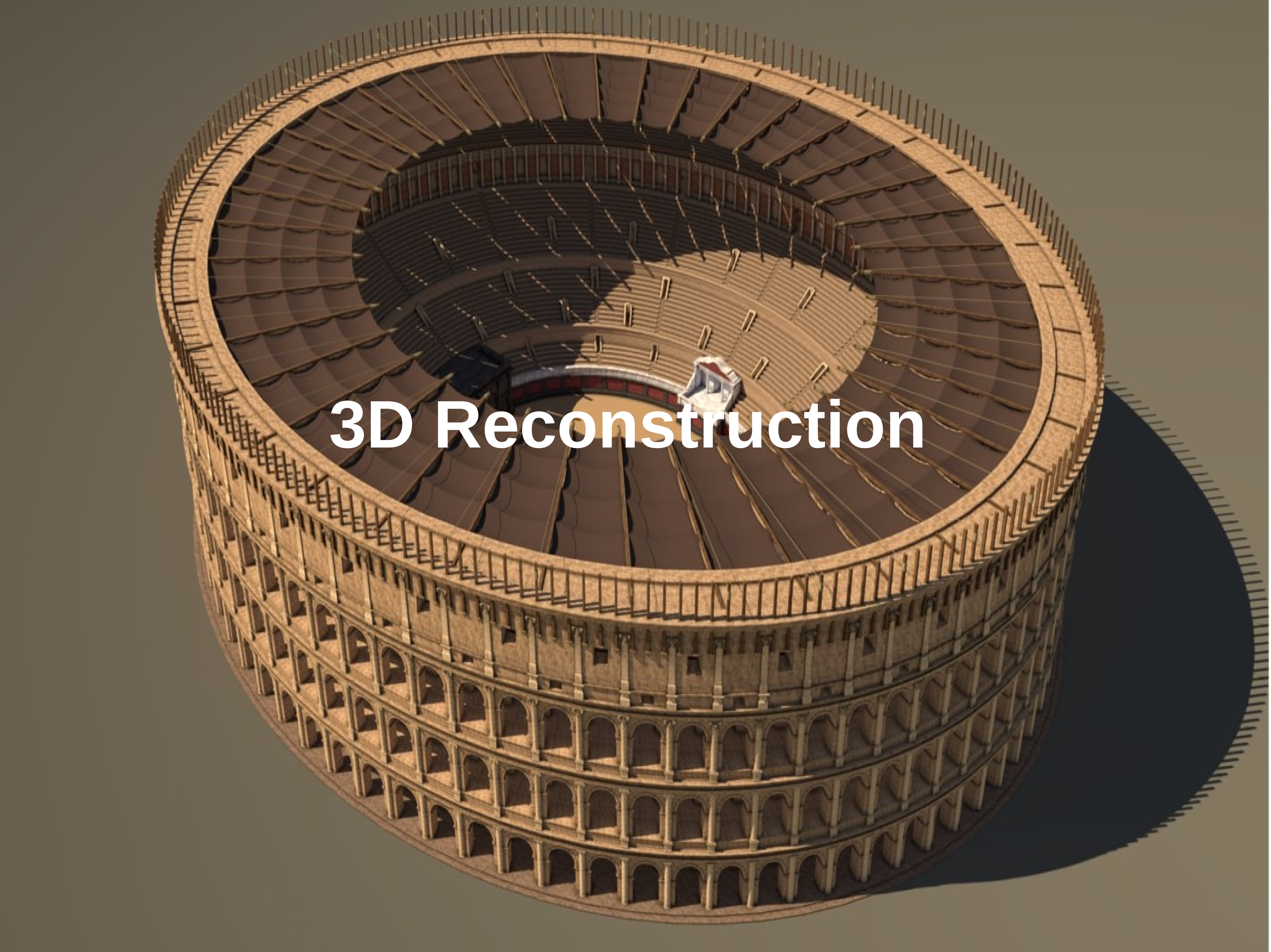

3D Reconstruction

Contents

Question

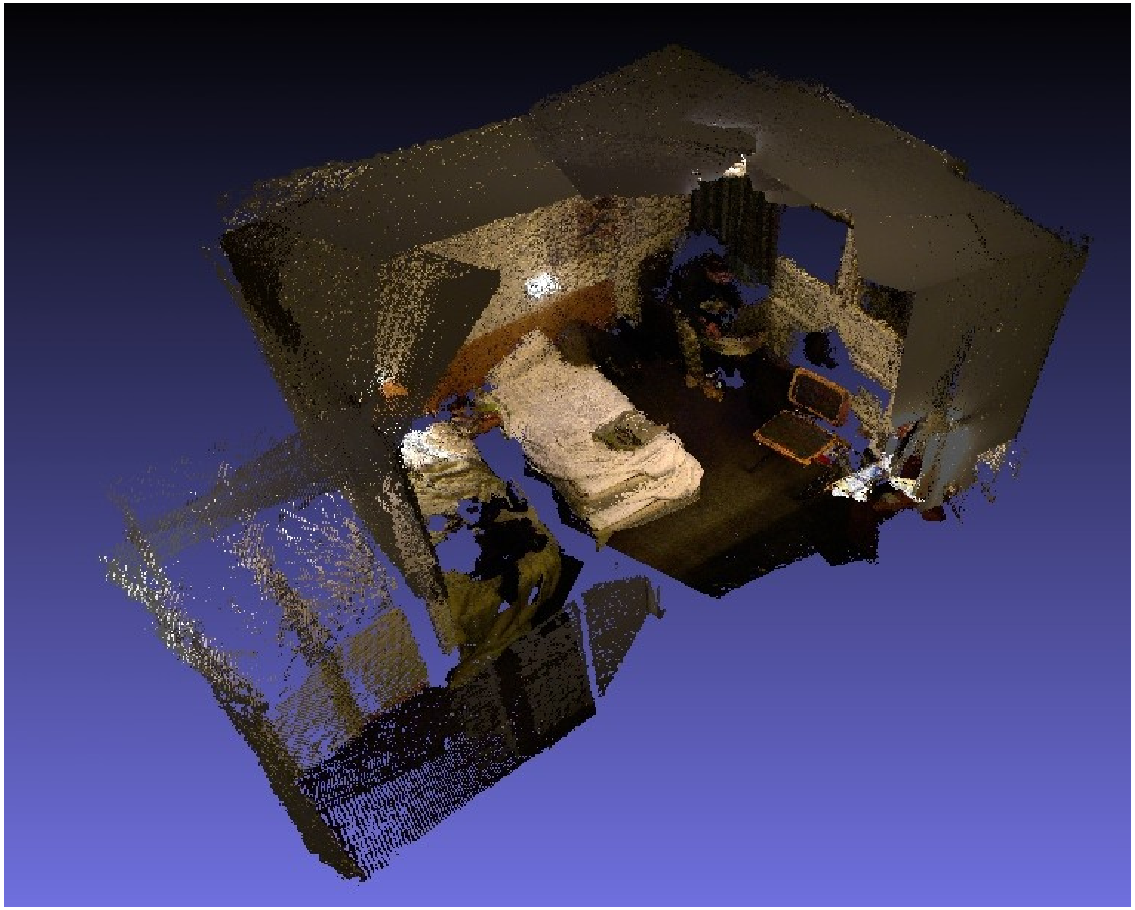
What is the goal of computer vision?

VISION

2D to 3D

3D Reconstruction in Our Life

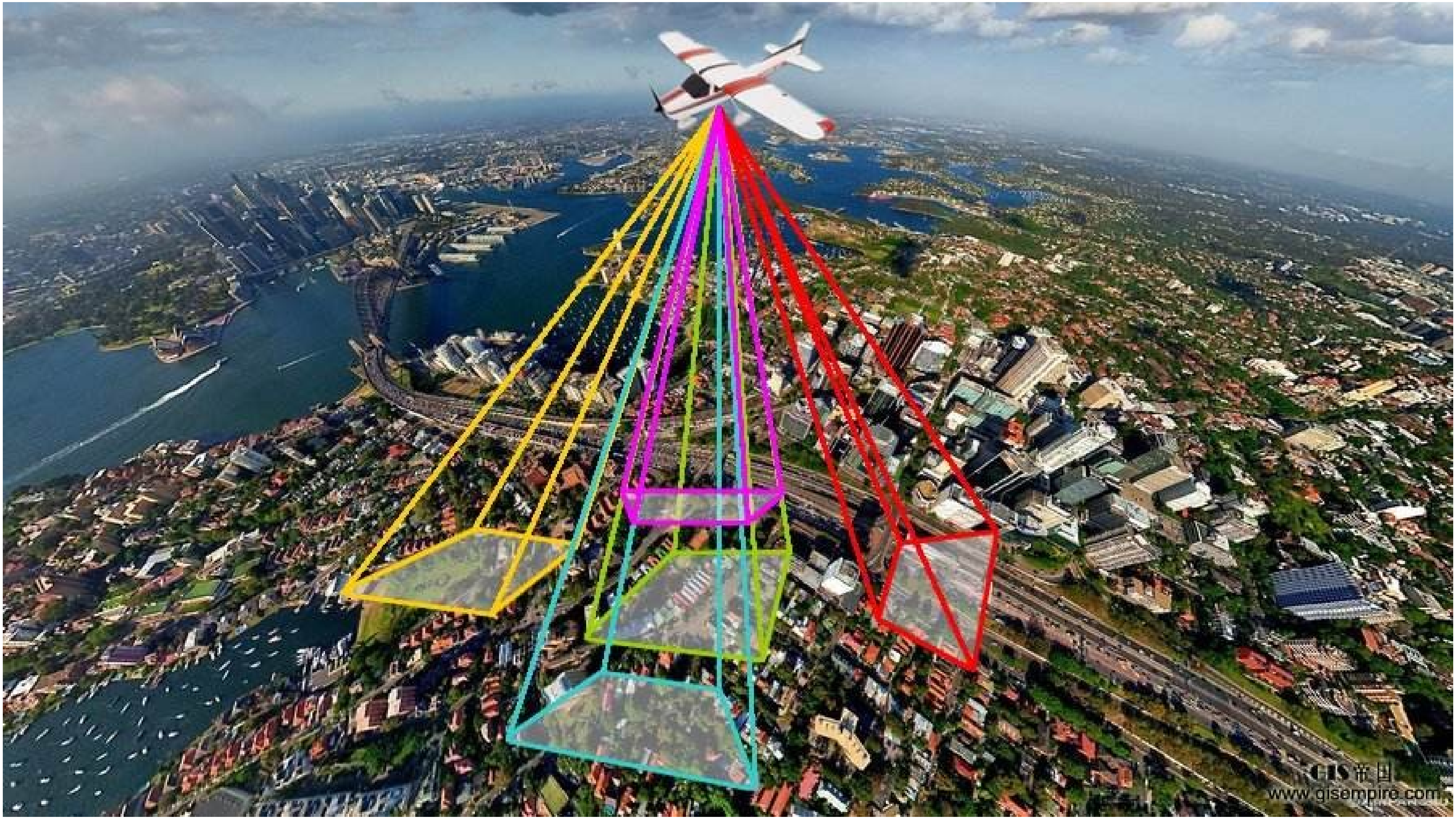
SLAM(Simultaneous localization and mapping)



