2, x_3)^{\mathsf{T}}$ but only 2DOF

Inhomogeneous coordinates $(x, y)^{\mathsf{T}}$

DUALITY

$$\mathbf{x} \longleftrightarrow \mathbf{l}$$

$$\mathbf{x}^{\mathsf{T}}\mathbf{l} = 0 \longleftrightarrow \mathbf{l}^{\mathsf{T}}\mathbf{x} = 0$$

$$\mathbf{x} = \mathbf{l} \times \mathbf{l}' \longleftrightarrow \mathbf{l} = \mathbf{x} \times \mathbf{x}'$$

Duality principle:

To any theorem of 2-dimensional projective geometry there corresponds a dual theorem, which may be derived by interchanging the role of points and lines in the original theorem

CAMERA MATRIX

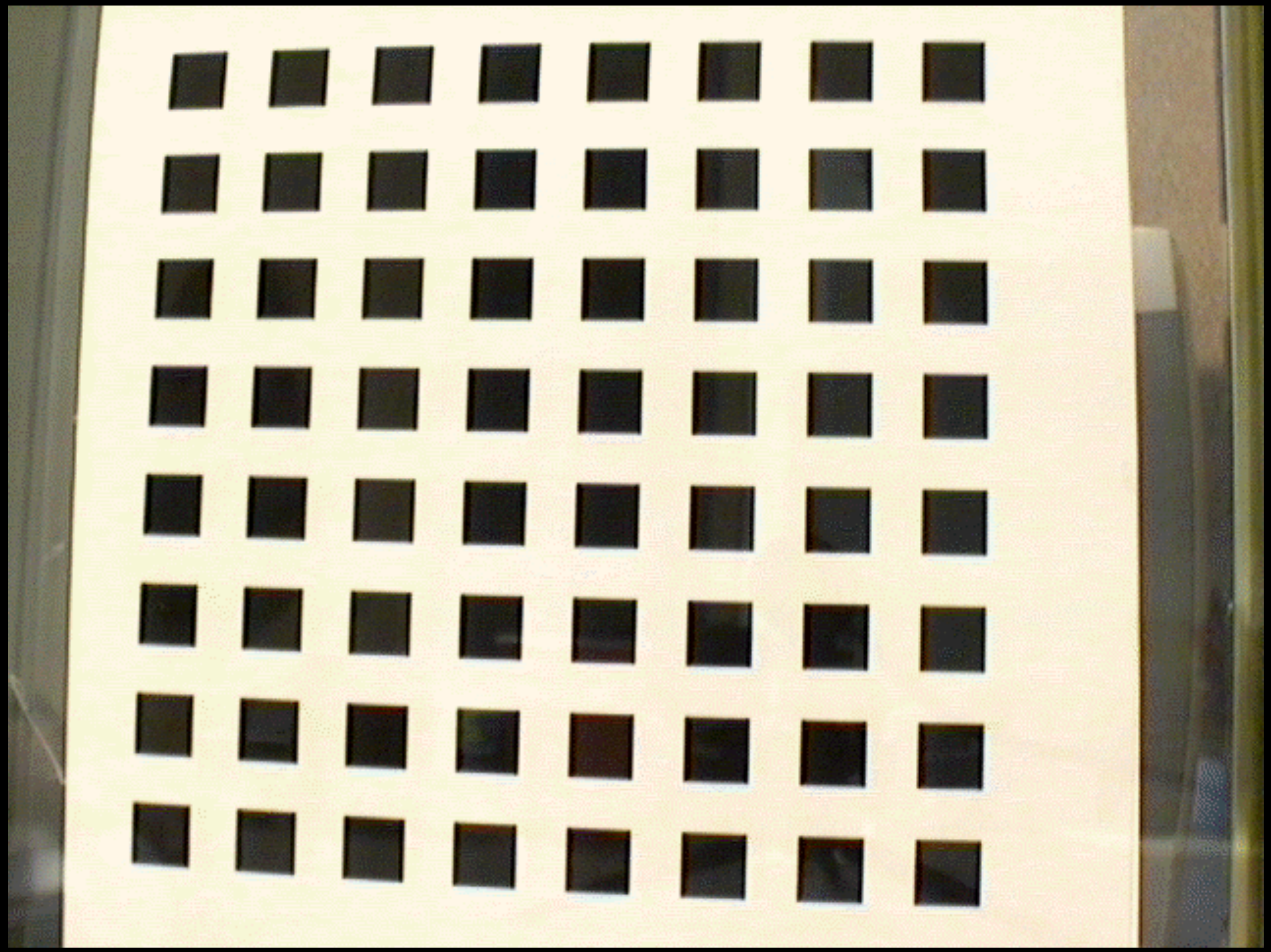
$$\mathbf{x} = \mathbf{K}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}\mathbf{X}$$

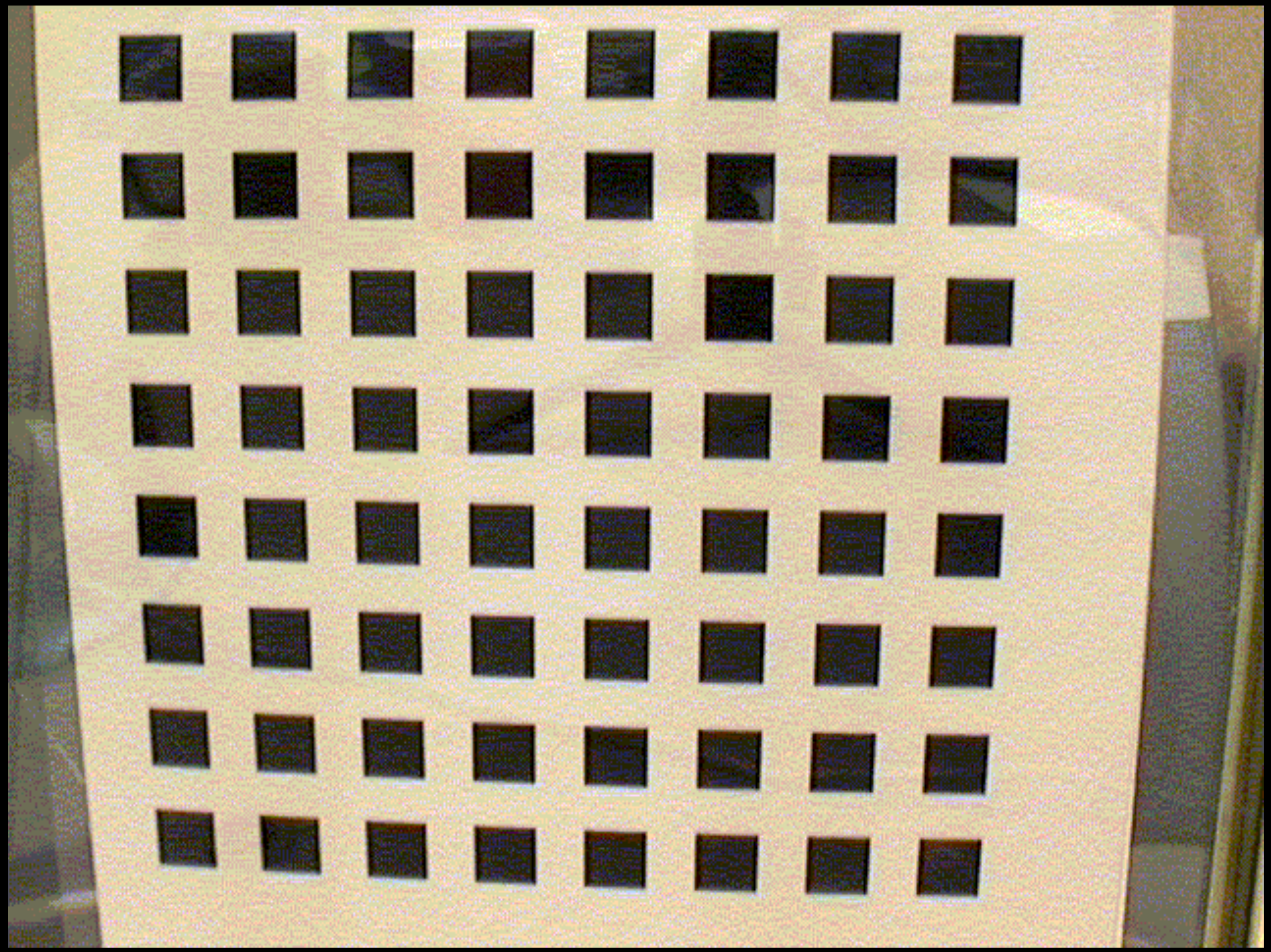
$$\Downarrow$$

$$w\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \overset{5}{\begin{bmatrix} \alpha & s & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}} \overset{6}{\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix}} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

CAMERA CALIBRATION AND LENS DISTORTION

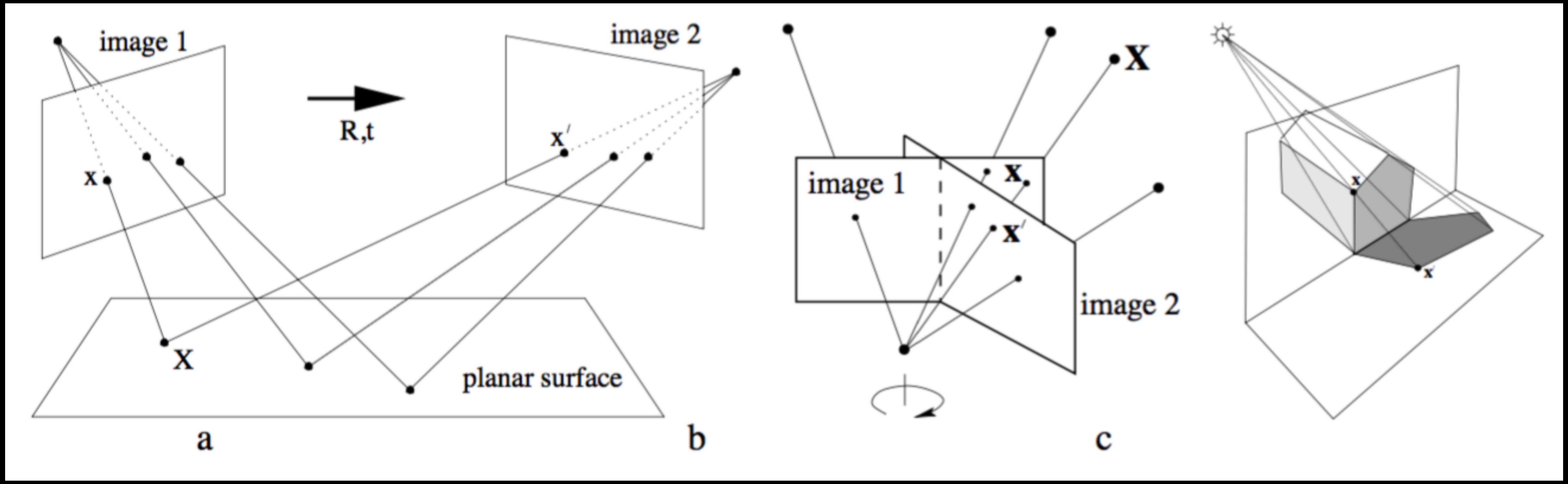
- Capture several images of a flat surface with a chessboard pattern