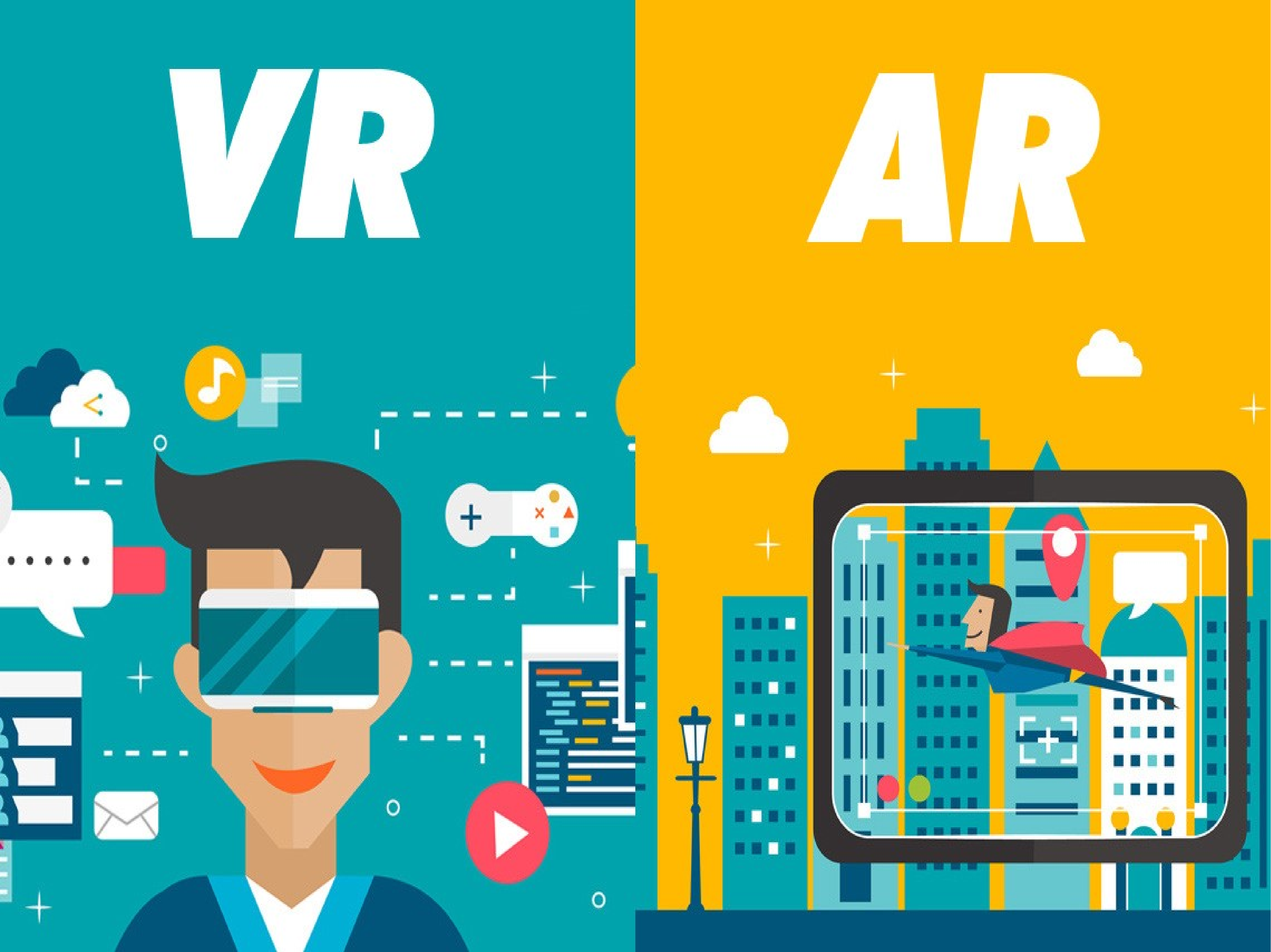
Video

Autonomous Vehicles


ZINGTO


www.gisempire.com

Video


VR
AR

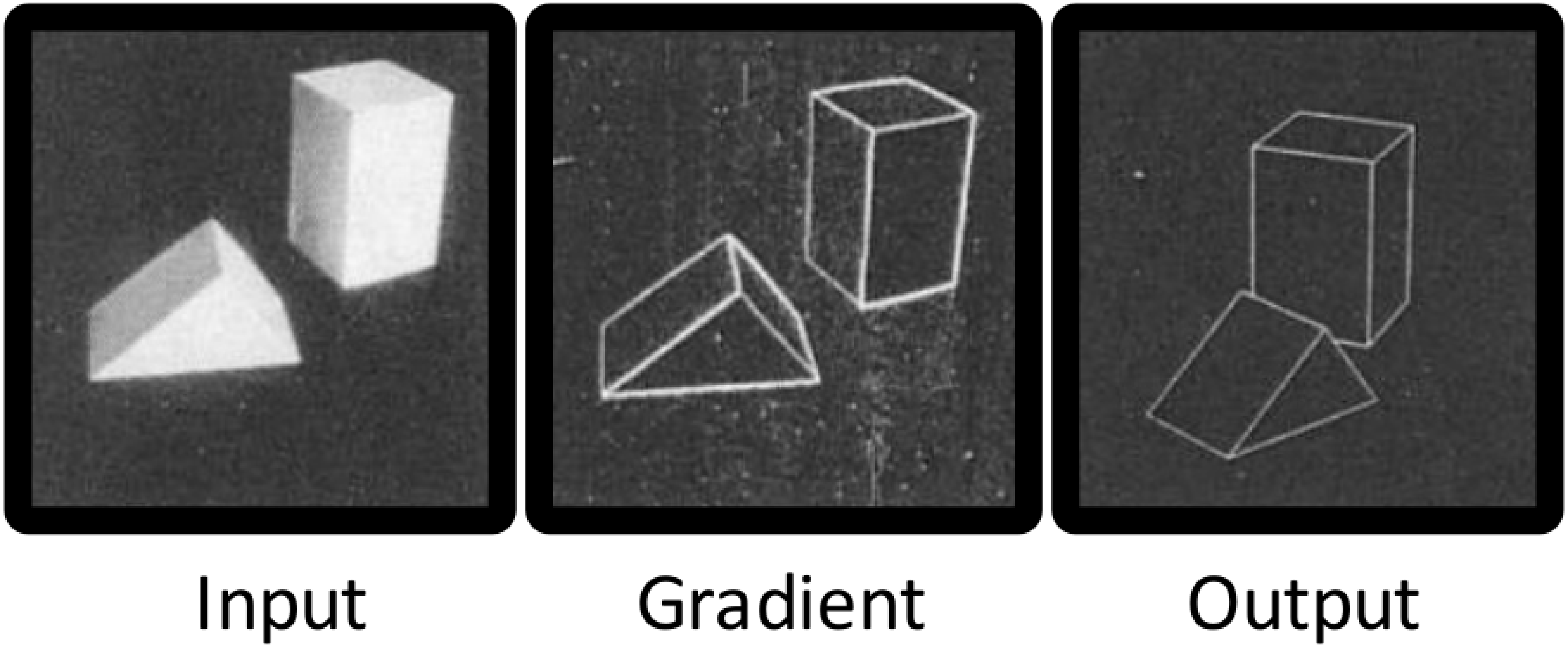
Virtual Reality


move and two fingers to rotate

History of 3D Reconstruction

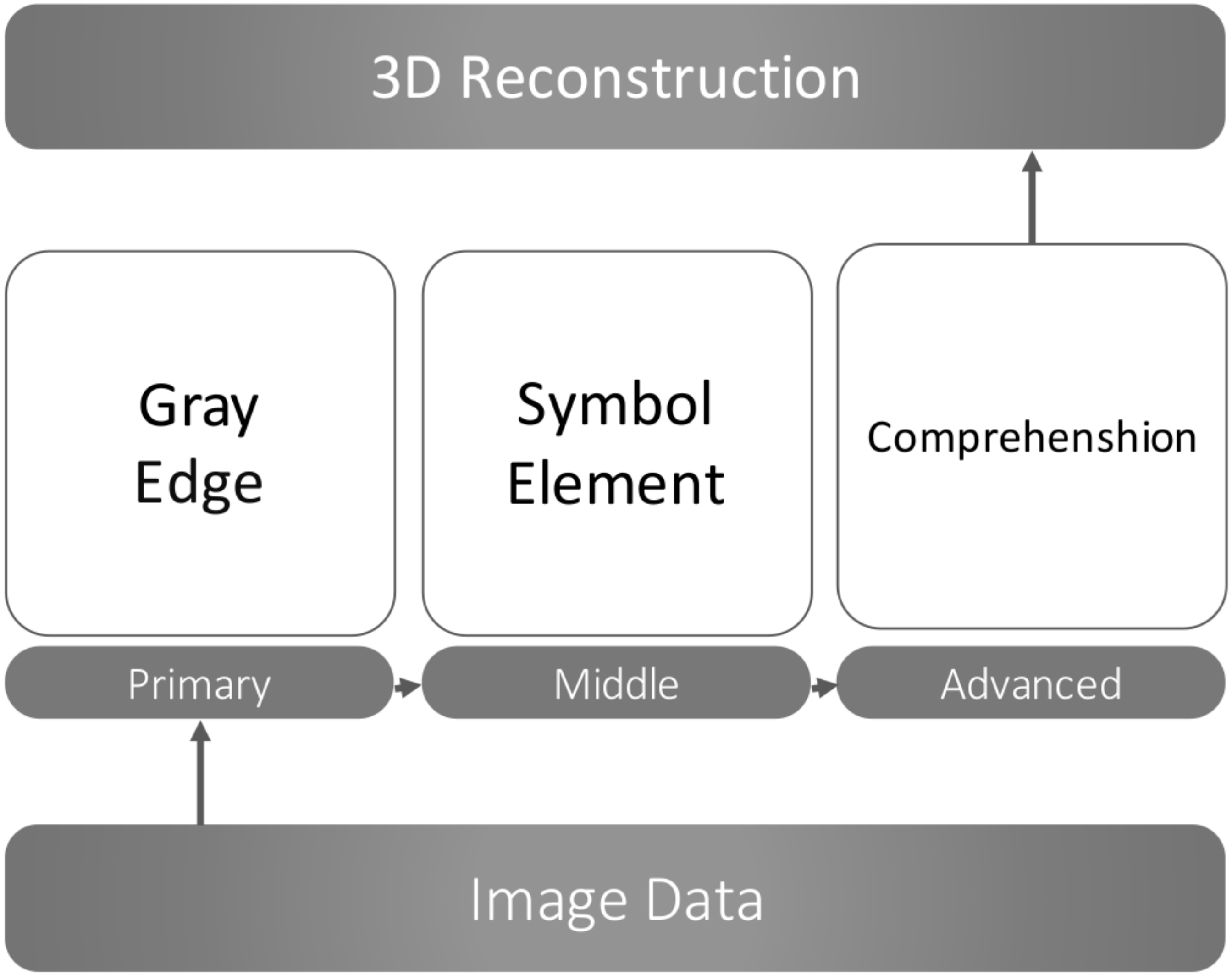
- 1960s Problem Definition
- 1970s Image Formulation
- 1990s Geometry
- 2000s Reconstruction
- 2010s Innovation

1960s   Problem Definition



Roberts L G. Machine perception of three-dimensional soups. Massachusetts Institute of Technology, 1963.

1970s Image Formulation



D. Marr: Computational Theory of Vision

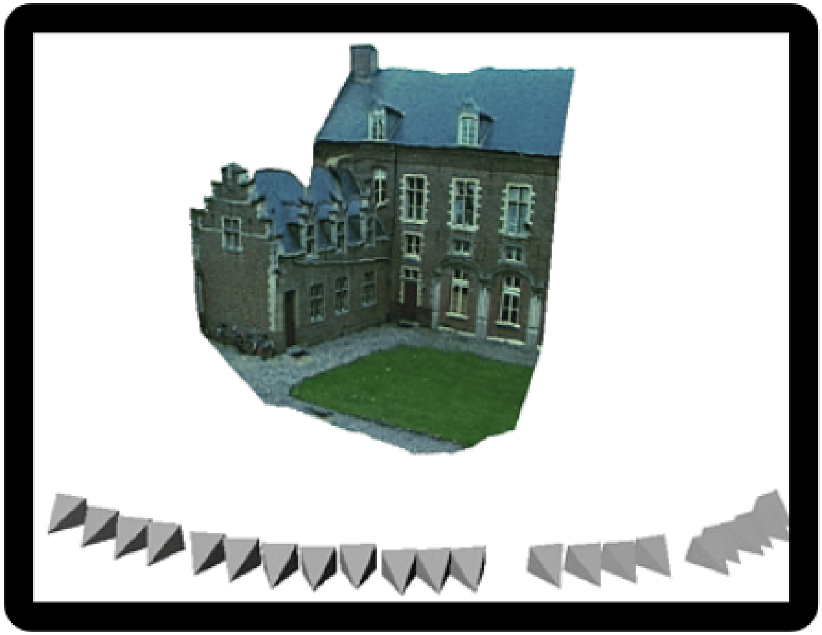
1990s   Geometry

8 Point Algorithm

Longuet-Higgins H C. A computer algorithm for reconstructing a scene from two projections. Readings in Computer Vision: Issues, Problems, Principles, and Paradigms, MA Fischler and O. Firschein, 1987.

2000s Reconstruction

Structure from Motion



Pollefeys et al

Multi-view Stereo

Furukawa & Ponce

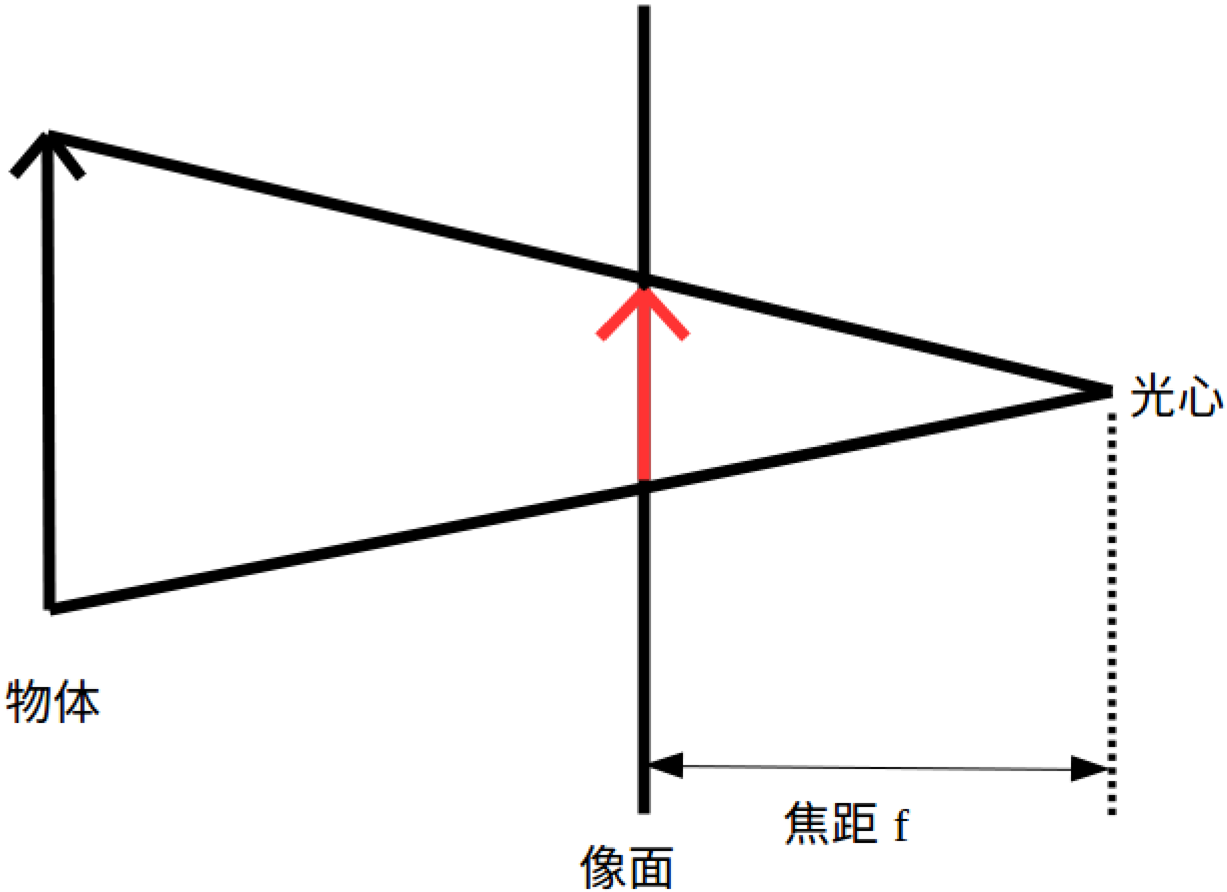
2010s Innovation

- Big Data
- Real Time
- Deep Learning

Classification

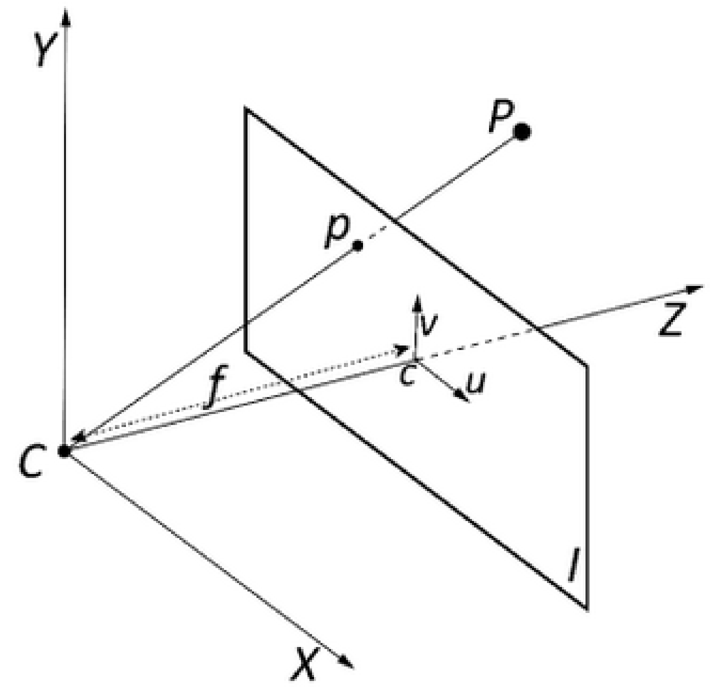
- Passive methods
  - Recover from shadow
  - Recover from stereoscopic
  - Structure from motion
  - ……
- Active methods
  - Structure light
  - ……

Camera Model

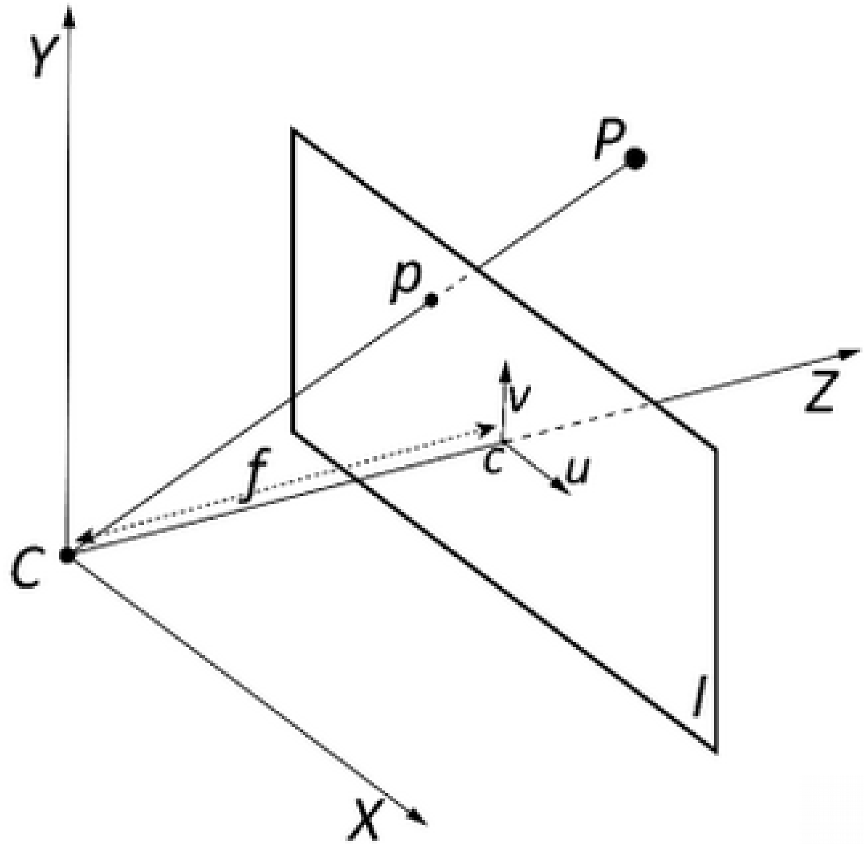


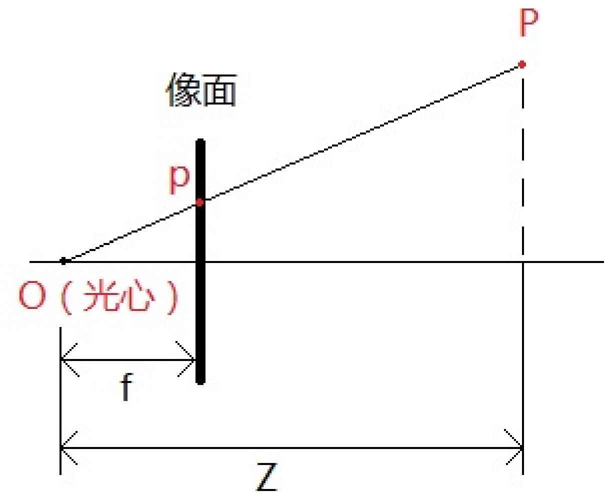
Goal: Establish a map: P to p

Goal: Establish a map: P to p



$$P = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \longrightarrow p = \begin{bmatrix} x \\ y \end{bmatrix}$$





$$x = \frac{fX}{Z}, y = \frac{fY}{Z}$$

$$x = \frac{fX}{Z} + c_x, y = \frac{fY}{Z} + c_y$$

E→H

$$(x, y) \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

homogeneous image coordinates

$$(x, y, z) \Rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

homogeneous scene coordinates

H→E

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w)$$

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \Rightarrow (x/w, y/w, z/w)$$

homogeneous coordinates

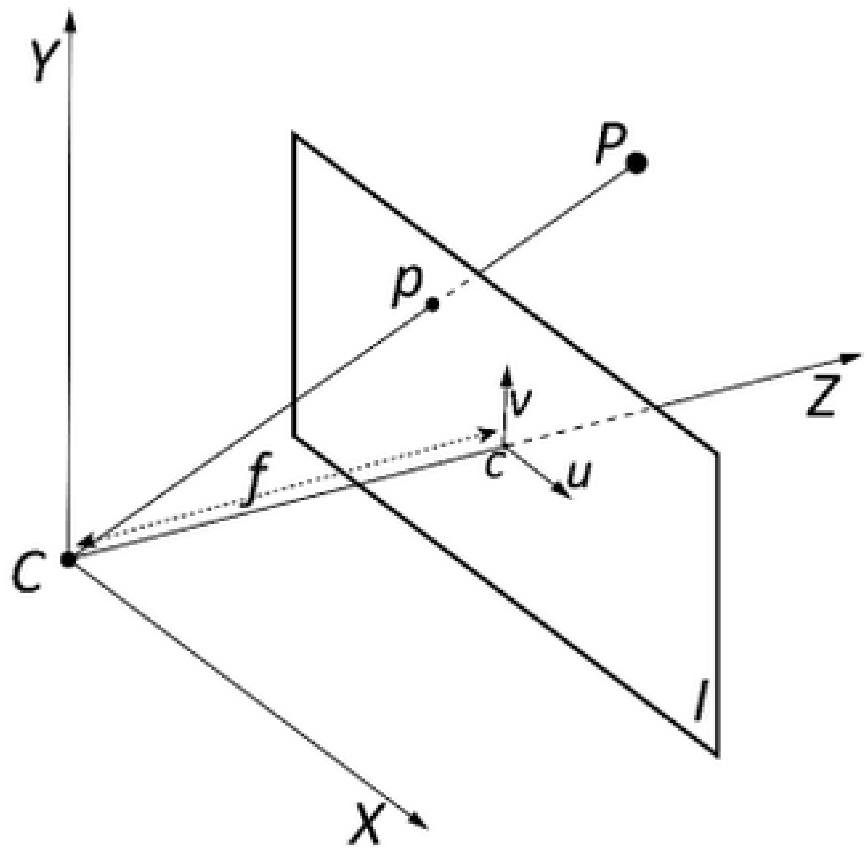
$$P = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \longrightarrow p = \begin{bmatrix} \frac{fX}{Z} + c_x \\ \frac{fY}{Z} + c_y \\ 1 \end{bmatrix}$$

$$P = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \longrightarrow p = \begin{bmatrix} fX + Zc_x \\ fY + Zc_y \\ Z \end{bmatrix}$$

$$\begin{bmatrix} fX + Zc_x \\ fY + Zc_y \\ Z \end{bmatrix} = \begin{bmatrix} f & 0 & c_x & 0 \\ 0 & f & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$p = MP$$

Parameter k,l: something like pixels/cm



$$x = \frac{fX}{Z}, y = \frac{fY}{Z}$$

$$x = \frac{fkX}{Z}, y = \frac{flY}{Z}$$

$$x = \frac{\alpha X}{Z}, y = \frac{\beta Y}{Z}$$

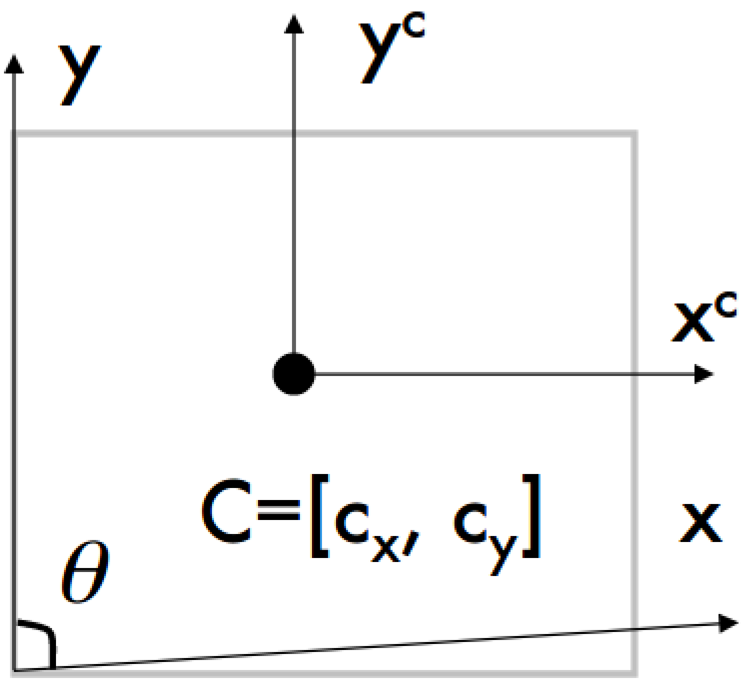
$$x = \frac{\alpha X}{Z} + c_x, y = \frac{\beta Y}{Z} + c_y$$

$$\begin{bmatrix} \alpha X + Z c_x \\ \beta Y + Z c_y \\ Z \end{bmatrix} = \begin{bmatrix} \alpha & 0 & c_x & 0 \\ 0 & \beta & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

p = MP

$$M = \begin{bmatrix} \alpha & 0 & c_x & 0 \\ 0 & \beta & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Camera Skewness



$$M = \begin{bmatrix} \alpha & -\alpha cot\theta & c_x & 0 \\ 0 & \frac{\beta}{sin\theta} & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$p = \begin{bmatrix} \alpha & -\alpha cot\theta & c_x & 0 \\ 0 & \frac{\beta}{sin\theta} & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} P$$

$$p = \begin{bmatrix} \alpha & -\alpha cot\theta & c_x \\ 0 & \frac{\beta}{sin\theta} & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} P$$

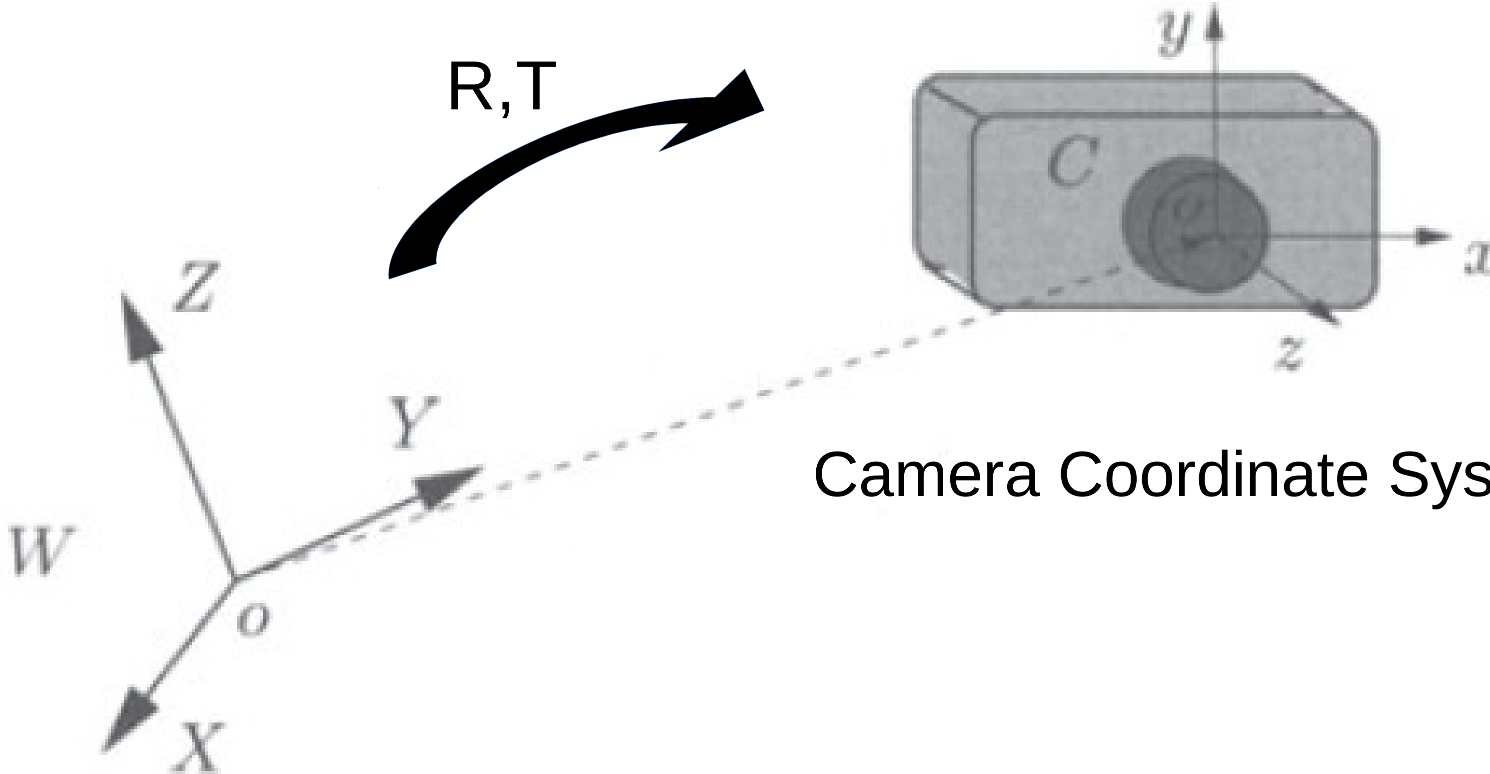
Camera Matrix

External Matrix

$$p = K[I|0]P$$

K have 5 degrees of freedom

Camera Coordinate System and World Coordinate System



Camera Coordinate System

World Reference System

3D Translation and Rotation

$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix}$$

$$R_y(\beta) = \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix}$$

$$R_z(\gamma) = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R = R_x(\alpha)\ R_y(\beta)\ R_z(\gamma)$$