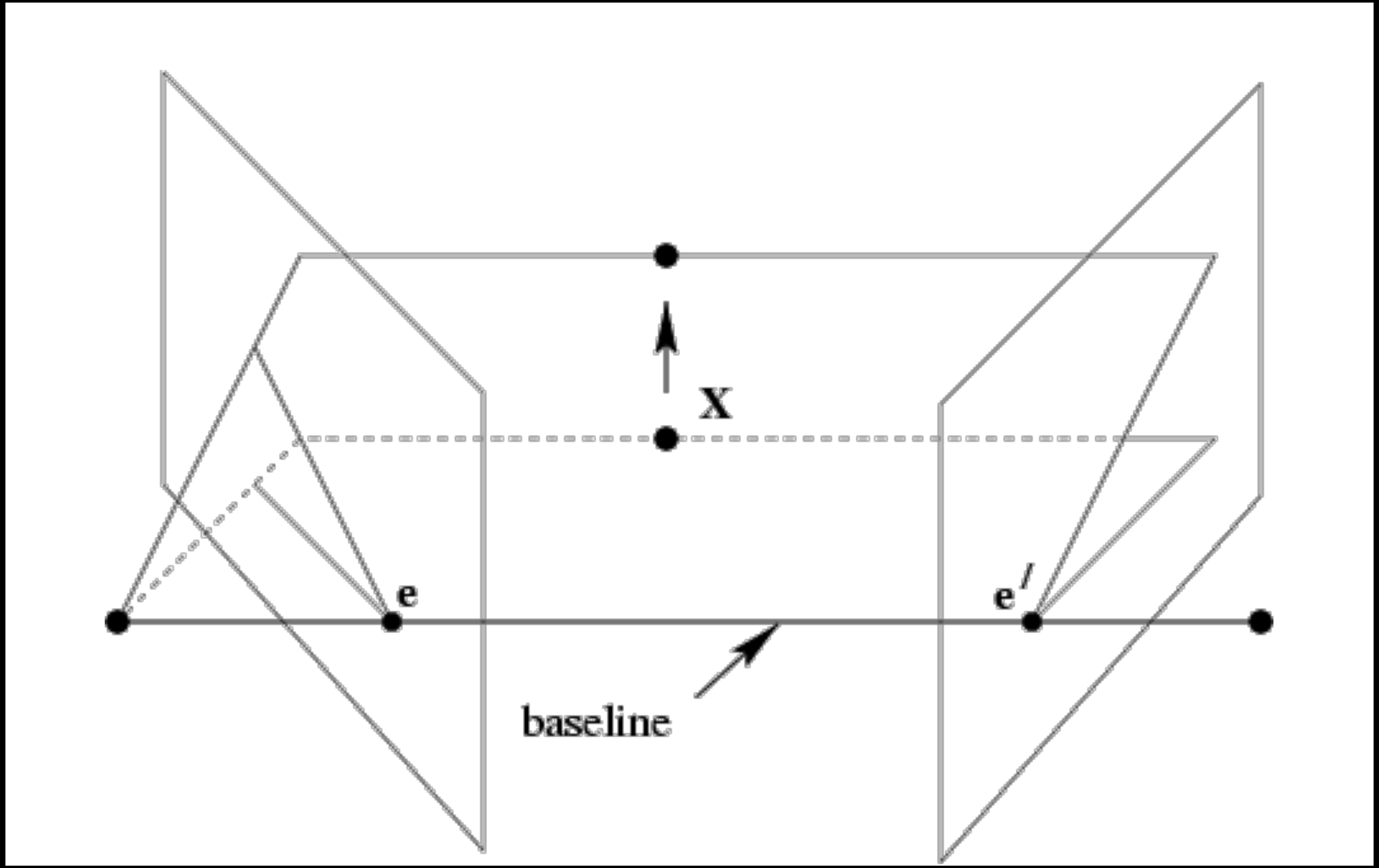


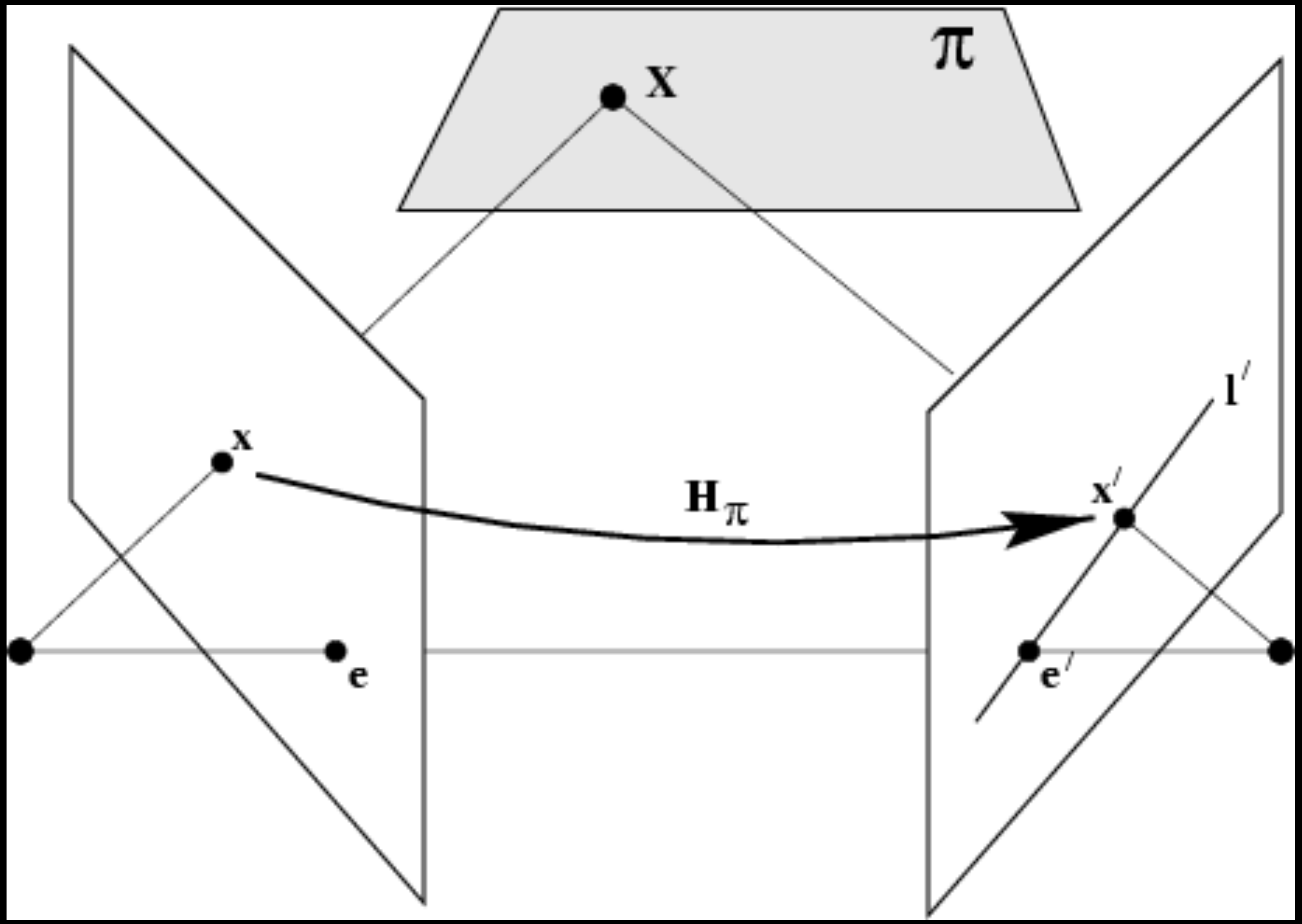
HOMOGRAPHIES

- Rotation and Translation in 3D of a plane
- Rotation on camera center
- Shadows from a point light-source



THE EPIPOLAR GEOMETRY





$$x' = H_\pi x$$

$$l' = e' \times x' = [e']_\times H_\pi x = Fx$$

FAMILY OF PLANES Π AND LINES L AND L'
INTERSECTION IN E AND E'

THE FUNDAMENTAL MATRIX F

CORRESPONDENCE CONDITION

The fundamental matrix satisfies the condition that for any pair of corresponding points $x \leftrightarrow x'$ in the two images

$$x'^{T} F x = 0 \qquad \left(x'^{T} l' = 0 \right)$$

THE FUNDAMENTAL MATRIX F

F is the unique 3x3 rank 2 matrix that satisfies $x'^TFx=0$ for all $x \leftrightarrow x'$

(i) **Transpose:** if F is fundamental matrix for (P,P'), then F^T is fundamental matrix for (P',P)

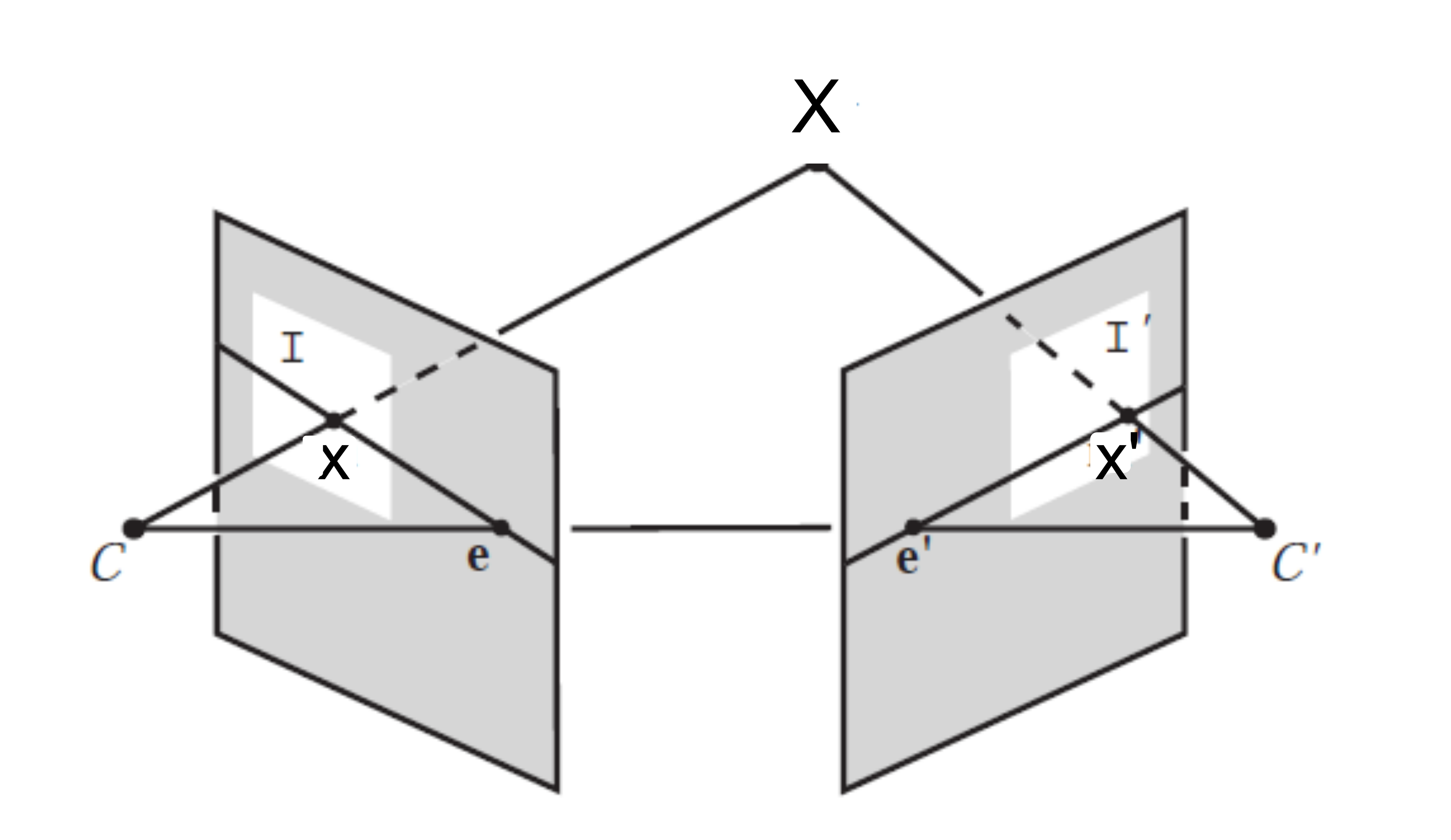
(ii) **Epipolar lines:** $l'=Fx$ & $l=F^Tx'$

(iii) **Epipoles:** on all epipolar lines, thus $e'^TFx=0, \forall x \Rightarrow e'^TF=0$, similarly $Fe=0$

(iv) **F** has 7 d.o.f. , i.e. 3x3-1(homogeneous)-1(rank2)

"GOLD STANDARD" ALGORITHM

- Use 8-point algorithm to get initial value of F
- Use F to solve for P and P′
- Jointly solve for 3d points **X** and **F** that minimize the squared re-projection error



SEE ALGORITHM 11.2 AND ALGORITHM 11.3 IN HZ (PAGES 284–285) FOR DETAILS

ESSENTIAL MATRIX AND TRIANGULATION

- Similar to the fundamental Matrix, we have the Essential Matrix
- We can compute a P1 and P2 from F or E
- and triangulate the points based on these.
- With F -> Projective Reconstruction
- With E -> Similarity Reconstruction

LINKING MORE THAN 2 VIEWS

- Get P matrices from F or E.
- Triangulate Points
- Compute a new P from the 3D points.
- This produces a sparse point cloud.
- If we have a calibrated case, this is a metric reconstruction, if not it is a projective reconstruction, which might contain severe distortions

TAKING PICTURES
FOR 3D RECONSTRUCTION

TAKING PICTURES FOR RECONSTRUCTION

- Depends on features
- Necessary Overlap
  - SIFT features allow for “wide baseline” (up to 30°)
- Materials that change with viewpoint will cause problems
- Blur, movement, will cause problems.
- Pure rotation or small baseline and fully planar scenes will cause problems, need to be taken into account.