R: 3 degrees of freedom

$$T = \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix}$$

T: 3 degrees of freedom

World Coordinate System to Camera Coordinate System

In homogeneous coordinates:

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \rightarrow P = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}_{4\times 4} P_w \leftarrow \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

Internal parameters

External parameters

$$p = K\begin{bmatrix} I & 0 \end{bmatrix} P = K\begin{bmatrix} I & 0 \end{bmatrix}\begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}_{4\times 4} P_w = \underbrace{\boxed{K\begin{bmatrix} R & T \end{bmatrix}}}_{M} P_w$$

M: 5+3+3=11 degrees of freedom

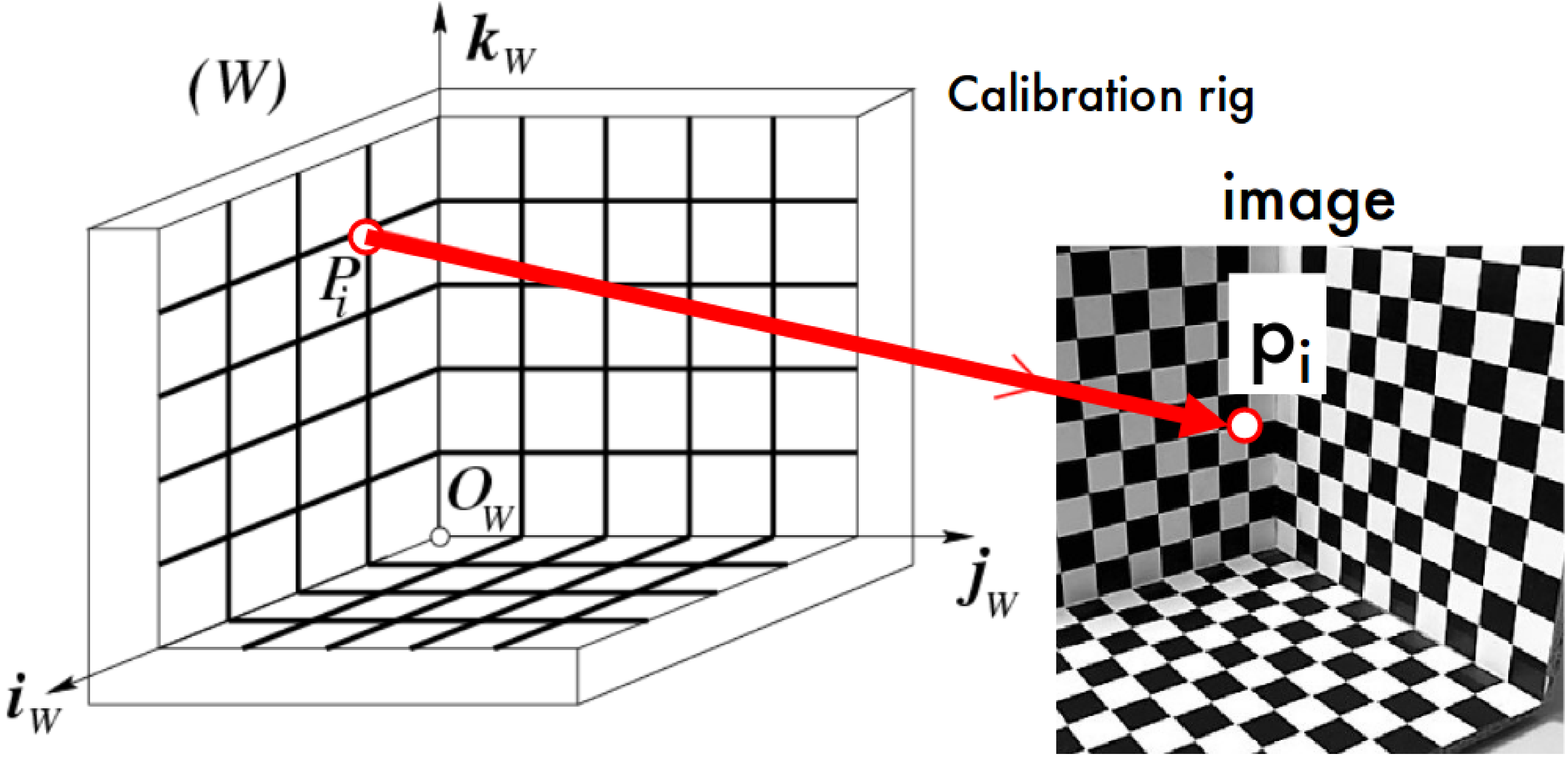
$$p_{3\times1} = M\ P_w \ = K_{3\times3}\begin{bmatrix} R & T \end{bmatrix}_{3\times4} P_{w4\times1} \qquad M = \begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \\ \mathbf{m}_3 \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \\ \mathbf{m}_3 \end{bmatrix} P_W = \begin{bmatrix} \mathbf{m}_1 P_W \\ \mathbf{m}_2 P_W \\ \mathbf{m}_3 P_W \end{bmatrix} \xrightarrow{\mathbf{E}} \left(\frac{\mathbf{m}_1 P_w}{\mathbf{m}_3 P_w}, \frac{\mathbf{m}_2 P_w}{\mathbf{m}_3 P_w}\right)$$

Camera Calibration

$$P' = M\,P_w = \underbrace{K}_{\text{Internal parameters}}\underbrace{\begin{bmatrix} R & T \end{bmatrix}}_{\text{External parameters}} P_w$$

- Goal: Estimate intrinsic and extrinsic parameters from 1 or multiple images
- Why use it: If given an arbitrary camera, we may or may not have access to these parameters



- $p_i = MP_i$
- p_i, P_i are known
- M have 11 degrees of freedom, we need 11 equations

How many correspondences do we need?

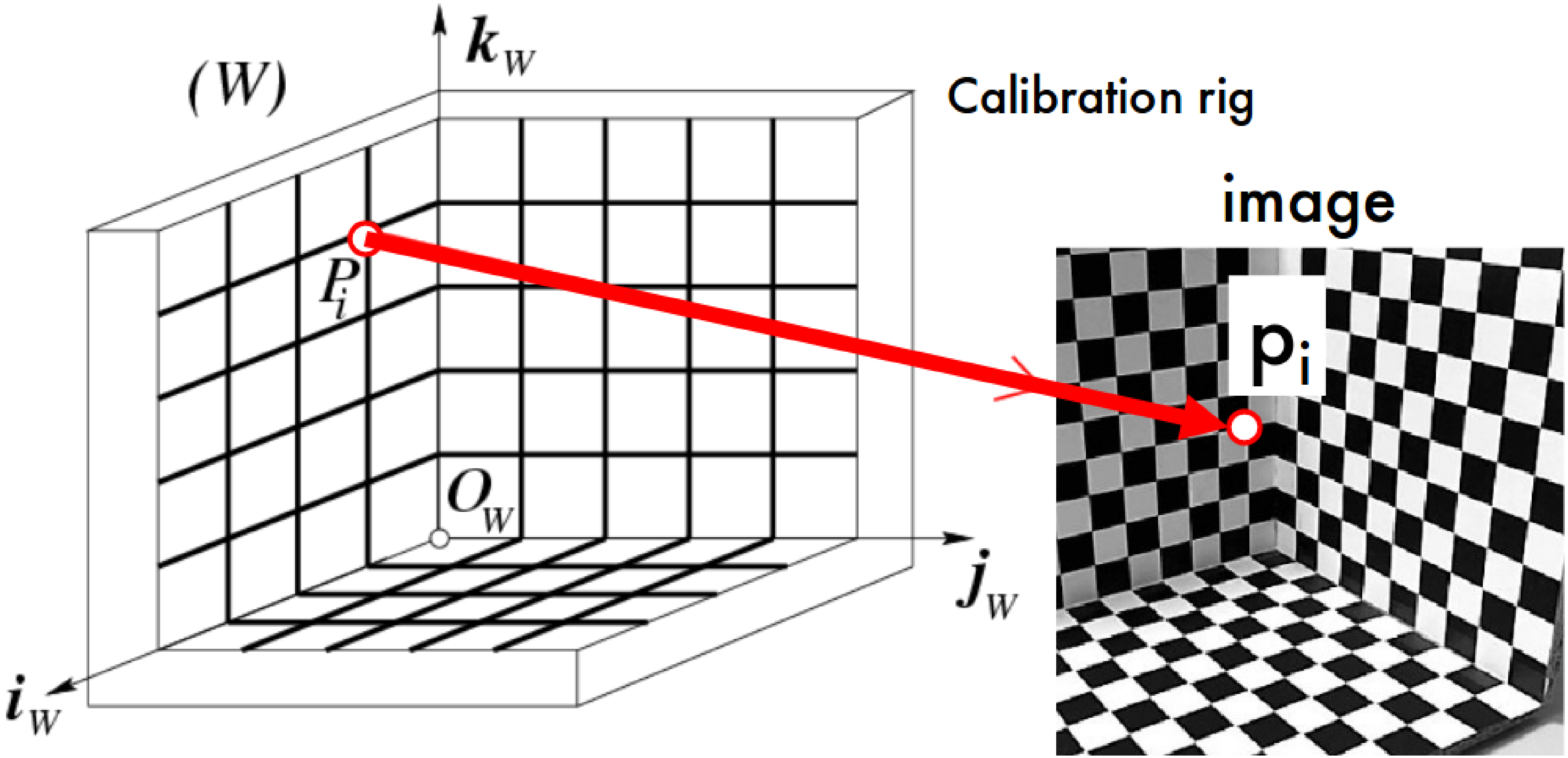
$$p_i = \begin{bmatrix} u_i \\ v_i \end{bmatrix} = \begin{bmatrix} \dfrac{\mathbf{m}_1 \, \mathrm{P_i}}{\mathbf{m}_3 \, \mathrm{P_i}} \\ \dfrac{\mathbf{m}_2 \, \mathrm{P_i}}{\mathbf{m}_3 \, \mathrm{P_i}} \end{bmatrix} = M \, P_i \qquad \mathrm{M} = \begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \\ \mathbf{m}_3 \end{bmatrix}$$

in pixels

$$\mathrm{u}_i(\mathbf{m}_3 \, \mathrm{P}_i) - \mathbf{m}_1 \, \mathrm{P}_i = 0$$

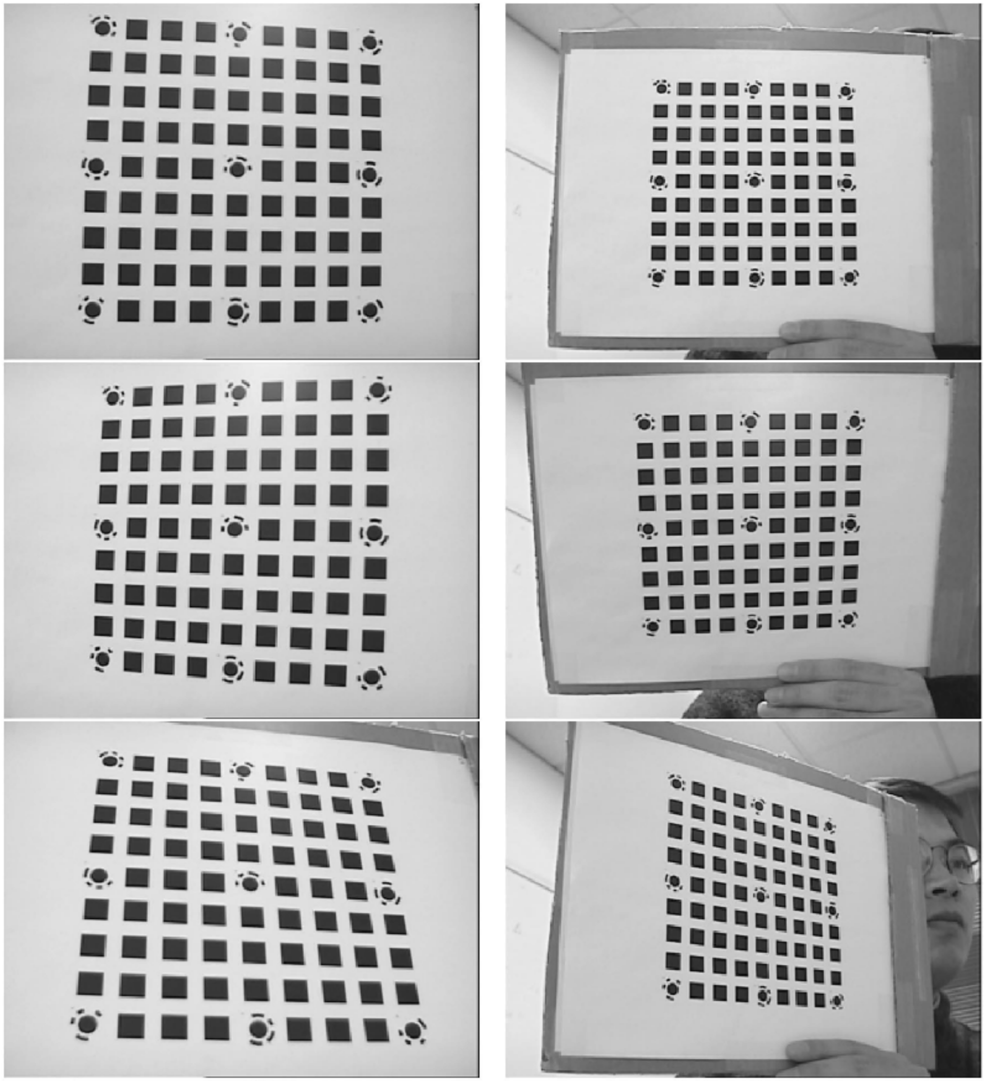
$$\mathrm{v}_i(\mathbf{m}_3 \, \mathrm{P}_i) - \mathbf{m}_2 \, \mathrm{P}_i = 0$$

We need at least 6 correspondences



In practice, using more than 6 correspondences enables more robust results

Zhang Zhengyou calibration method



Zhang Z Y. A flexible new technique for camera calibration. IEEE Transactions on pattern analysis and machine intelligence, 2000.

Compute Internal Matrix

↓

Compute External Matrix

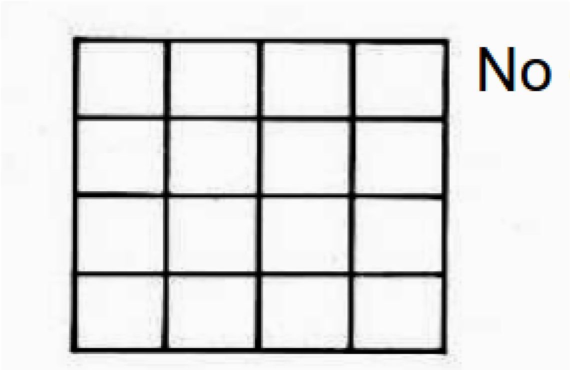
↓

Maximum-likelihood Criterion

↓

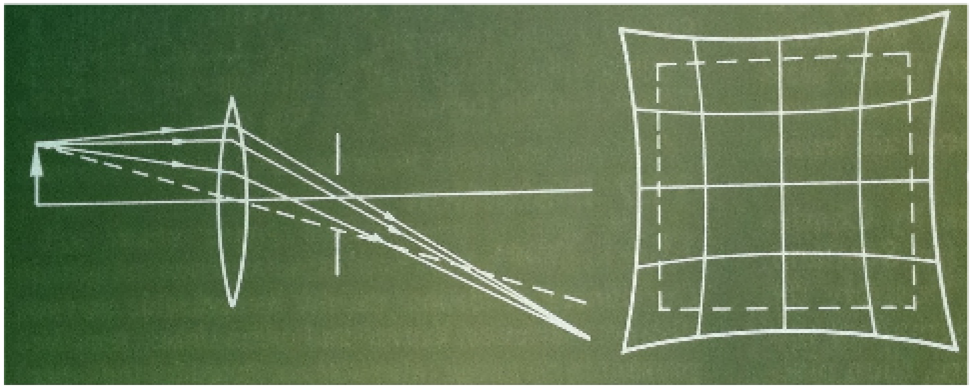
Radial Distortion

Zhang Z Y. A flexible new technique for camera calibration. IEEE Transactions on pattern analysis and machine intelligence, 2000.

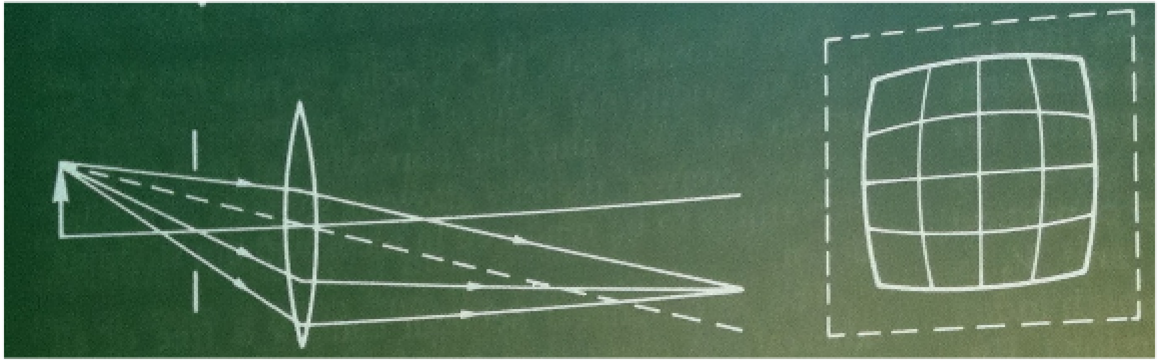


No distortion

Pin cushion



Barrel (fisheye lens)




Copyright (C) Yusuf Hashim

$$\hat{m}(K, R_i, t_i, M_{ij}) = K[R|t]M_{ij}$$

$$f(M_{ij}) = \frac{1}{\sqrt{2\pi}} e^{\frac{-(\hat{m}(K, R_i, t_i, M_i j) - m_{ij})^2}{\sigma^2}}$$

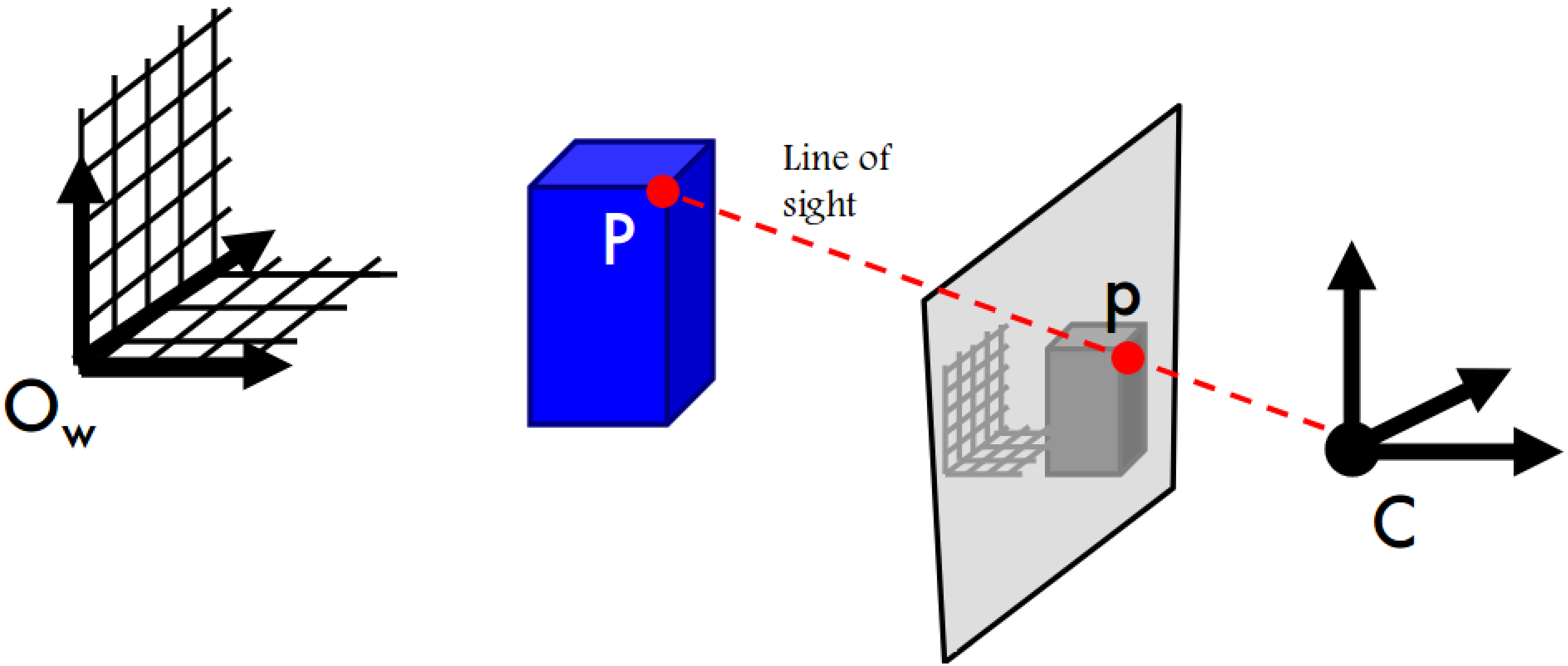
$$L(A, R_i, t_i, M_{ij}) = \prod_{i=1, j=1}^{n,m} f(M_{ij}) = \frac{1}{\sqrt{2\pi}} e^{\frac{-\sum_{i=1}^{n} \sum_{j=1}^{m} (\hat{m}(K, R_i, t_i, M_i j) - m_{ij})^2}{\sigma^2}}$$

$$\sum_{i=1}^{n} \sum_{j=1}^{m} \|\hat{m}(K, R_i, t_i, M_{ij}) - m_{ij}\|^2$$

Conclusion

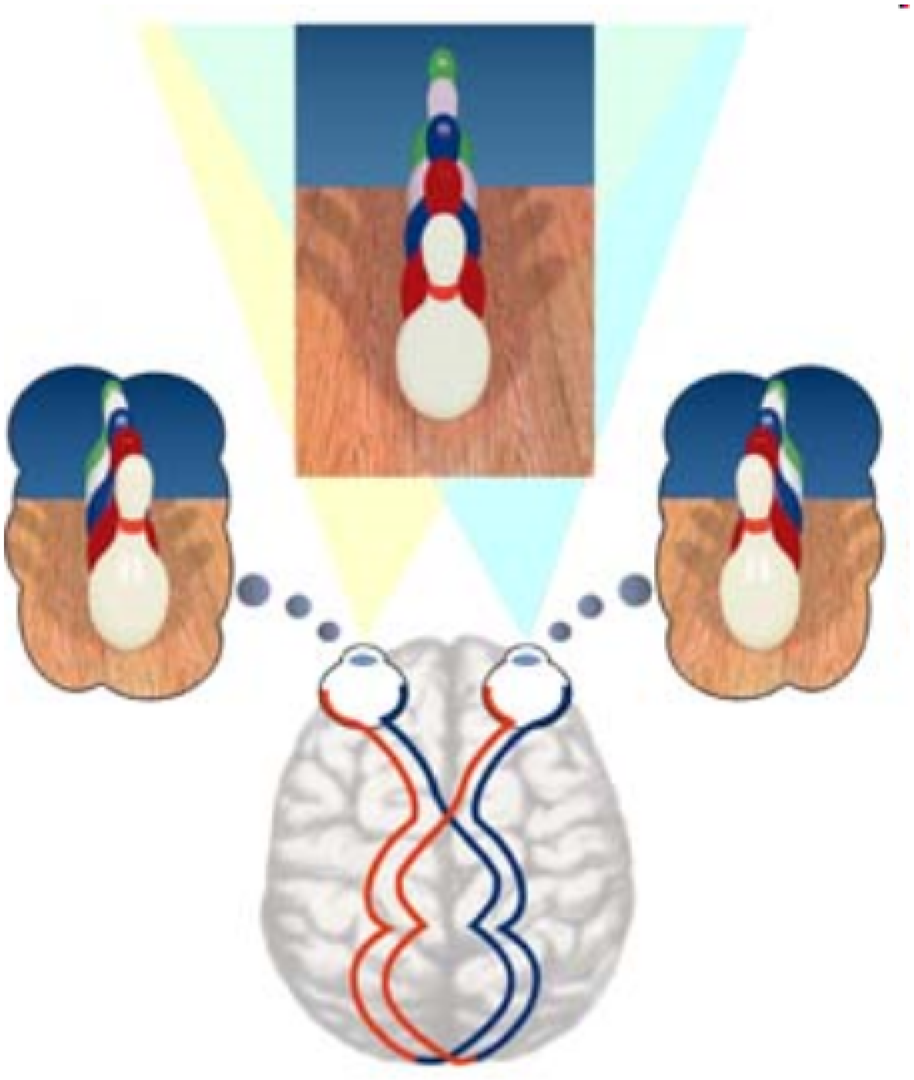
- Camera Model ------->p=MP
- M=K[R T]
- Camera Calibration ------>K
- Known: p, K
- Unknown: R, T, P
- Goal: P