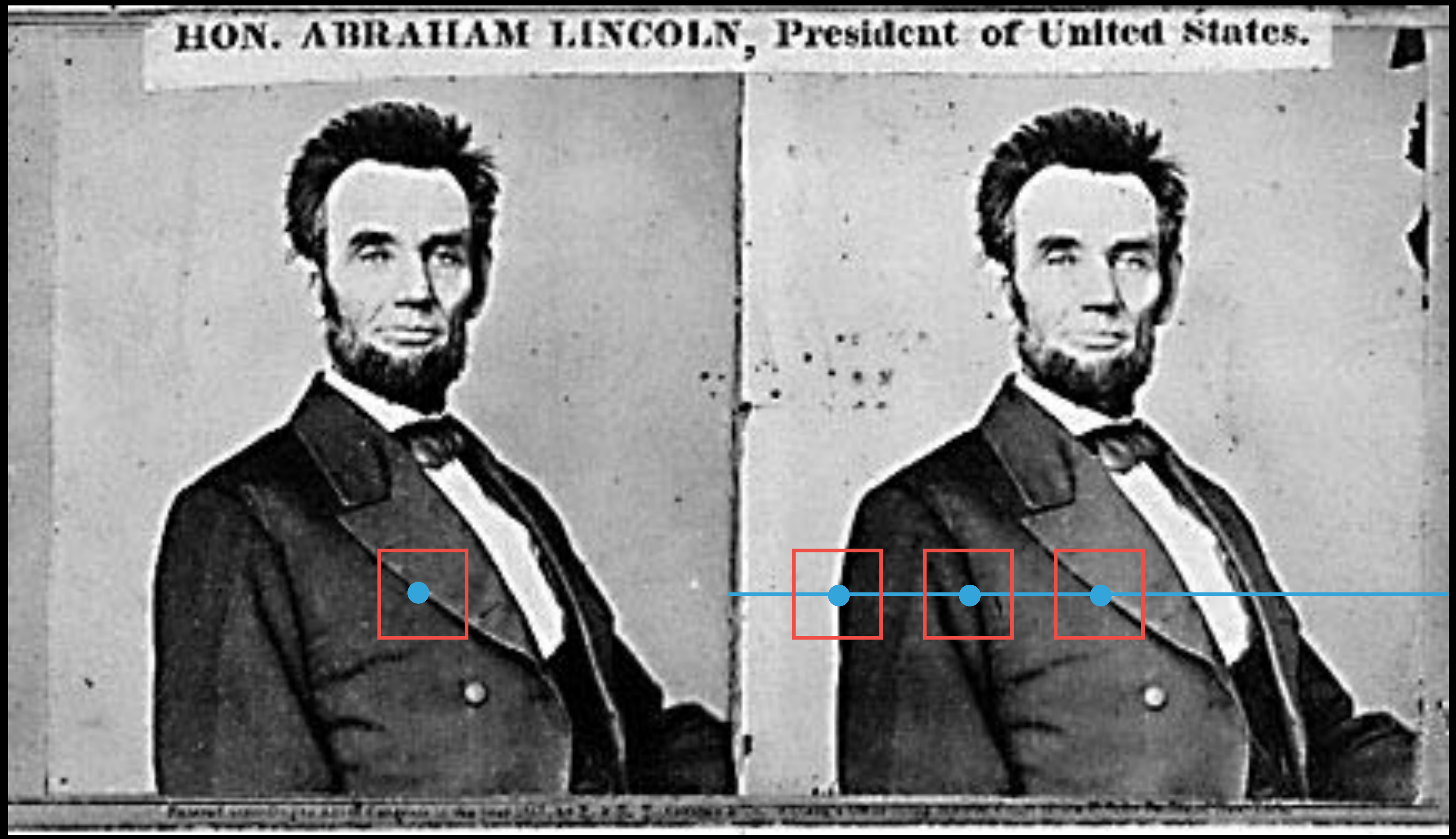
SAMPLES OF IMAGES

DENSE MATCHING
STEREO

BASIC STEREO MATCHING ALGORITHM

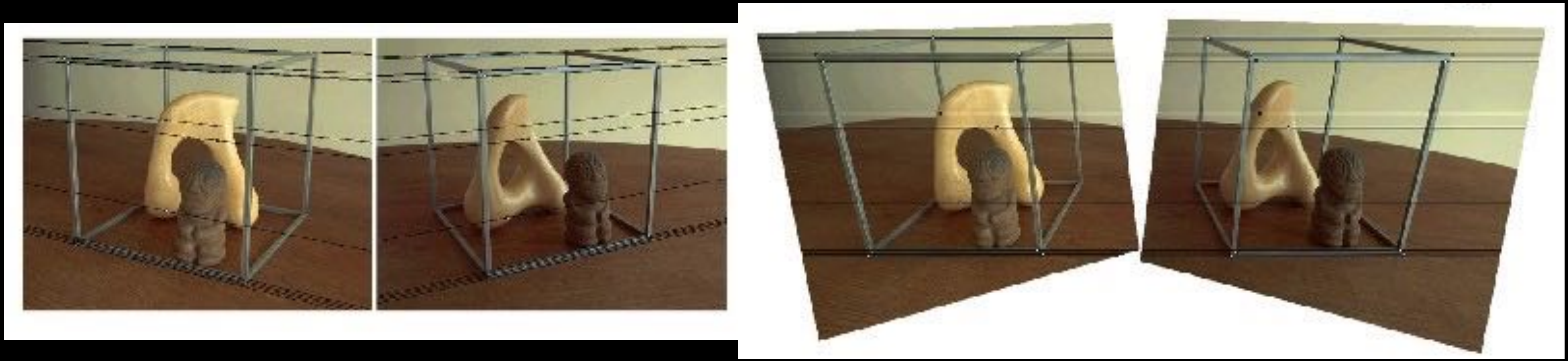
- For each pixel in the first image
  - Find corresponding epipolar line in the right image
  - Examine all pixels on the epipolar line and pick the best match
  - Triangulate the matches to get depth information

- Simplest case: epipolar lines are scanlines
  - When does this happen?



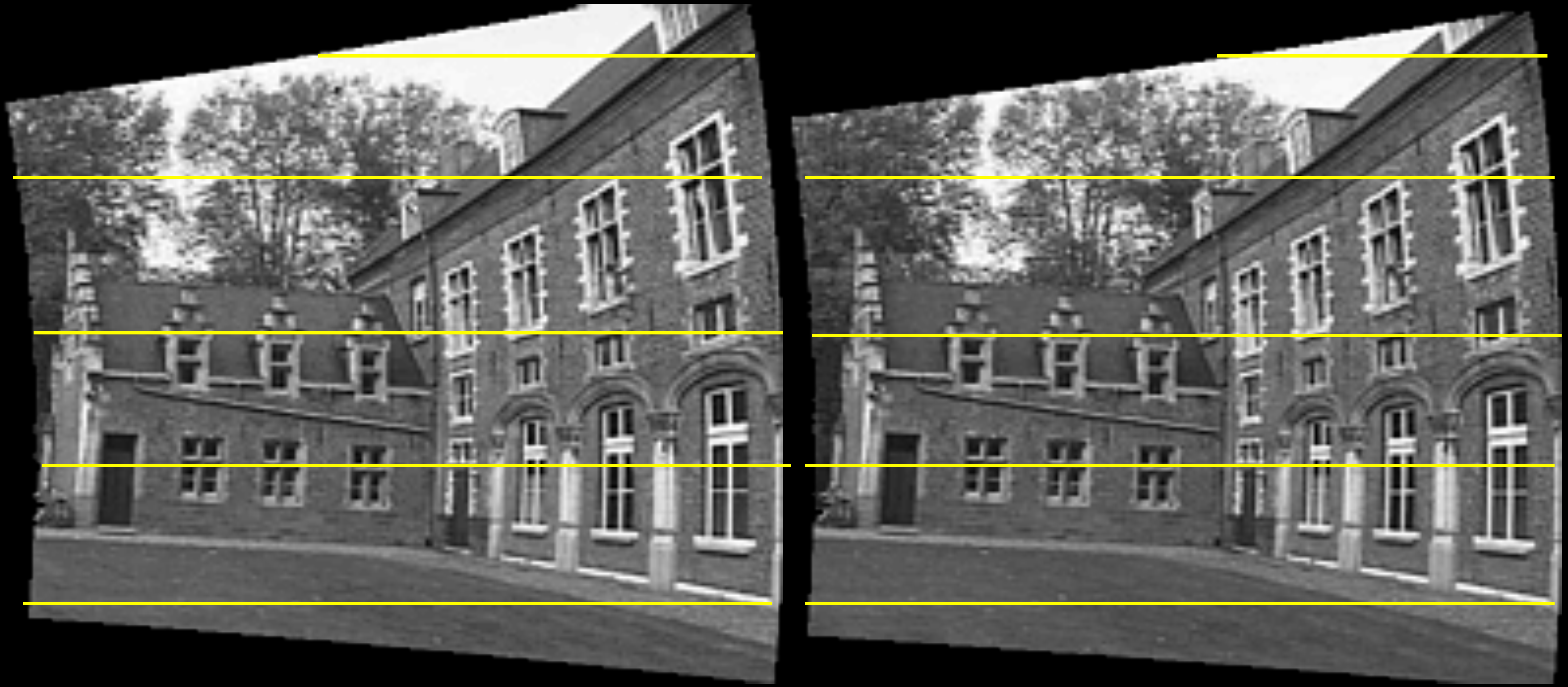
RECTIFICATION

- An homography so epipolar lines are horizontal and at the same vertical coordinate
- Try to minimise distortion



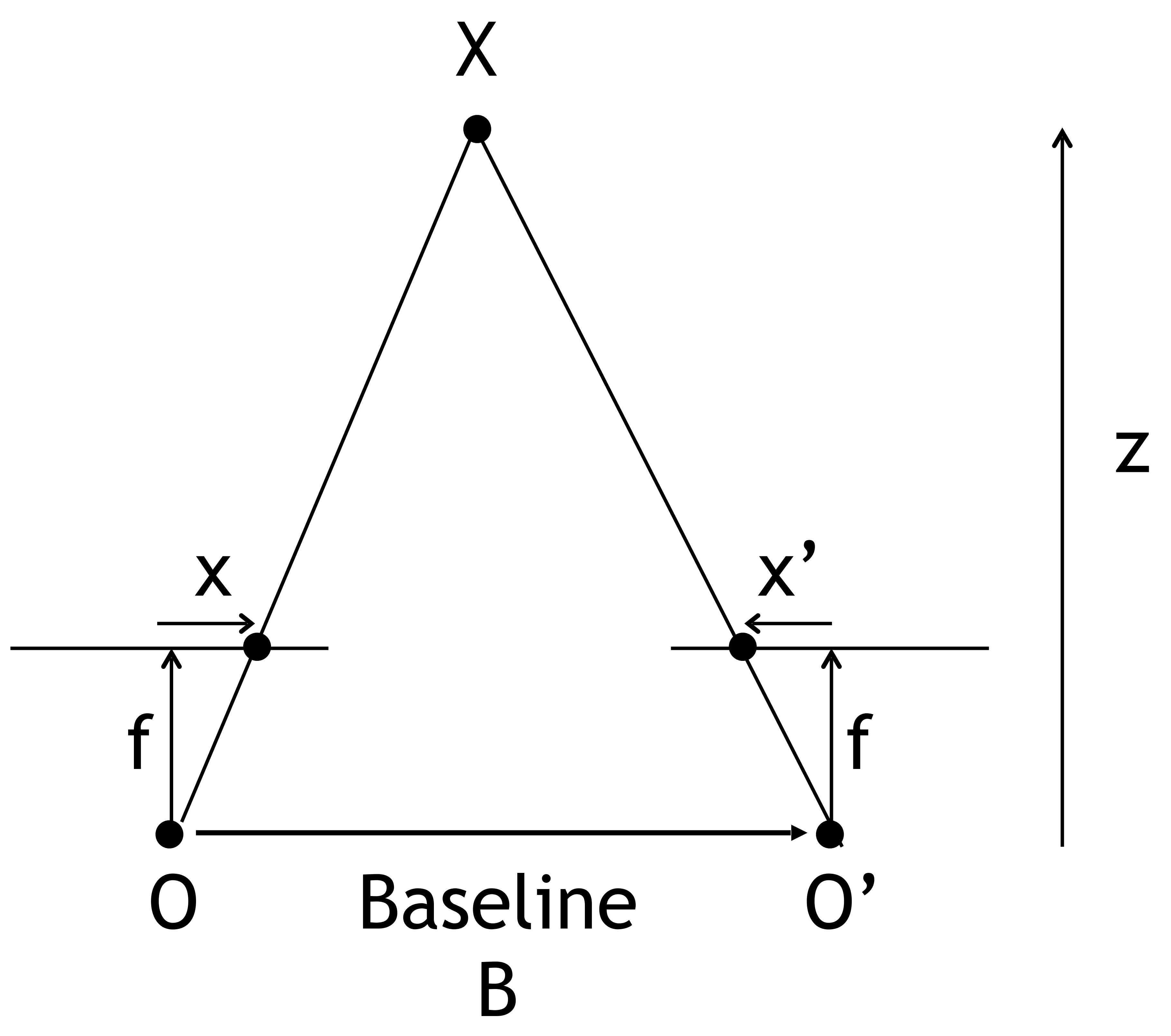
POLAR RECTIFICATION: EXAMPLE

- Works with epipoles inside the image



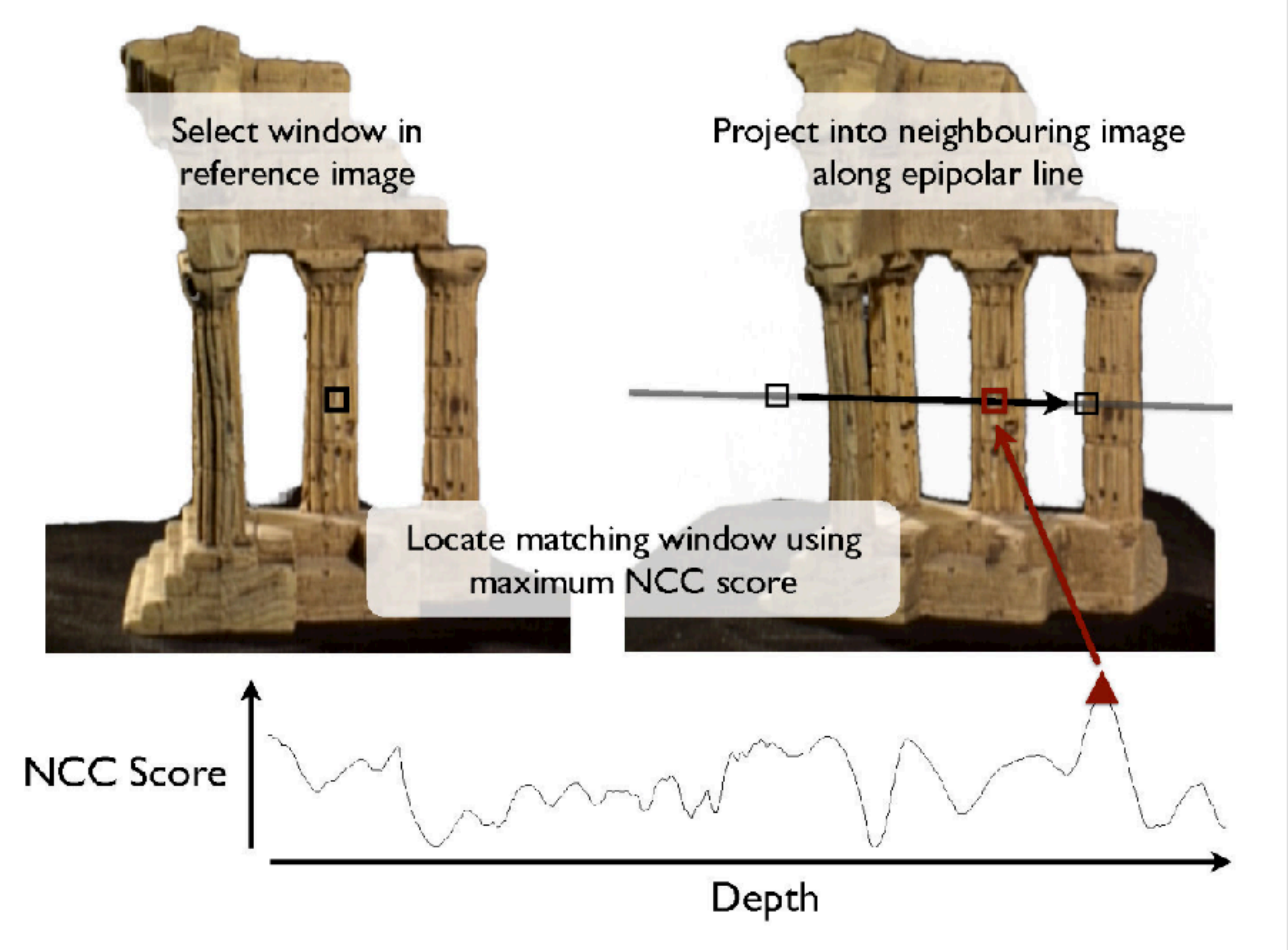
DEPTH FROM DISPARITY

- Disparity is inversely proportional to depth!