Single View

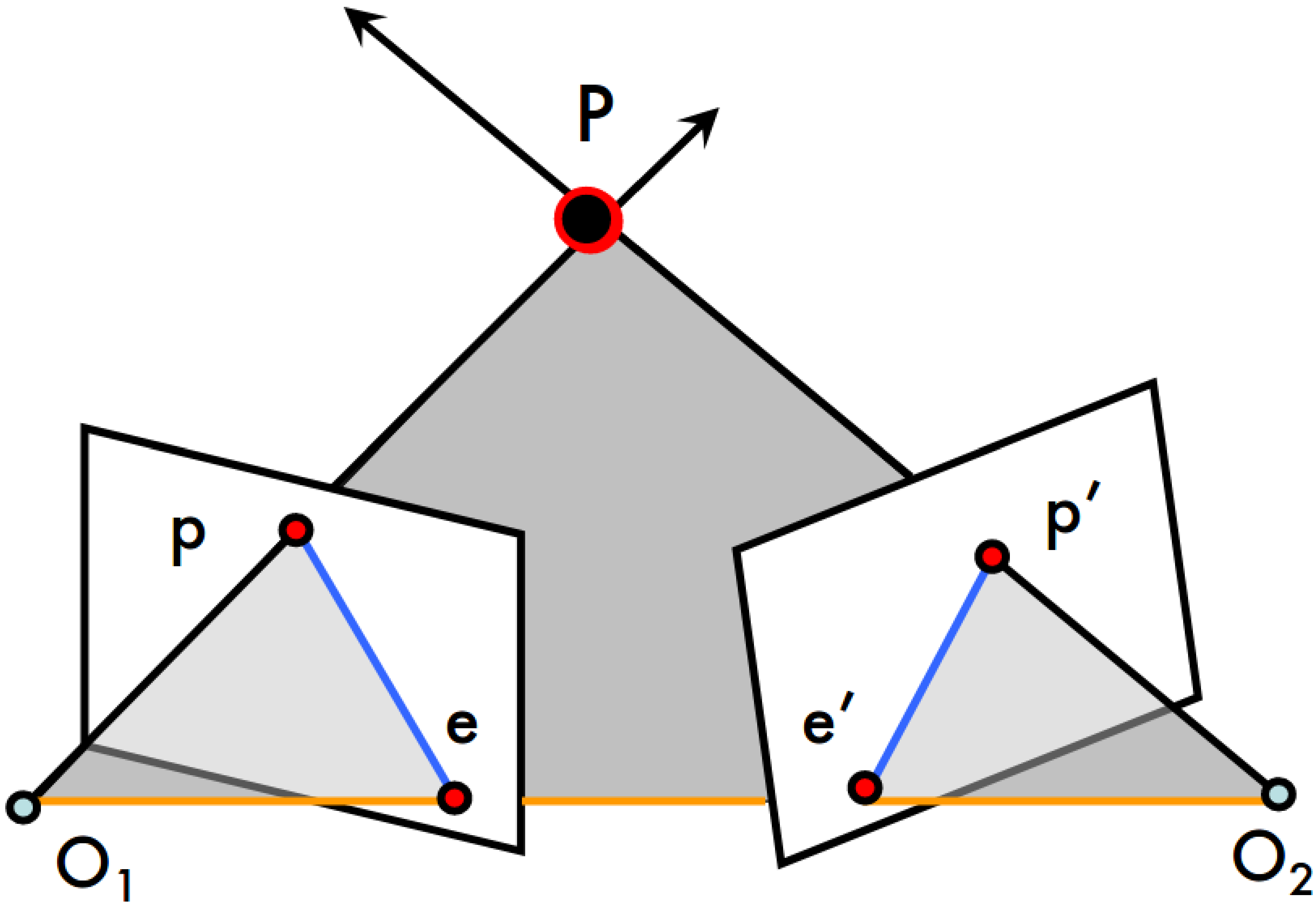


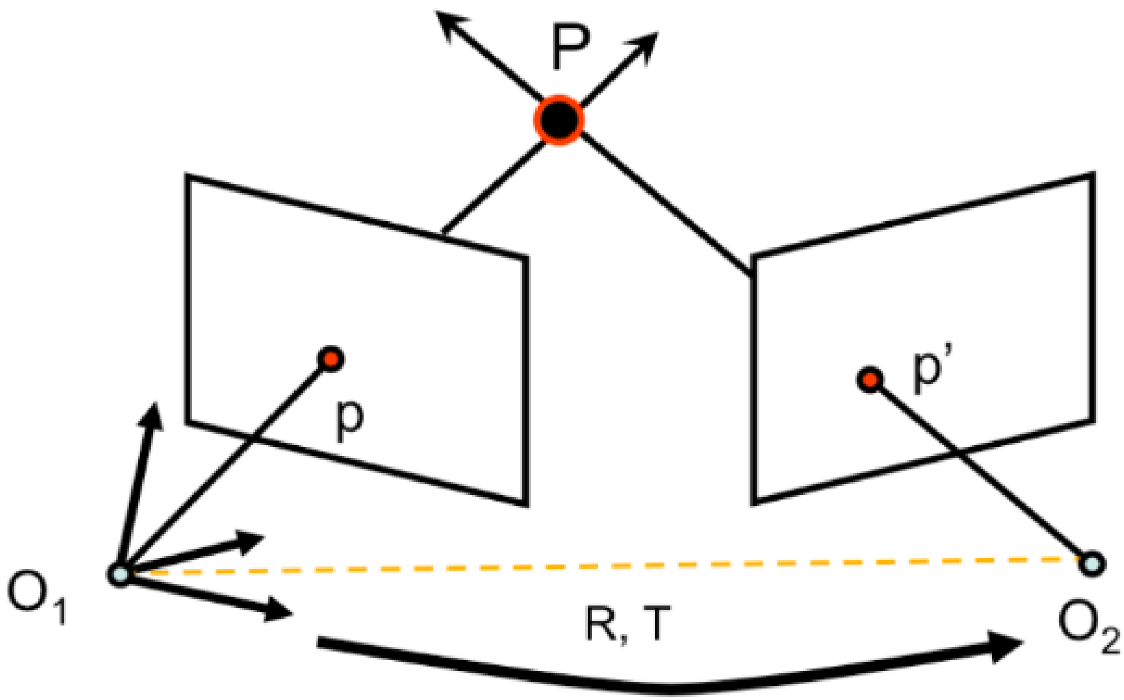
If K is known, can we get P from single view?

Multi (stereo)-view geometry



Epipolar Geometry



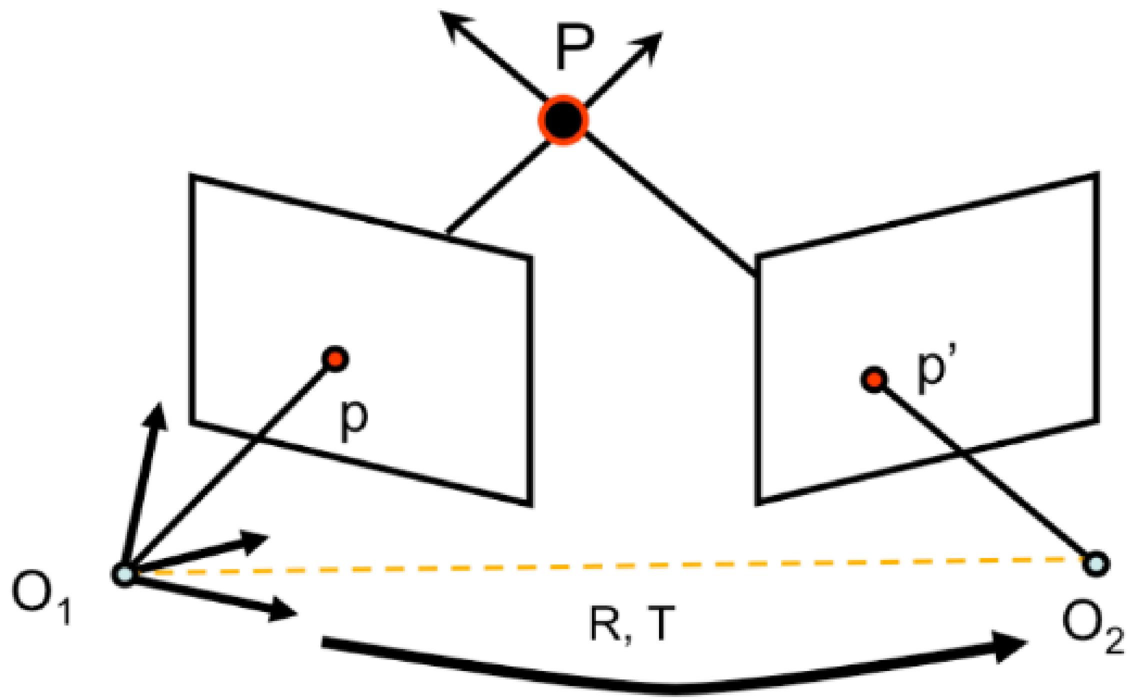


$$p = MP \qquad p' = M'P$$

$$M = K[I\ \ 0] \qquad M' = K'[R\ \ T]$$

Let K=K’=I(canonical cameras)

$$M = [I\ \ 0] \qquad M' = [R\ \ T]$$



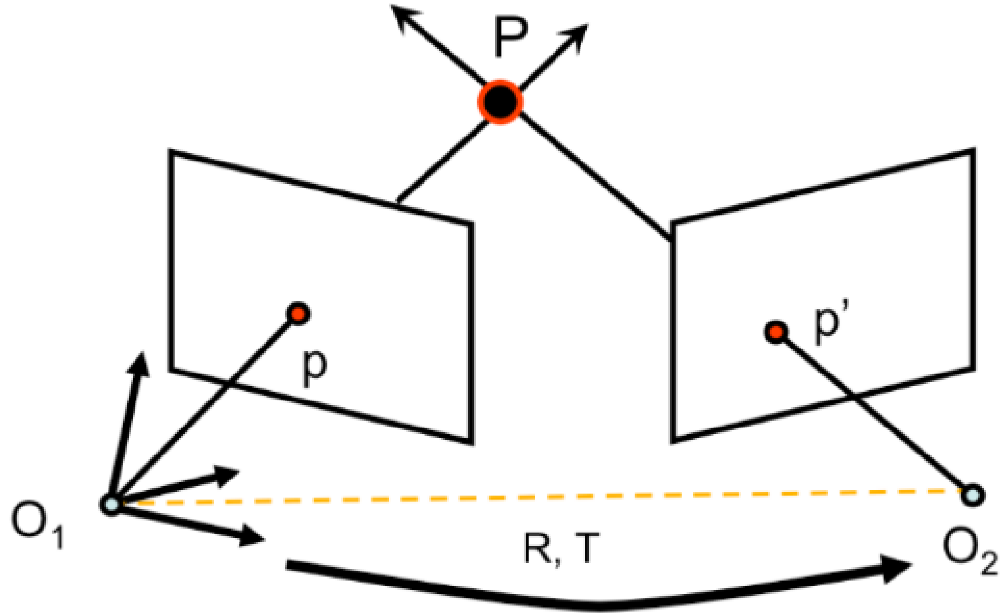
p′ in first camera reference system is $= R\,p' + T$

$T \times ((R\,p') + T) = T \times (R\,p')$ is perpendicular to epipolar plane

$$\rightarrow p^T \cdot [T \times (R\,p')] = 0$$

Recall: Cross product as matrix multiplication

$$\mathbf{a} \times \mathbf{b} = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix} \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix} = [\mathbf{a}_\times]\mathbf{b}$$



$$p^T \cdot [T \times (R\,p')] = 0 \rightarrow p^T \cdot \boxed{[T_\times] \cdot R}\, p' = 0$$

E = Essential matrix

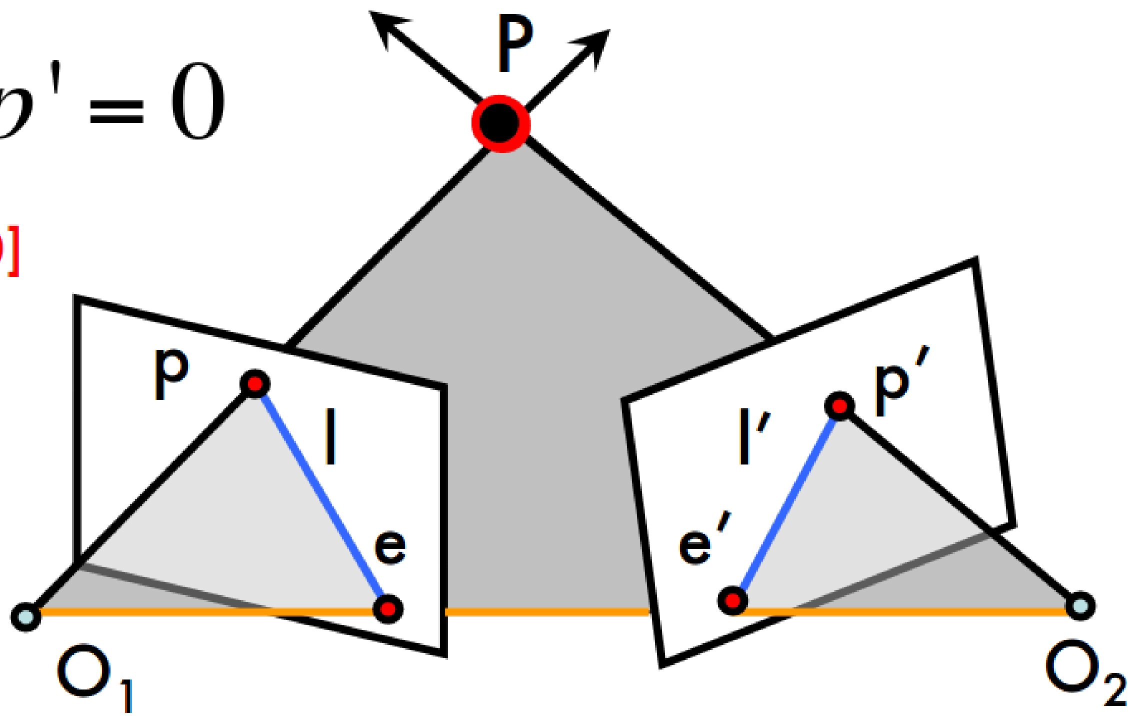
(Longuet-Higgins, 1981)

$$p^T \cdot E\, p' = 0$$ **Epipolar Constraint**

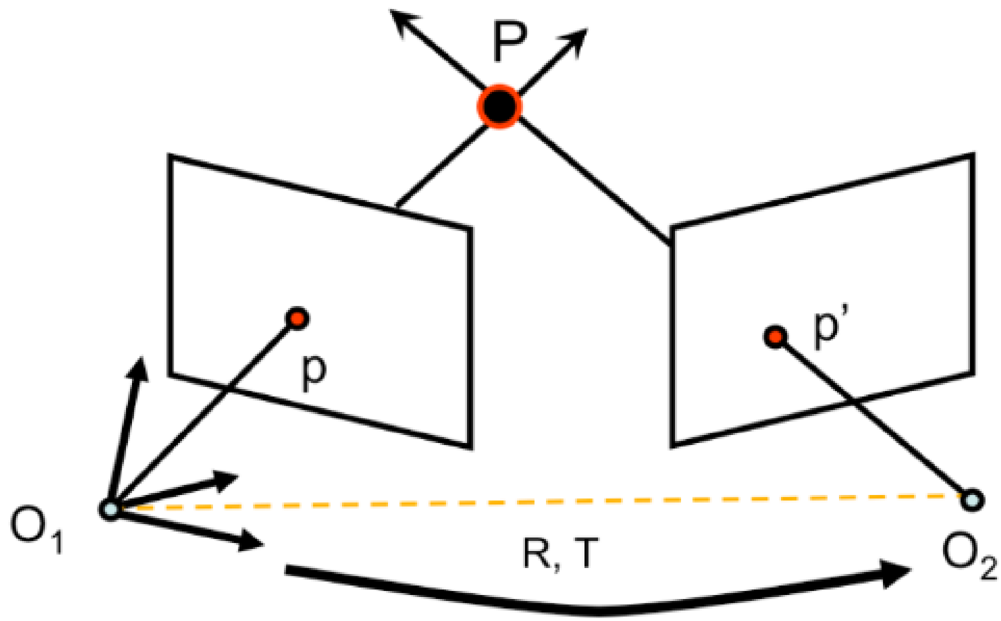
Epipolar Constraint

$$p^T \cdot E\, p' = 0$$

[Eq. 10]