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

MULTI-VIEW STEREO

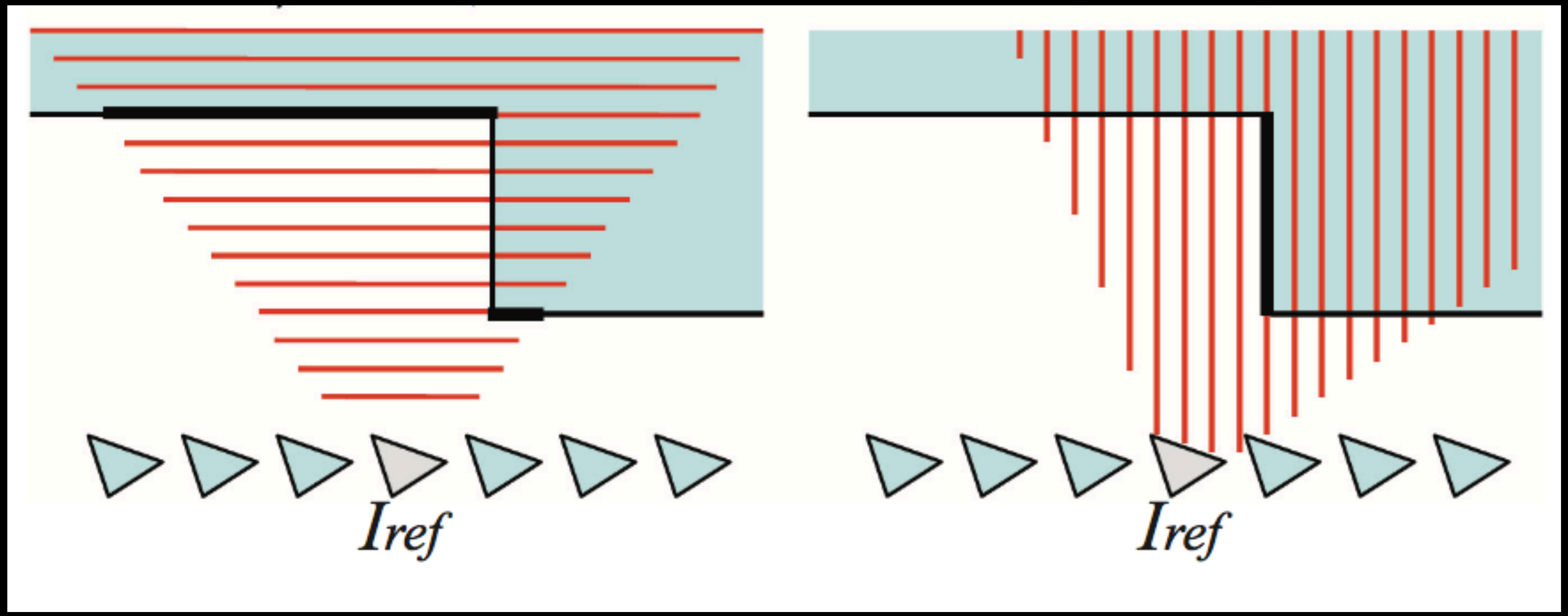


Algorithm 1 Simple depthmap estimation algorithm

procedure SIMPLEDEPTHMAP(I_{ref},$N(I_{ref})$,$\mathbb{D}$)
// I_{ref}: reference image
// $N(I_{ref})$: neighboring images of I_{ref}
// $\mathbb{D}$: depth range
    for all $p \in I_{ref}$ **do** // p: pixel
        $\mathcal{C}_{best} \leftarrow -\infty$
        for all $d \in \mathbb{D}$ **do**
            if $\mathcal{C}(p,d) > \mathcal{C}_{best}$ **then**
                $\mathcal{C}_{best} \leftarrow \mathcal{C}(p,d)$
                $d_{best} \leftarrow d$
                $f_{best} \leftarrow \mathcal{F}(p,d)$ // confidence value
            end if
        end for
        return (p, d_{best}, f_{best})
    end for
end procedure

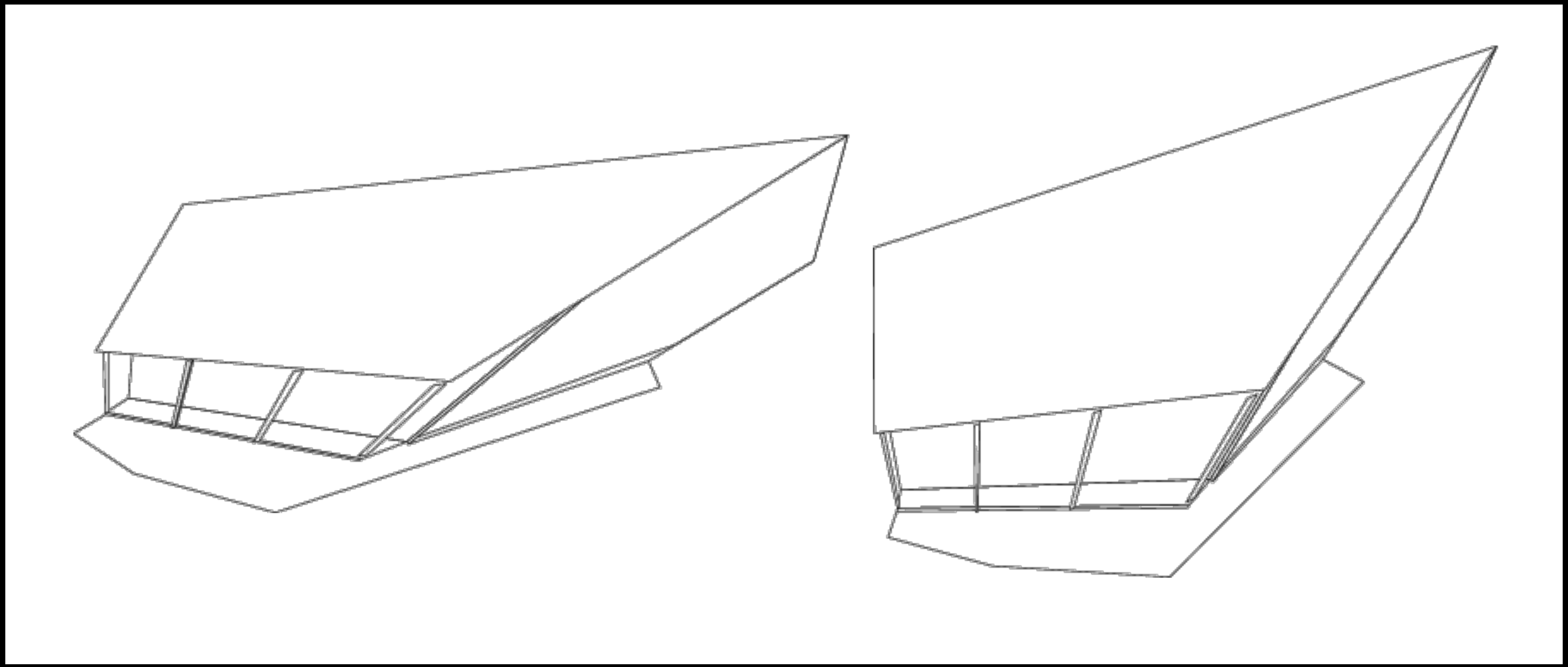
PLANE SWEEPING DEPTHMAP RECONSTRUCTION

- Projects images onto a plane via planar homographies, then evaluates photo-consistency values on each plane.
- Directions, which can be extracted from sparse 3D point cloud reconstructed by the SfM system.

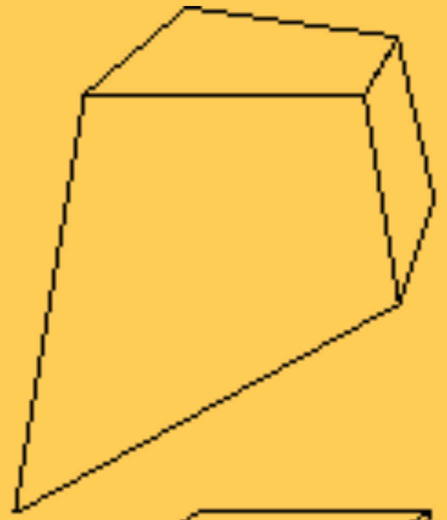
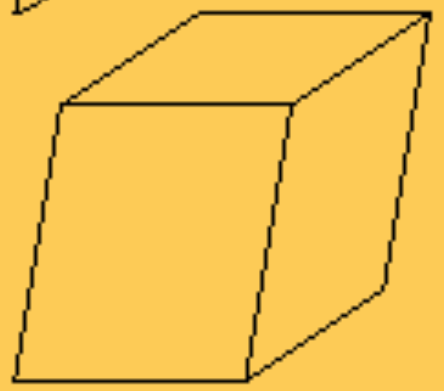
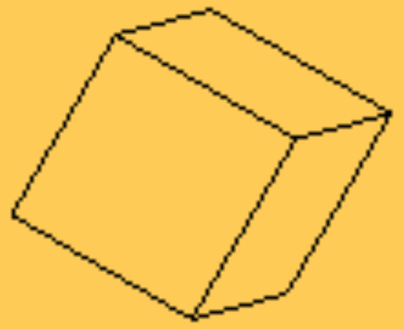
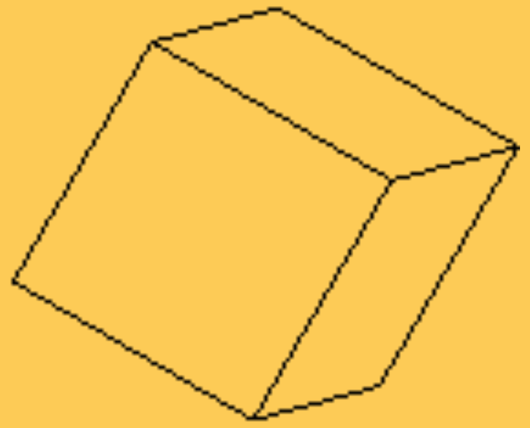


BUNDLE ADJUSTMENT, AUTO-CALIBRATION

TEXT



HIERARCHY OF TRANSFORMATIONS

Projective 15dof	$\begin{bmatrix} A & t \\ v^T & v \end{bmatrix}$		Intersection and tangency
Affine 12dof	$\begin{bmatrix} A & t \\ 0^T & 1 \end{bmatrix}$		Parallellism of planes, Volume ratios, centroids, **The plane at infinity π_∞**
Similarity 7dof	$\begin{bmatrix} sR & t \\ 0^T & 1 \end{bmatrix}$		**The absolute conic Ω_∞**
Euclidean 6dof	$\begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix}$		Volume

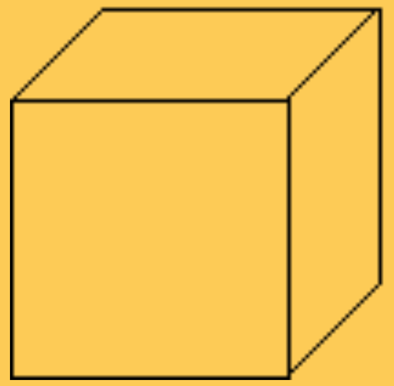
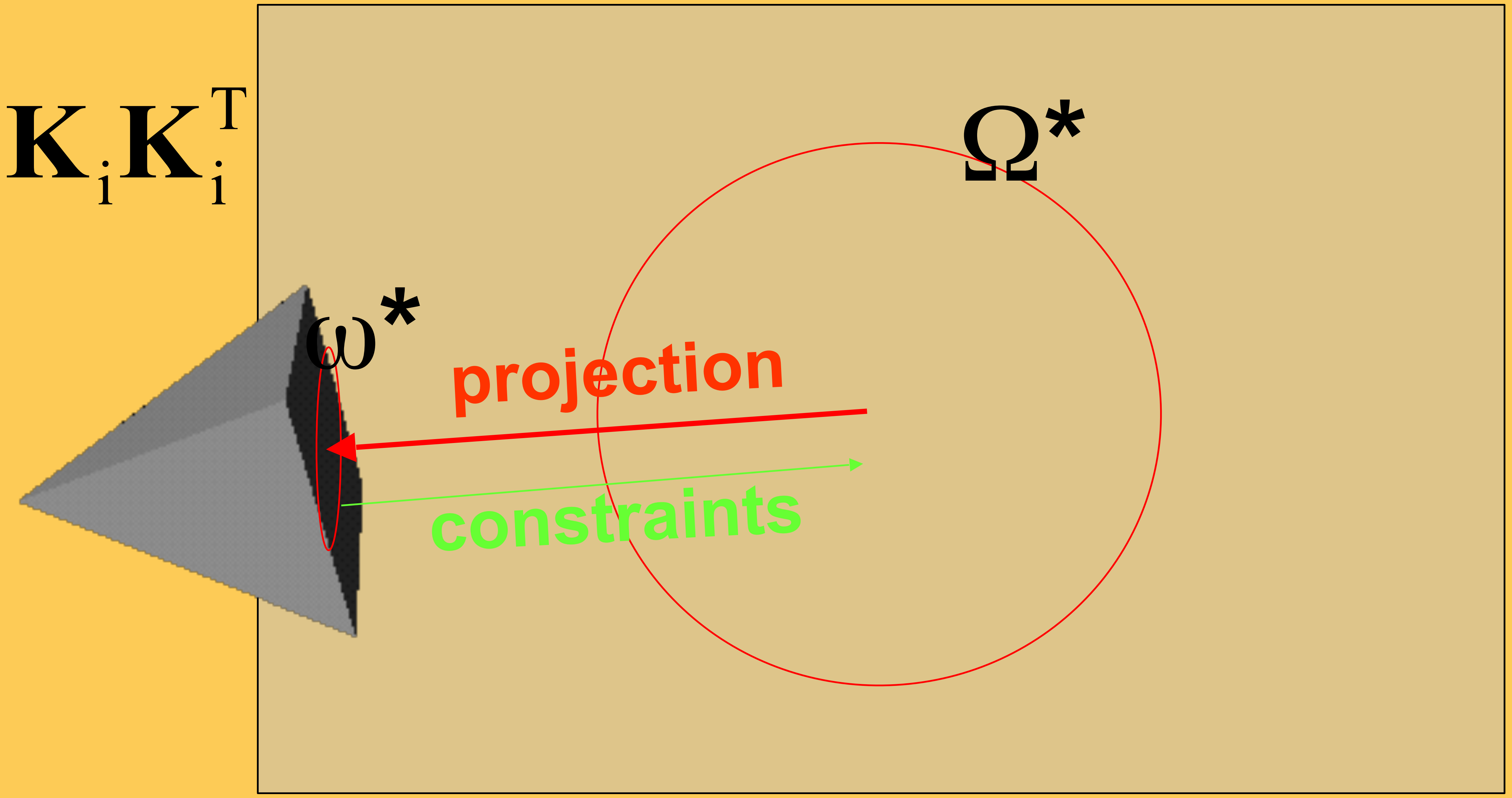


IMAGE OF THE ABSOLUTE CONIC

Projection equation:

$$\omega_i^* \propto P_i \Omega^* P_i^T \propto K_i K_i^T$$

Translate constraints on **K** through projection equation to constraints on Dual Absolute Quadric Ω^*



Absolute conic = calibration object which is always present but can only be observed through constraints on the intrinsics

CONSTRAINTS ON Ω^*_∞

$$\omega^*_\infty = \begin{bmatrix} f_x^2 + s^2 + c_x^2 & sf_y + c_x c_y & c_x \\ sf_y + c_x c_y & f_y^2 + c_y^2 & c_y \\ c_x & c_y & 1 \end{bmatrix}$$

condition	constraint	type	#constraints
Zero skew	$\omega^*_{12}\omega^*_{33} = \omega^*_{13}\omega^*_{23}$	quadratic	m
Principal point	$\omega^*_{13} = \omega^*_{23} = 0$	linear	$2m$
Zero skew (& p.p.)	$\omega^*_{12} = 0$	linear	m
Fixed aspect ratio (& p.p.& Skew)	$\omega^*_{11}\omega'^*_{22} = \omega^*_{22}\omega'^*_{11}$	quadratic	$m-1$
Known aspect ratio (& p.p.& Skew)	$\omega^*_{11} = \omega^*_{22}$	linear	m
Focal length (& p.p. & Skew)	$\omega^*_{33} = \omega^*_{11}$	linear	m

PROJECTIVE TO METRIC

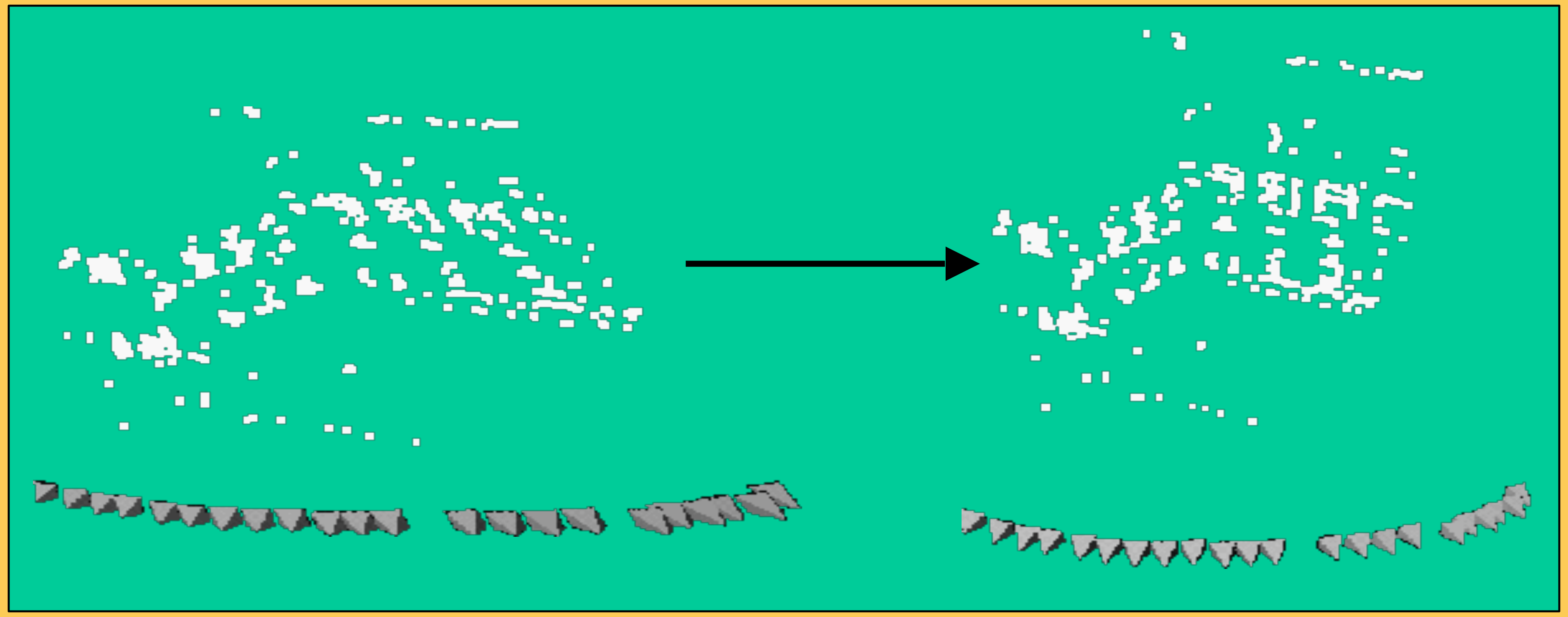
Compute $\mathbf{T}$ from

$$\tilde{\mathbf{I}} = \mathbf{T}\Omega^*_\infty \mathbf{T}^\mathsf{T} \text{ or } \mathbf{T}^{-1}\tilde{\mathbf{I}}\mathbf{T}^{-\mathsf{T}} = \Omega^*_\infty \text{ with } \tilde{\mathbf{I}} = \begin{bmatrix} \mathbf{I} & 0 \\ 0^\mathsf{T} & 0 \end{bmatrix}$$

using eigenvalue decomposition of Ω^*_∞

and then obtain metric reconstruction as

$$\mathbf{PT}^{-1} \text{ and } \mathbf{TM}$$



REFINING STRUCTURE AND MOTION

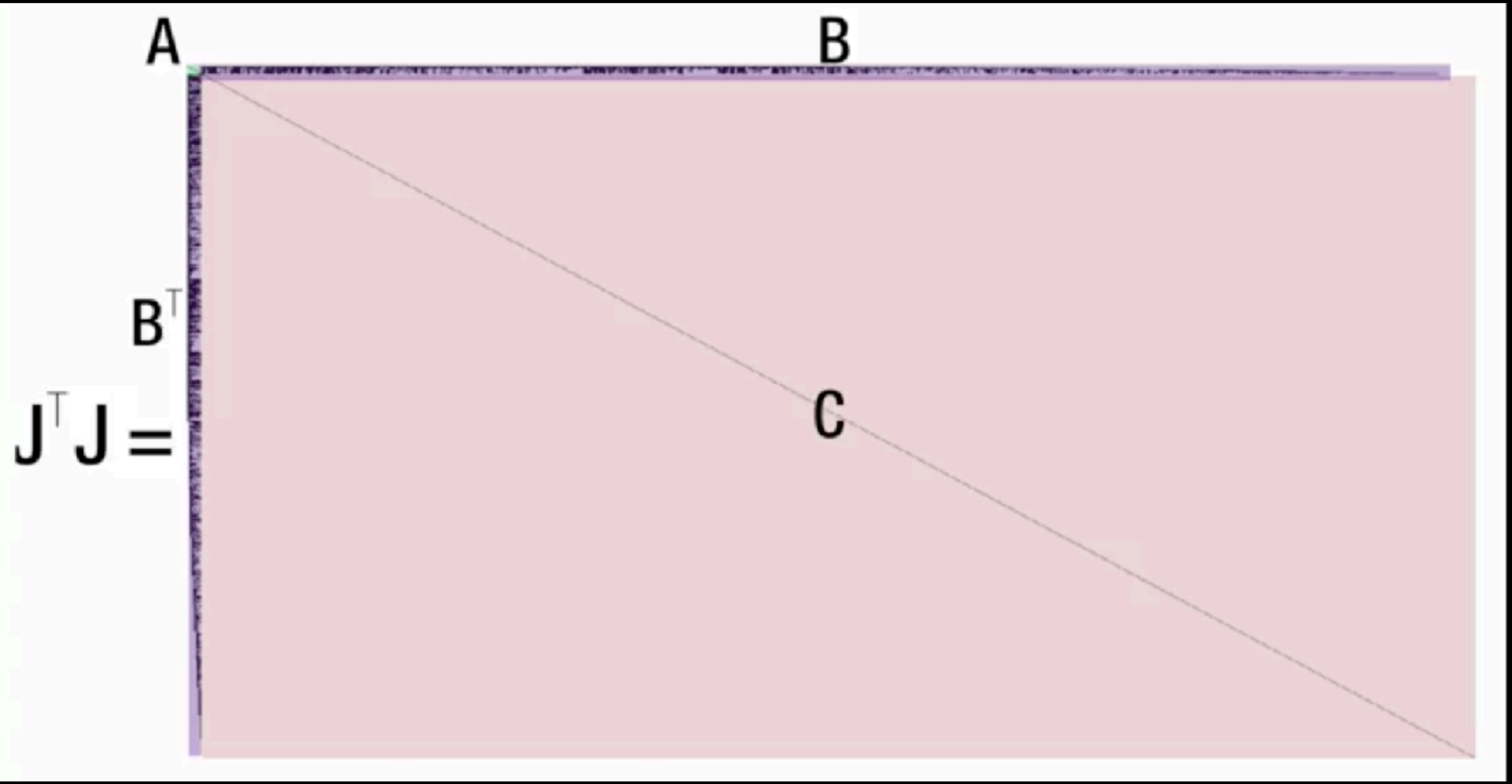
- Minimize reprojection error

$$\min_{\hat{P}_k, \hat{M}_i} \sum_{k=1}^{m} \sum_{i=1}^{n} D\left(m_{ki}, \hat{P}_k \hat{M}_i\right)^2$$

  - Maximum Likelyhood Estimation (if error zero-mean Gaussian noise)
  - Huge problem but can be solved efficiently (Bundle adjustment)
  - For a typical sequence of 20 views and 2000 points, a minimization problem in more than 6000 variables has to be solved

SPARSE LEVENBERG-MARQUARDT

- Takes advantage of the sparsity of the system, cameras are independent among them and points too.