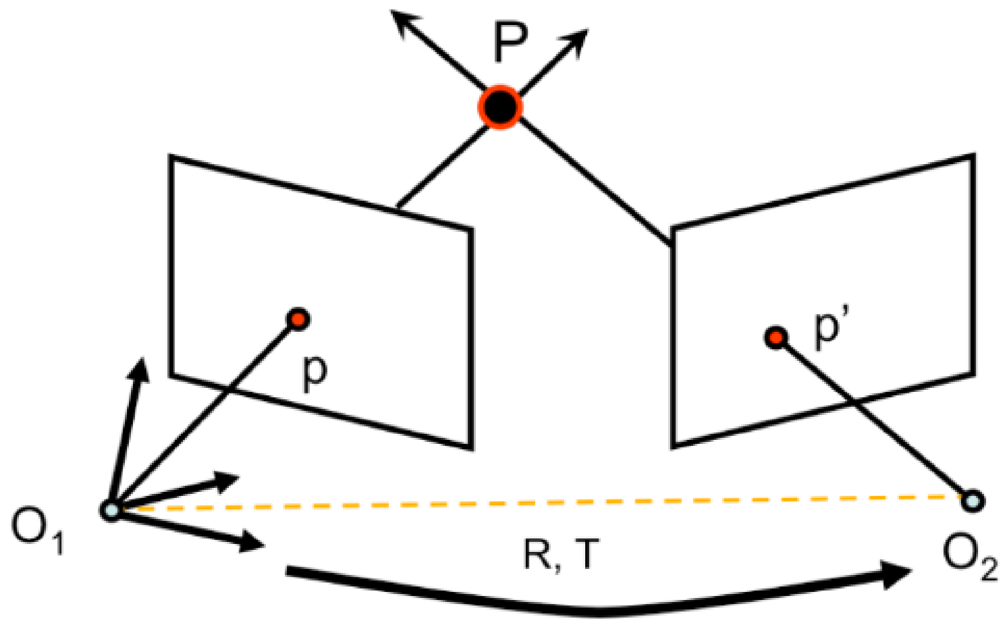


- $l = E\, p'$ is the epipolar line associated with p'
- $l' = E^T p$ is the epipolar line associated with p
- $E\, e' = 0$ and $E^T e = 0$
- E is 3x3 matrix; 5 DOF
- E is singular (rank two)



$$M = K\begin{bmatrix} I & 0 \end{bmatrix} \qquad M' = K'\begin{bmatrix} R^T & -R^T T \end{bmatrix}$$

$$p_c = K^{-1}\ p \qquad p'_c = K'^{-1}\ p'$$



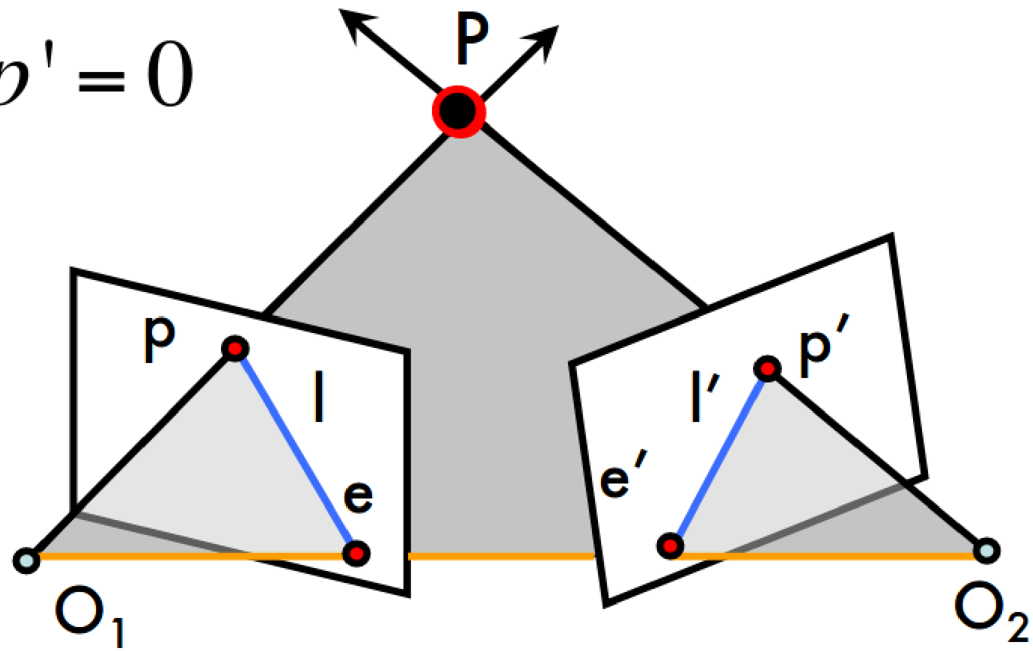
$$p_c^T \cdot [T_\times] \cdot R\, p_c' = 0 \rightarrow (\mathrm{K}^{-1}\,\mathrm{p})^{\mathrm{T}} \cdot [\mathrm{T}_\times] \cdot \mathrm{R}\,\mathrm{K}'^{-1}\,\mathrm{p}' = 0$$

$$\mathrm{p}^{\mathrm{T}}\, \boxed{\mathrm{K}^{-\mathrm{T}} \cdot [\mathrm{T}_\times] \cdot \mathrm{R}\,\mathrm{K}'^{-1}}\, \mathrm{p}' = 0 \rightarrow \mathrm{p}^{\mathrm{T}}\, \boxed{\mathrm{F}}\, \mathrm{p}' = 0$$

F = Fundamental Matrix (Faugeras and Luong, 1992)

Epipolar Constraint

$$p^T \cdot F\ p' = 0$$



- $l = F\ p'$ is the epipolar line associated with p'
- $l' = F^T p$ is the epipolar line associated with p
- $F\ e' = 0$ and $F^T e = 0$
- F is 3x3 matrix; 7 DOF
- F is singular (rank two)

Epipolar Constraint

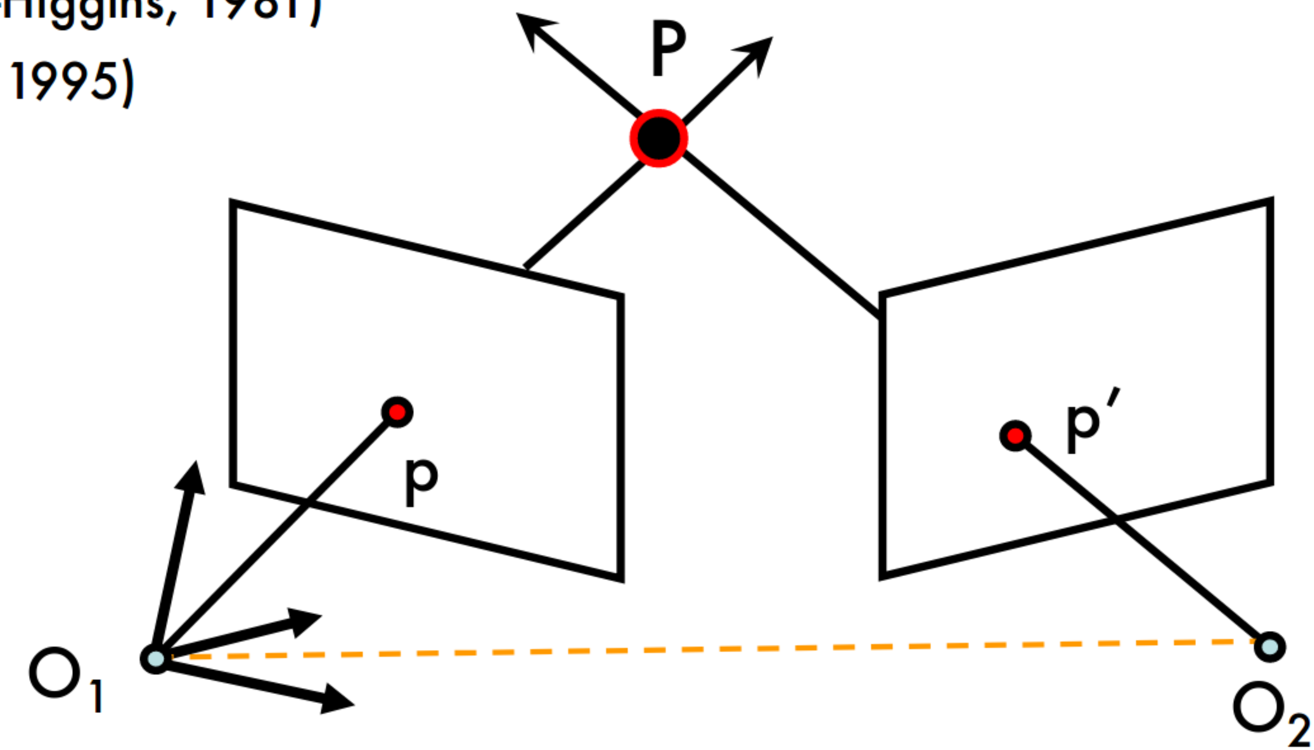
http://users.csc.calpoly.edu/~zwood/teaching/csc572/final08/kmerriman/

Estimating F

The Eight-Point Algorithm

(Longuet-Higgins, 1981)

(Hartley, 1995)



$$p^{T} F p' = 0$$

$$\mathrm{p}^{\mathrm{T}}\,\mathrm{F}\,\mathrm{p}'=0 \quad \Rightarrow \qquad p=\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad p'=\begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix}$$

$$(u,v,1)\begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix}\begin{pmatrix} u' \\ v' \\ 1 \end{pmatrix}=0$$

$$\Rightarrow \quad (uu',uv',u,vu',vv',v,u',v',1)\begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{pmatrix}=0$$

$$\underbrace{\begin{pmatrix} u_1u'_1 & u_1v'_1 & u_1 & v_1u'_1 & v_1v'_1 & v_1 & u'_1 & v'_1 & 1 \\ u_2u'_2 & u_2v'_2 & u_2 & v_2u'_2 & v_2v'_2 & v_2 & u'_2 & v'_2 & 1 \\ u_3u'_3 & u_3v'_3 & u_3 & v_3u'_3 & v_3v'_3 & v_3 & u'_3 & v'_3 & 1 \\ u_4u'_4 & u_4v'_4 & u_4 & v_4u'_4 & v_4v'_4 & v_4 & u'_4 & v'_4 & 1 \\ u_5u'_5 & u_5v'_5 & u_5 & v_5u'_5 & v_5v'_5 & v_5 & u'_5 & v'_5 & 1 \\ u_6u'_6 & u_6v'_6 & u_6 & v_6u'_6 & v_6v'_6 & v_6 & u'_6 & v'_6 & 1 \\ u_7u'_7 & u_7v'_7 & u_7 & v_7u'_7 & v_7v'_7 & v_7 & u'_7 & v'_7 & 1 \\ u_8u'_8 & u_8v'_8 & u_8 & v_8u'_8 & v_8v'_8 & v_8 & u'_8 & v'_8 & 1 \end{pmatrix}}_{\mathbf{W}} \underbrace{\begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{pmatrix}}_{\mathbf{f}} = 0$$

- Homogeneous system $\mathbf{W}\,\mathbf{f} = 0$
- Rank 8 ⟶ A non-zero solution exists (unique)
- If N>8 ⟶ Lsq. solution by SVD! ⟶ $\hat{\mathrm{F}}$

$\hat{F}$ satisfies: $\mathrm{p}^{\mathrm{T}}\,\hat{\mathrm{F}}\,\mathrm{p}' = 0$

and estimated $\hat{\mathrm{F}}$ may have full rank ($\det(\hat{\mathrm{F}}) \neq 0$)

But remember: fundamental matrix is Rank2

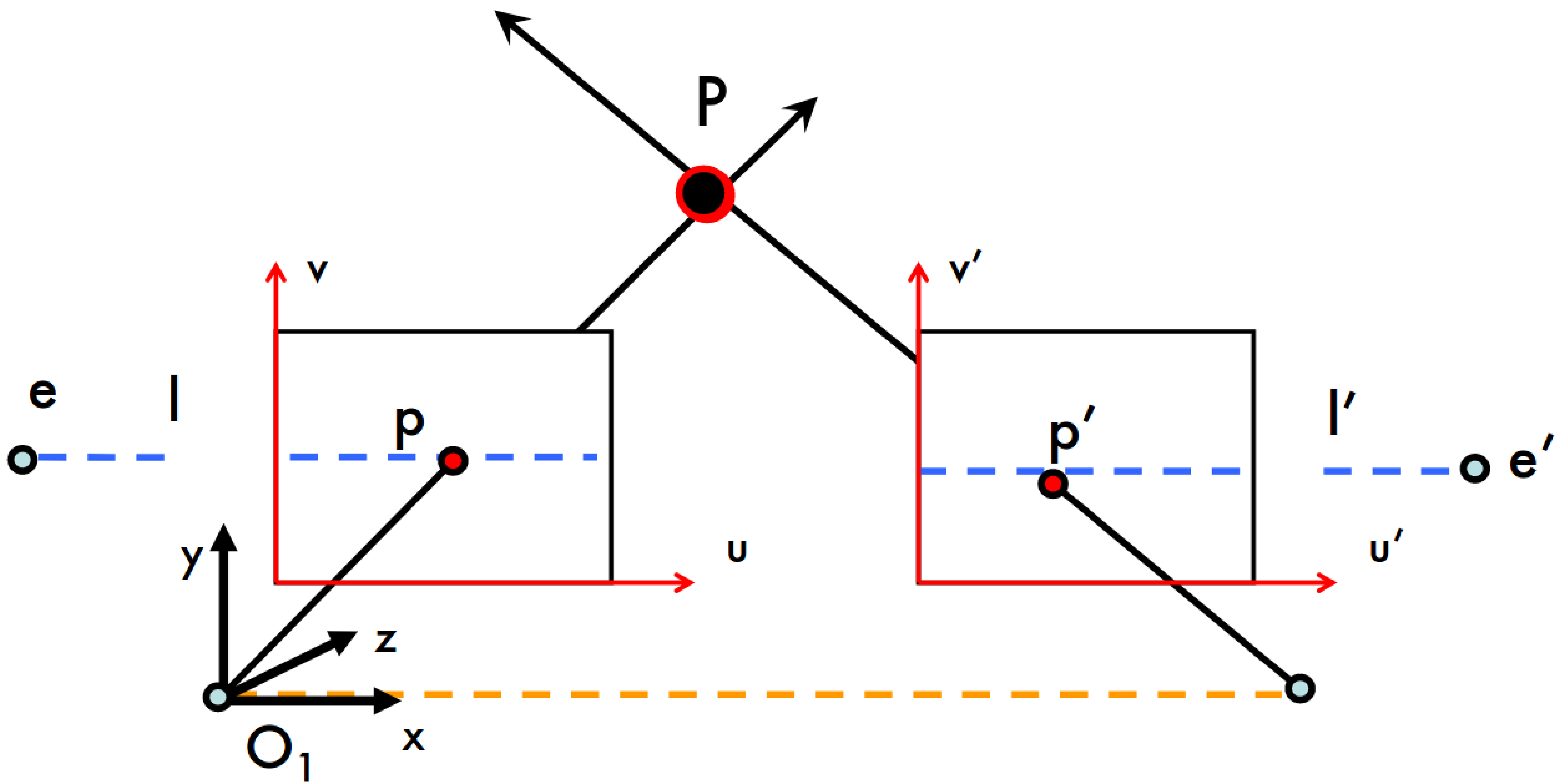
Find F that minimizes $\left\| \mathrm{F} - \hat{\mathrm{F}} \right\| = 0$