CONCLUSION

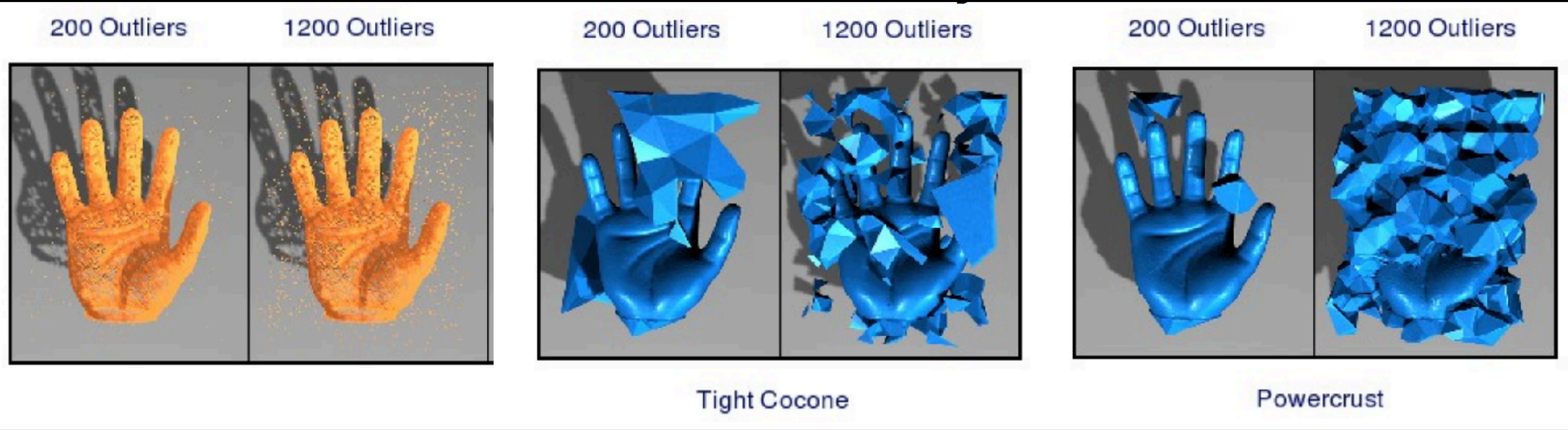
- Despite Linear Optimization can result in fair models, they are not optimal
- Non-linear optimisation techniques to refine the estimations of linear methods are typically used for camera calibration, fundamental matrix estimation, triangulation and BA

MESHING

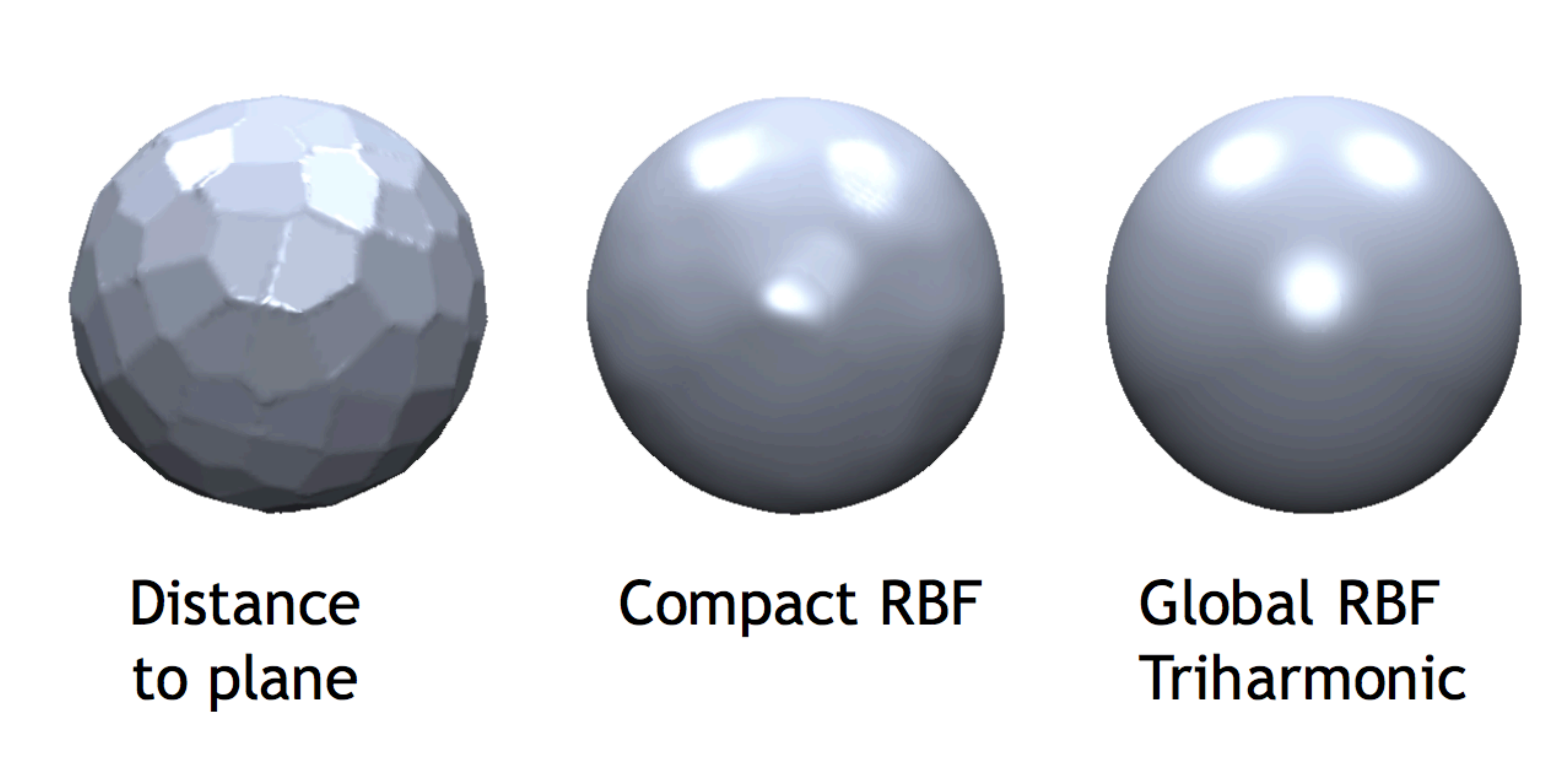
TWO MAIN APPROACHES

- Explicit:
  - Local surface connectivity estimation
  - Point interpolation
- Implicit:
  - Signed distance function estimation
  - Mesh approximation

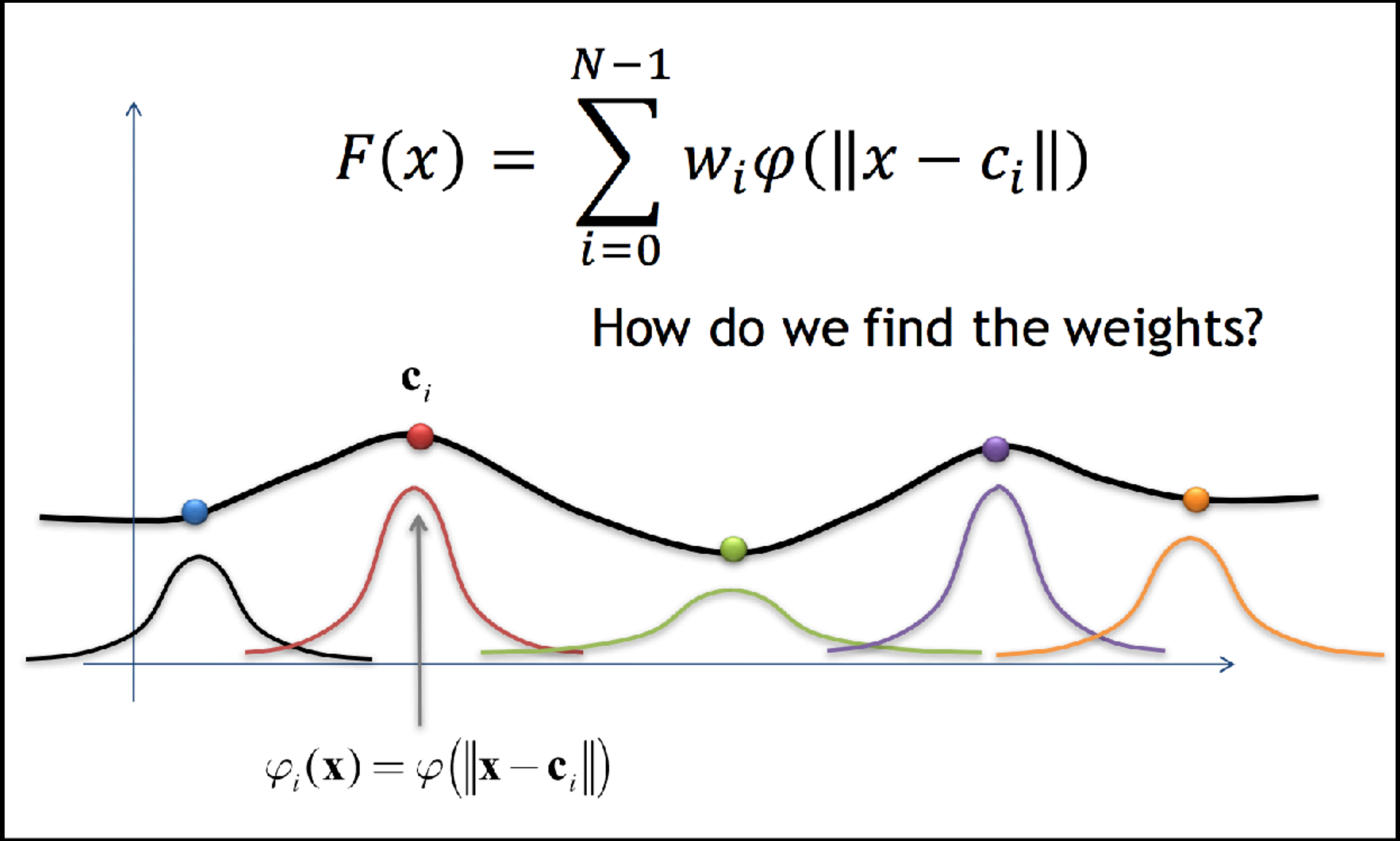
DELAUNAY TRIANGULATION



RBF KERNELS

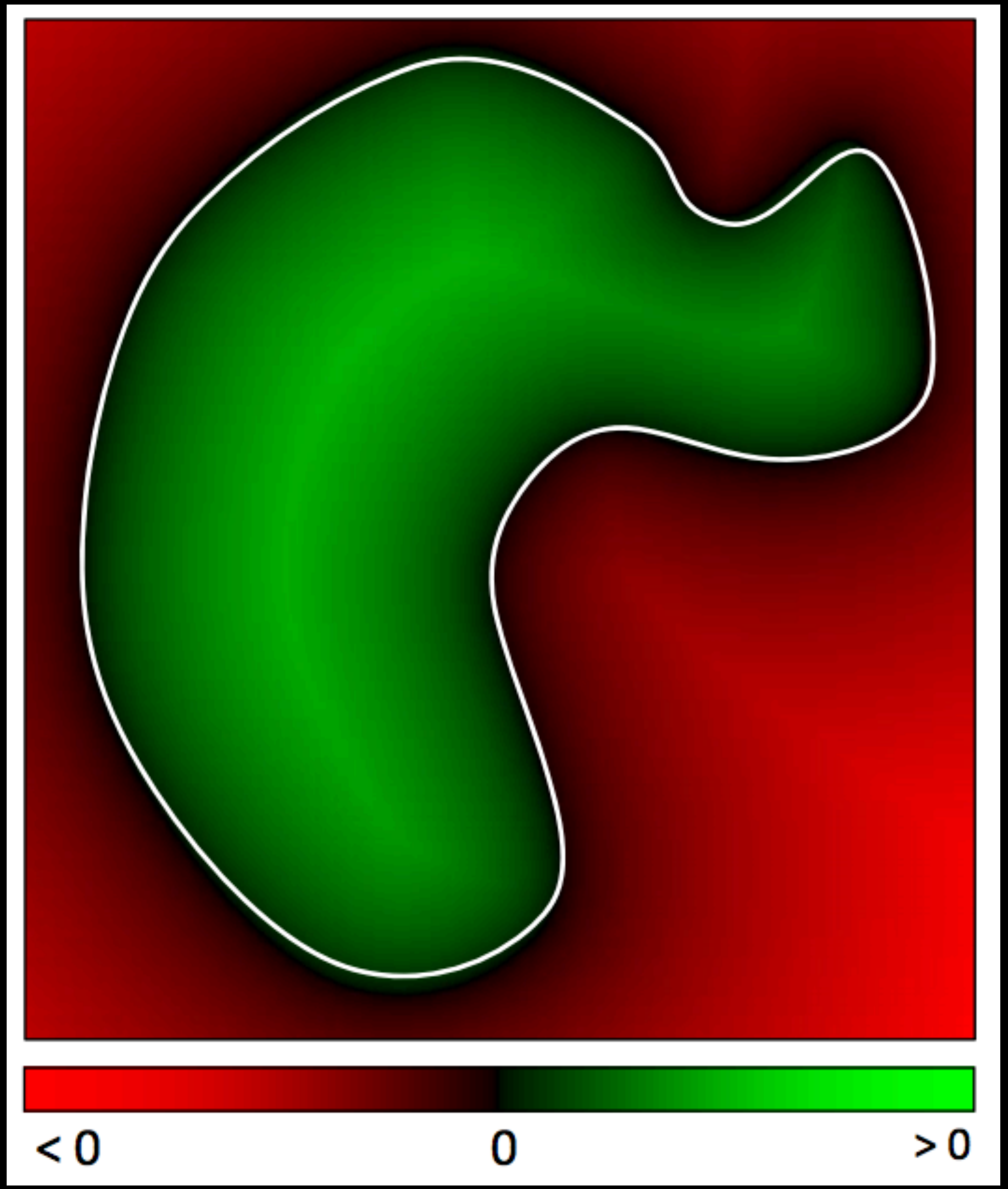


RADIAL BASIS FUNCTION INTERPOLATION



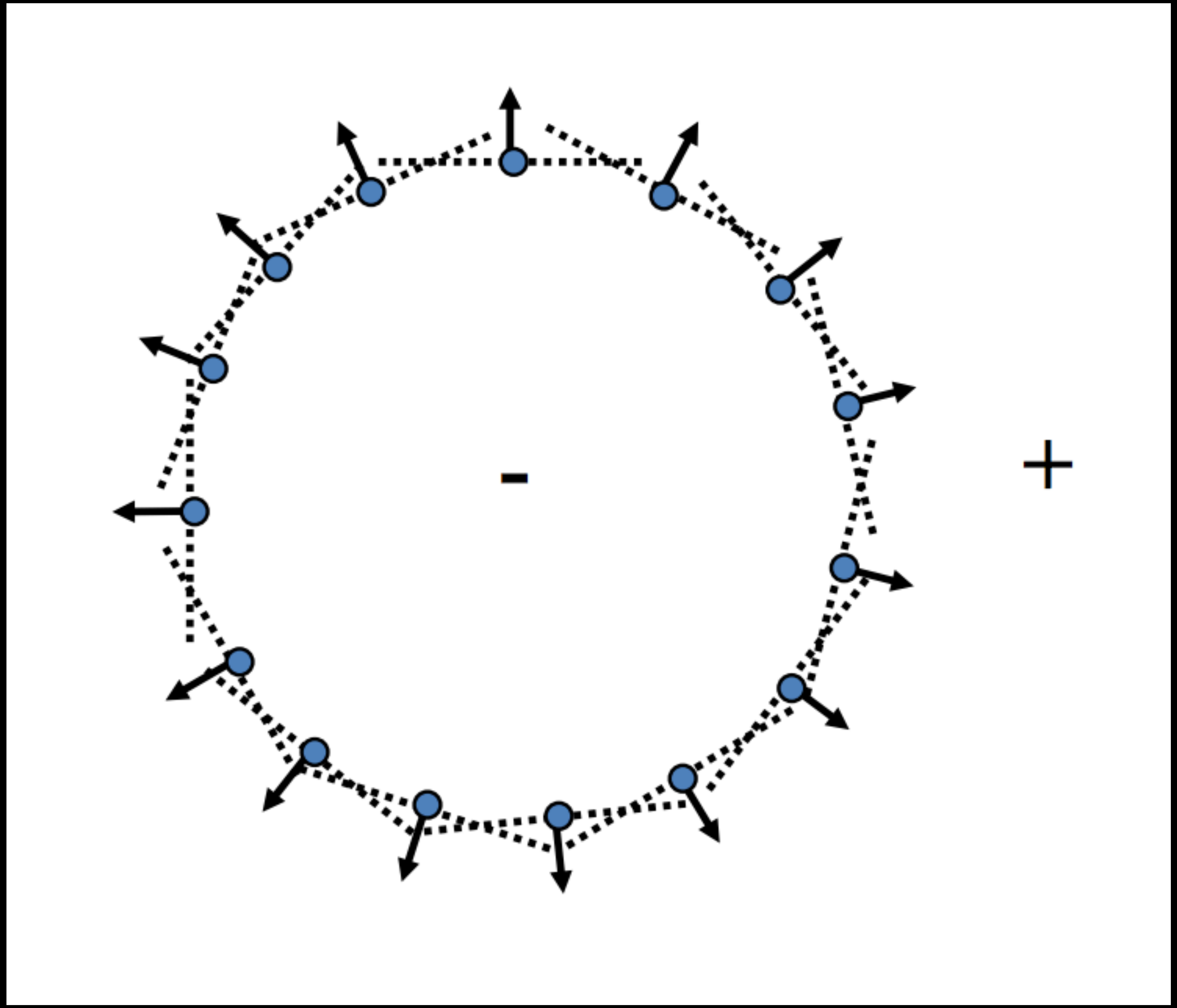
IMPLICIT RECONSTRUCTION

- Define a function
  - $f : R^3 \rightarrow R$
- with value < 0 outside the shape and > 0 inside
- Extract the 0 set
  - $\{ x : f(x) = 0 \}$



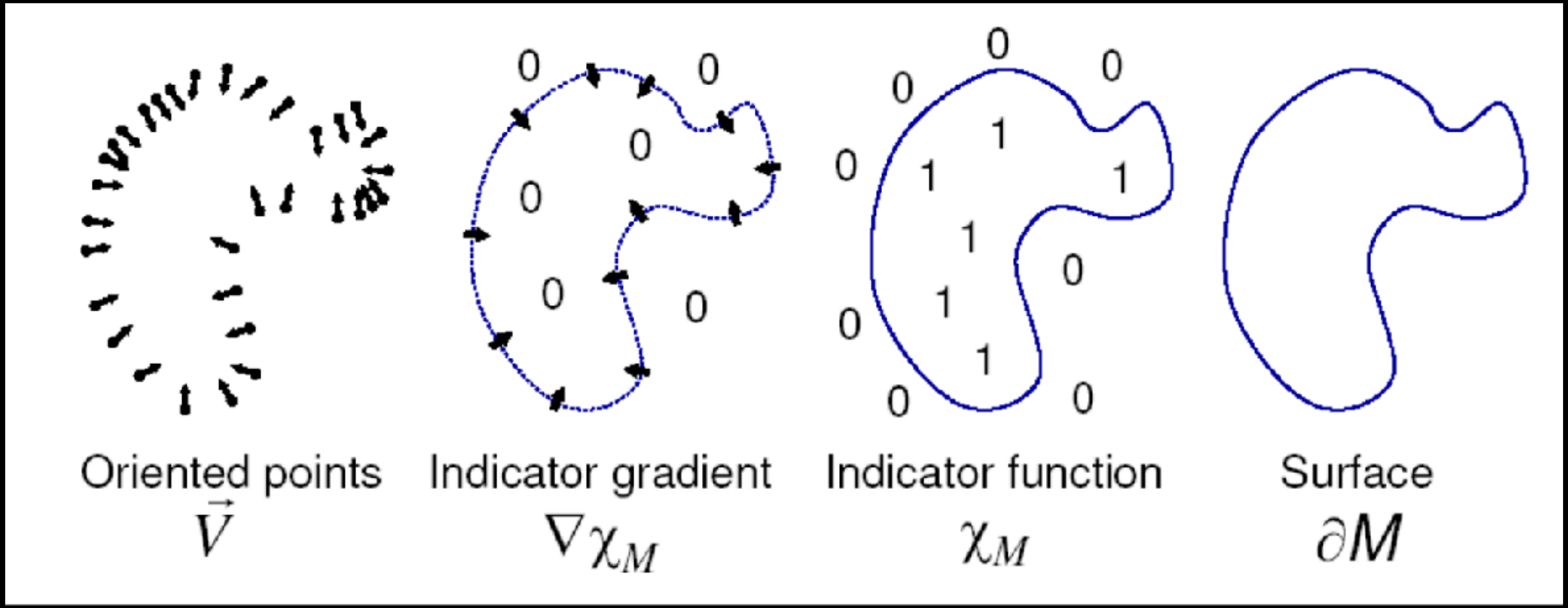
SDF FROM POINTS AND NORMALS

- Compute signed distance to the tangent plane of the closest point.
- Normals help to distinguish between inside and outside.

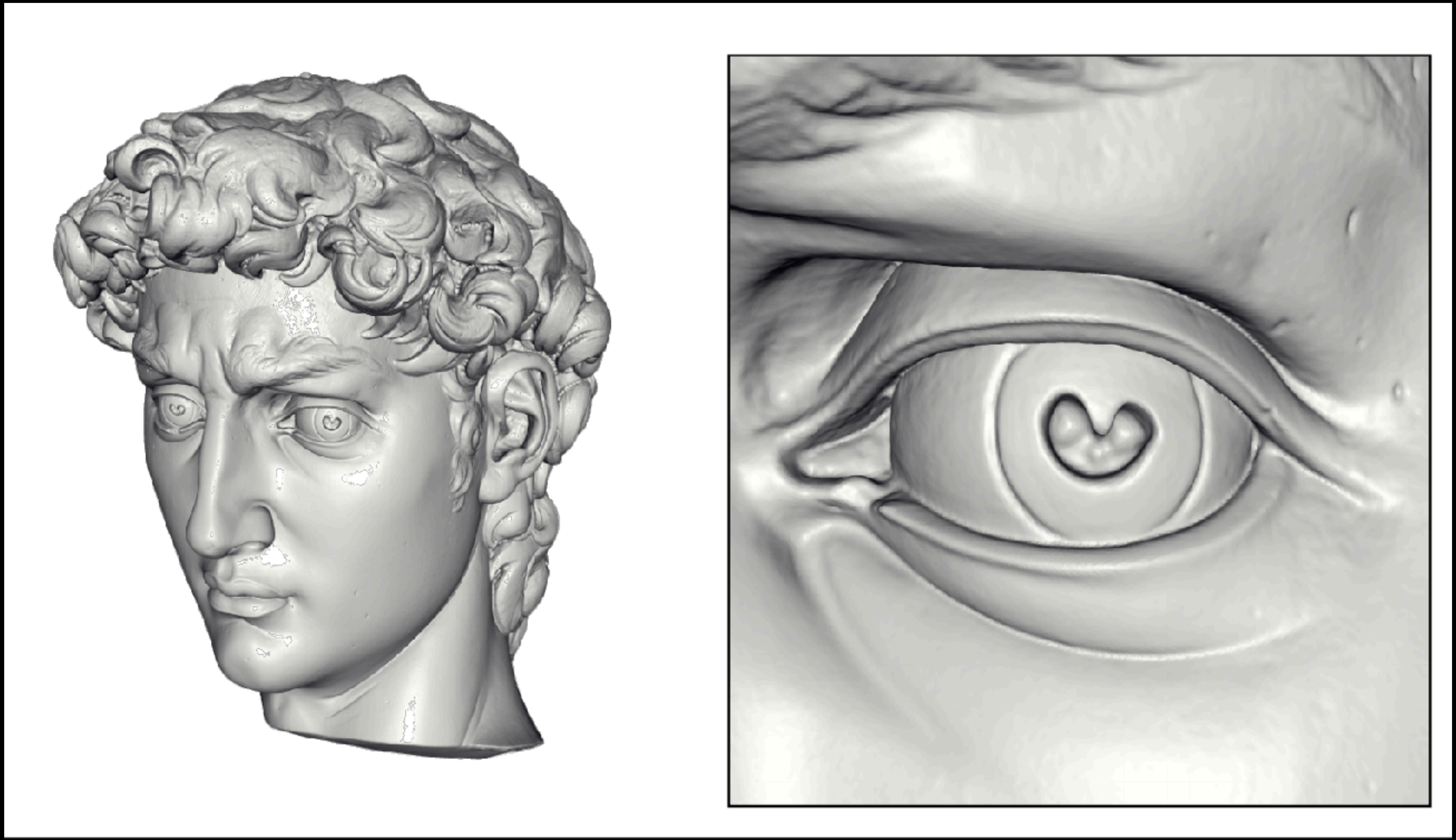


PROBLEM MODELING

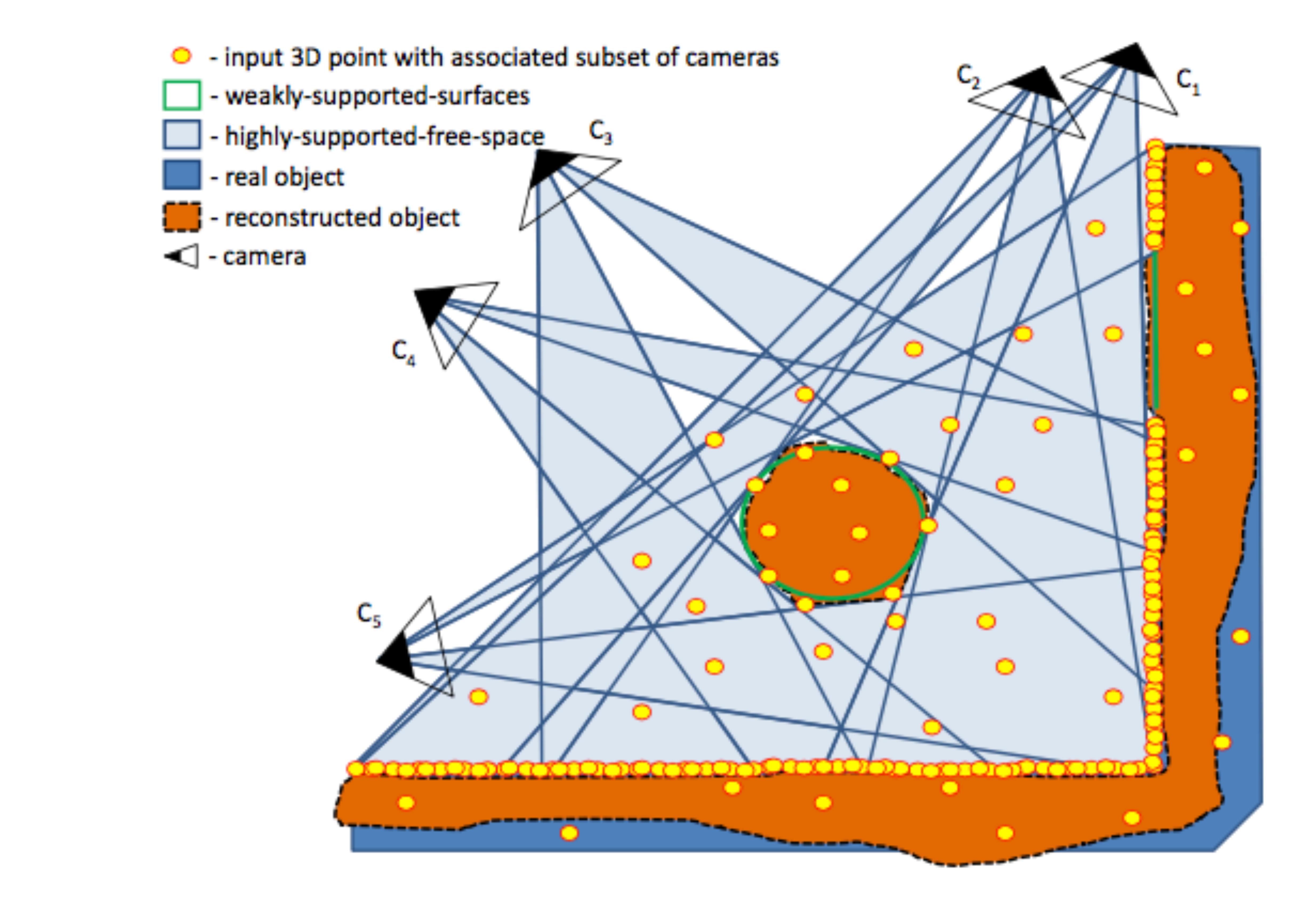
- Input: points with associated normals
- • Indicator function: 1 inside, 0 outside
- ➢ => gradient = 0 everywhere except near surface

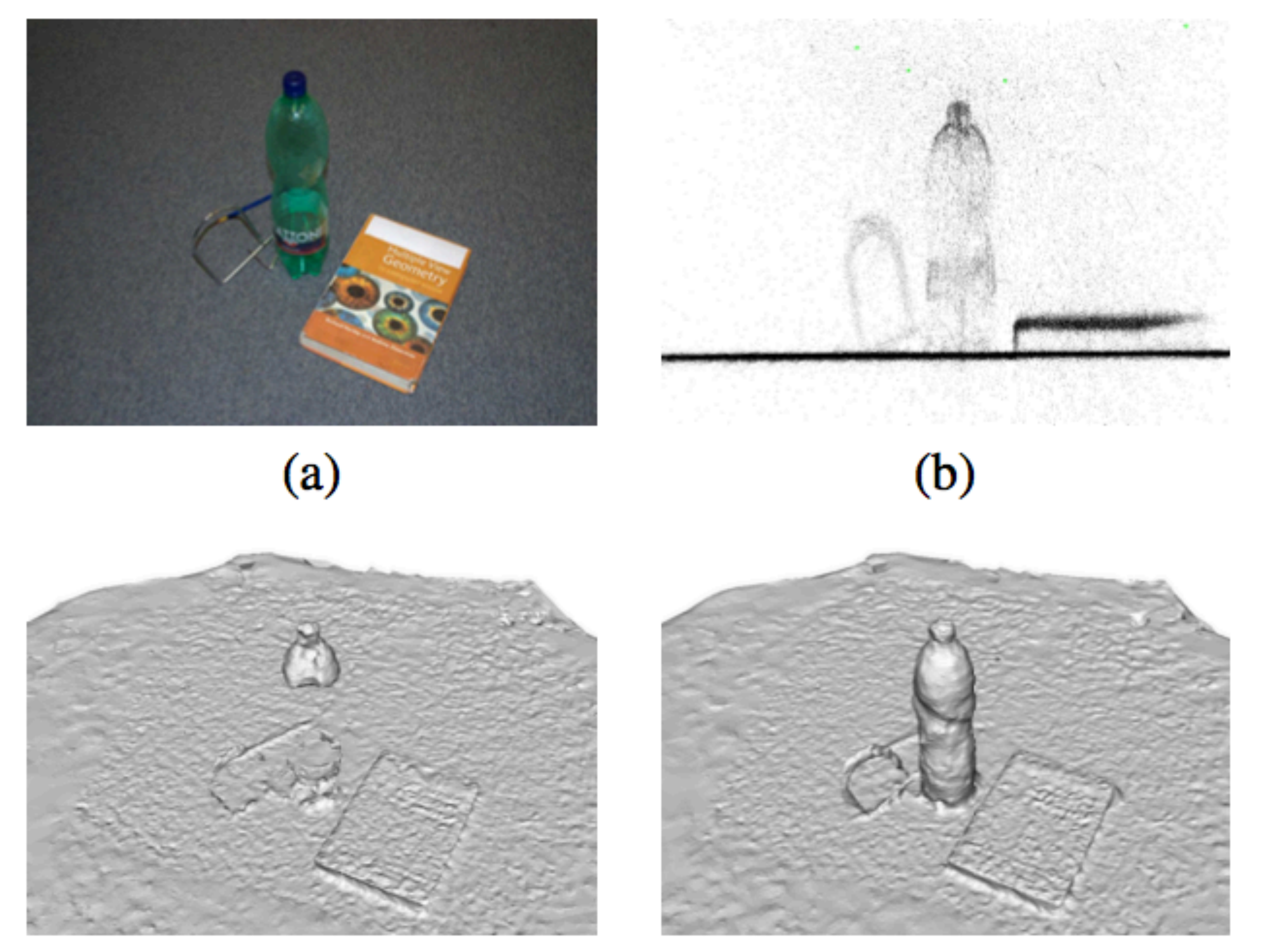


MICHELANGELO'S DAVID – EYE



MULTI-VIEW RECONSTRUCTION PRESERVING WEAKLY-SUPPORTED SURFACES

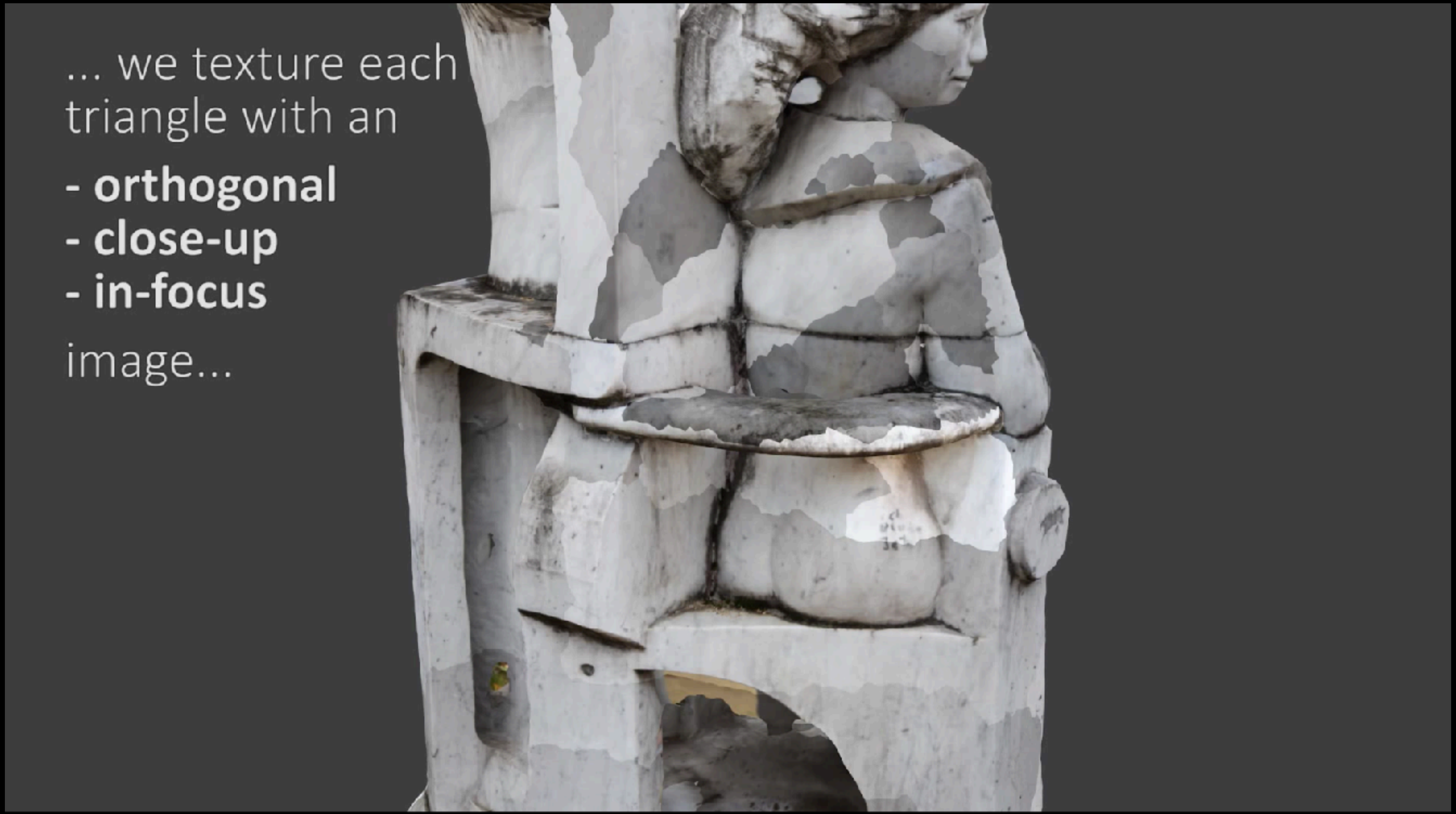




TEXTURING

MULTI-VIEW TEXTURING

VIEW SELECTION



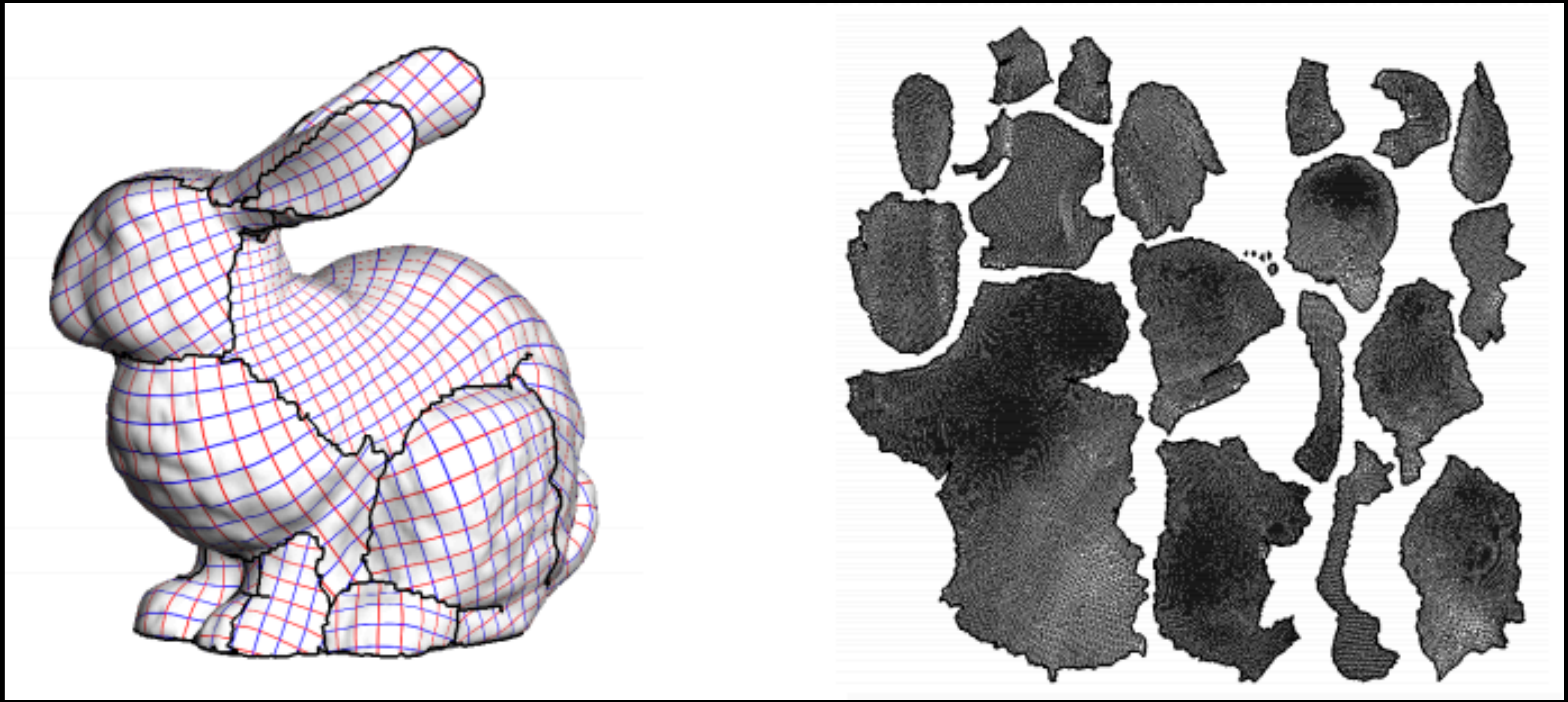
TYPICAL APPROACH

- Pretty much the same as with panorama stitching.
- Select view
  - Angle, distance, … (best view)
  - Blending (weghts depend on quality) not necessary
- Reduce seams
  - Color correction

TEXTURE ATLAS GENERATION

- Segmentation
- Parameterization
- Packing

PACKING CHARTS



http://ir-ltd.net/cracking-the-black-box-of-photogrammetry-scanning/

SOFTWARE

- VSFM + Meshlab
- VSFM + CMPMVS
  - http://ptak.felk.cvut.cz/sfmservice/websfm.pl?menu=cmpmvs
- Capturing Reality
- Agisoft Photoscan

EVALUATION OF THE COURSE

COMPUTER VISION AND PHOTOGRAMMETRY

- Image Formation and Camera Calibration
- Feature detection and matching
- Stereo and Structure from Motion
- Registration, Meshing and Texturing
- Shape from Shading and Photometric Stereo

KNOWLEDGE

- Theoretical background regarding 3d reconstruction
- Hands-on knowledge about 3d reconstruction
- Understanding the problems involved in 3d reconstruction
- and an overview of possible solutions

SKILLS

- Experience implementing image processing algorithms
- Experience with algebra and image processing libraries

COMPUTER VISION AND PHOTOGRAMMETRY

- Mathematical Tools
  - DLS - Solving with SVN
  - RANSAC
  - Bundle Adjustment
- Understanding about images and descriptors
- Geometry between Cameras