Frobenius norm (*)

Subject to det(F)=0

SVD (again!) can be used to solve this problem

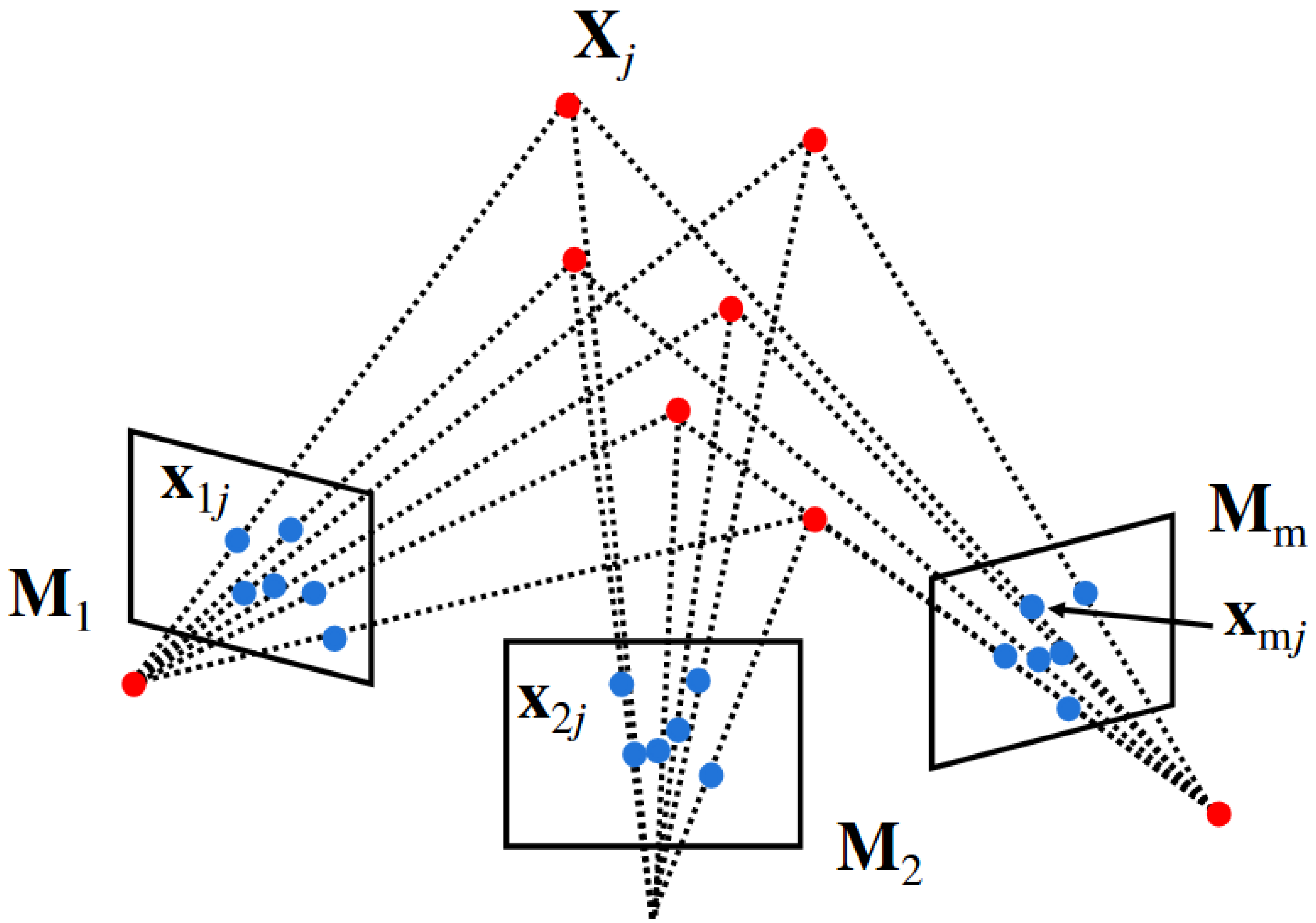
Parallel image planes



Conclusion

- Epipolar Constraint: pFp’=0
- F= $K^{-T} \cdot [T_{\times}] \cdot R\, K'^{-1}$
- p, p’ are known –----->SVD---->F
- F, K is known ------>R, T
- p=K[R T]P
- p, K, R, T are known----->P

Structure from Motion (SFM)



SFM(Structure from Motion)

- Advantages:

  RGB

  Cheap equipment

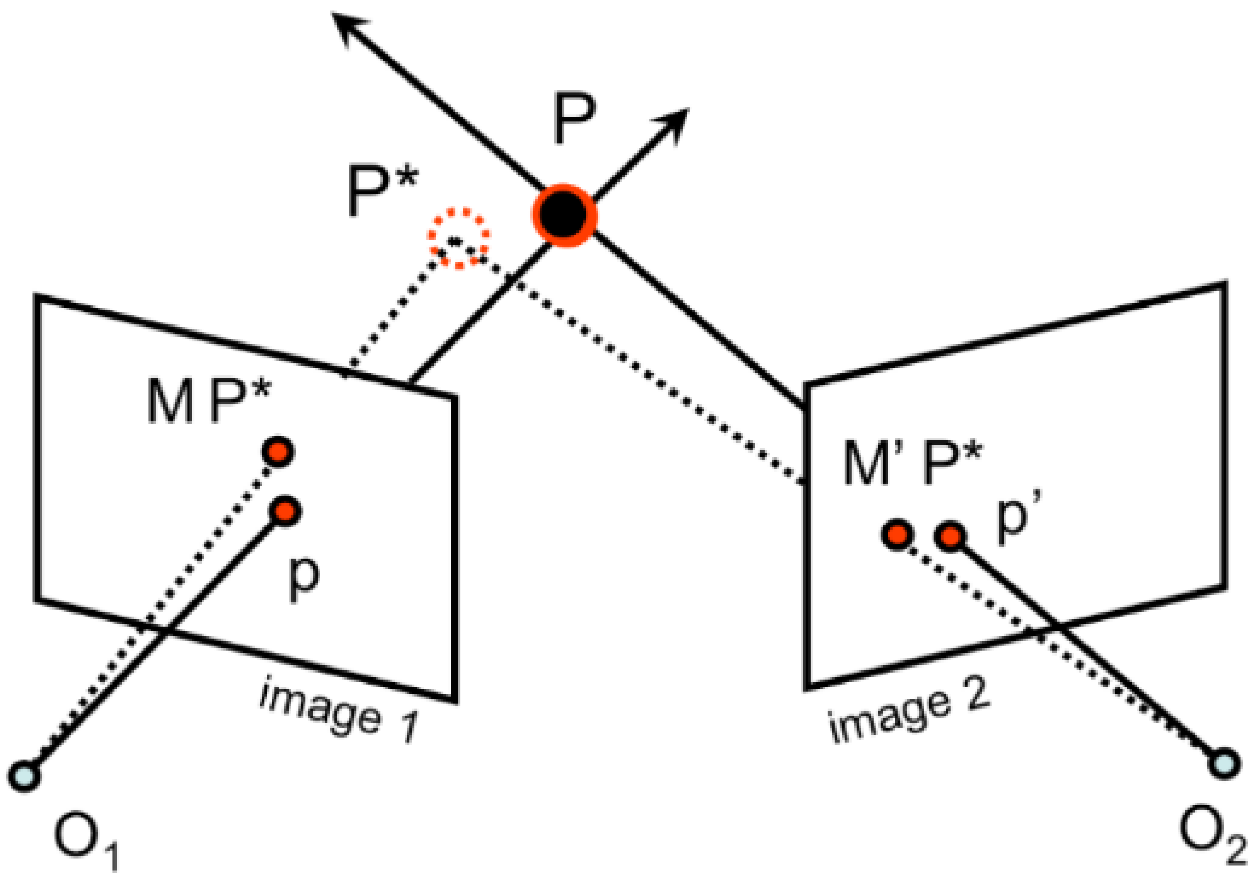
  Less limited by the environment

- Disadvantages:

  Accuracy

  Speed

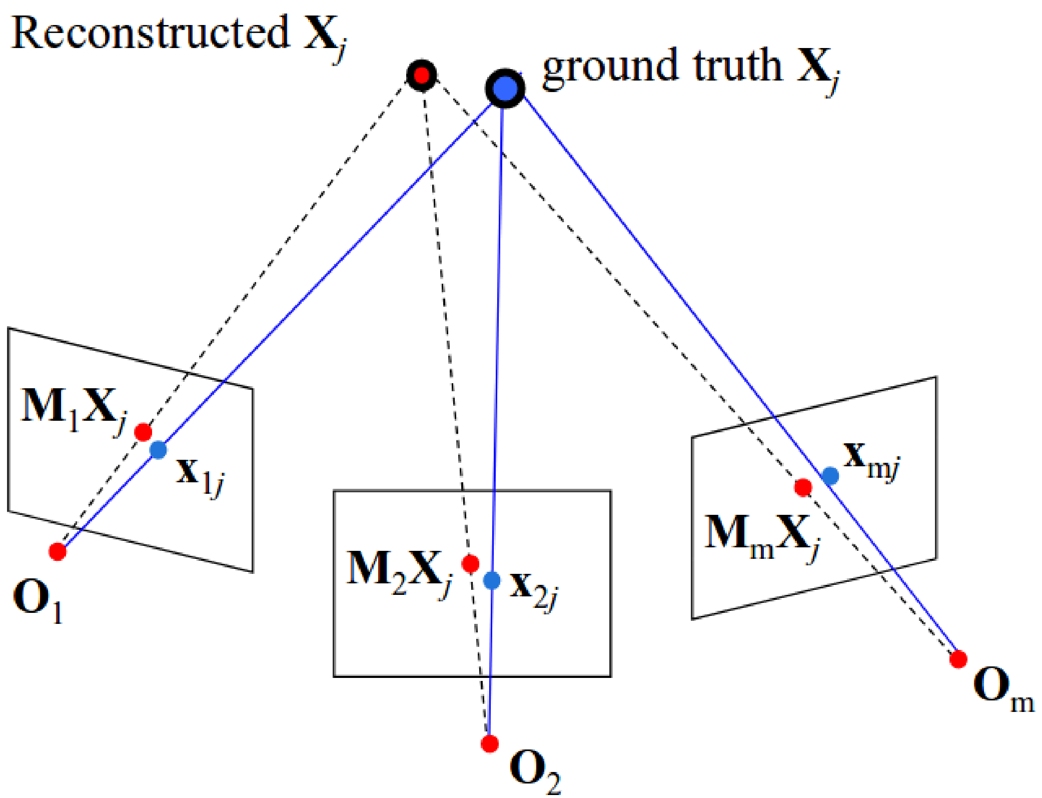
Triangulation



Bundle adjustment

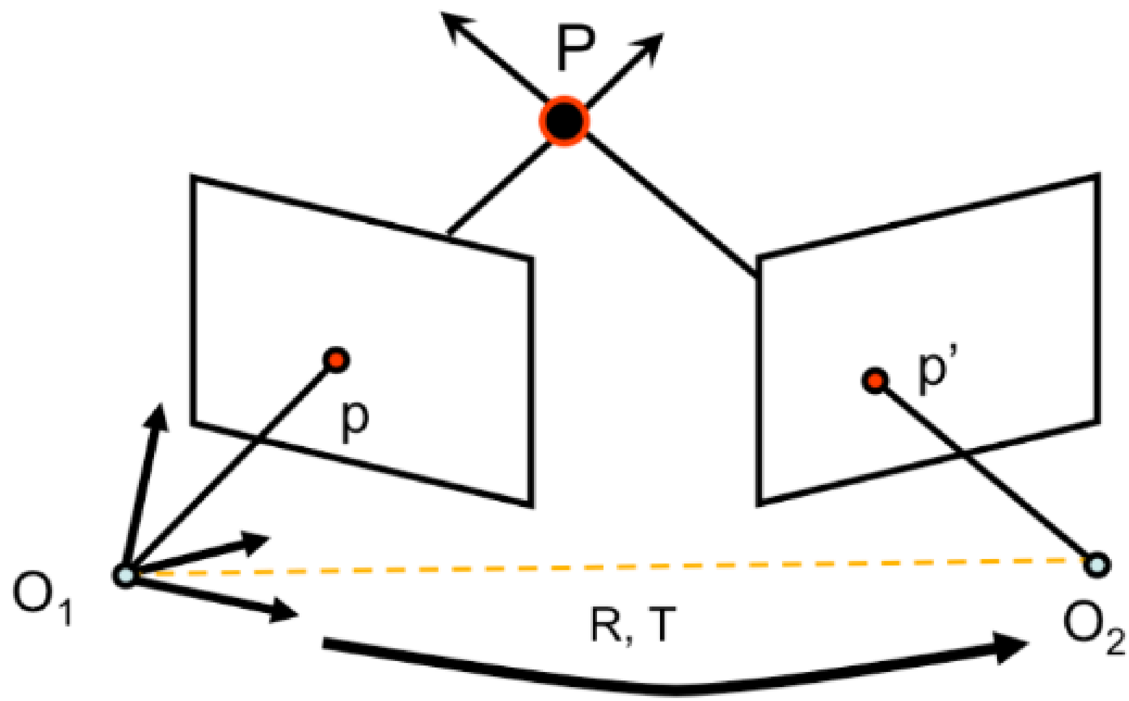
- Non-linear method for refining structure and motion
- Minimizes re-projection error

$$E(M, \mathbf{X}) = \sum_{i=1}^{m} \sum_{j=1}^{n} D\left(\mathbf{x}_{ij}, M_i \mathbf{X}_j\right)^2$$

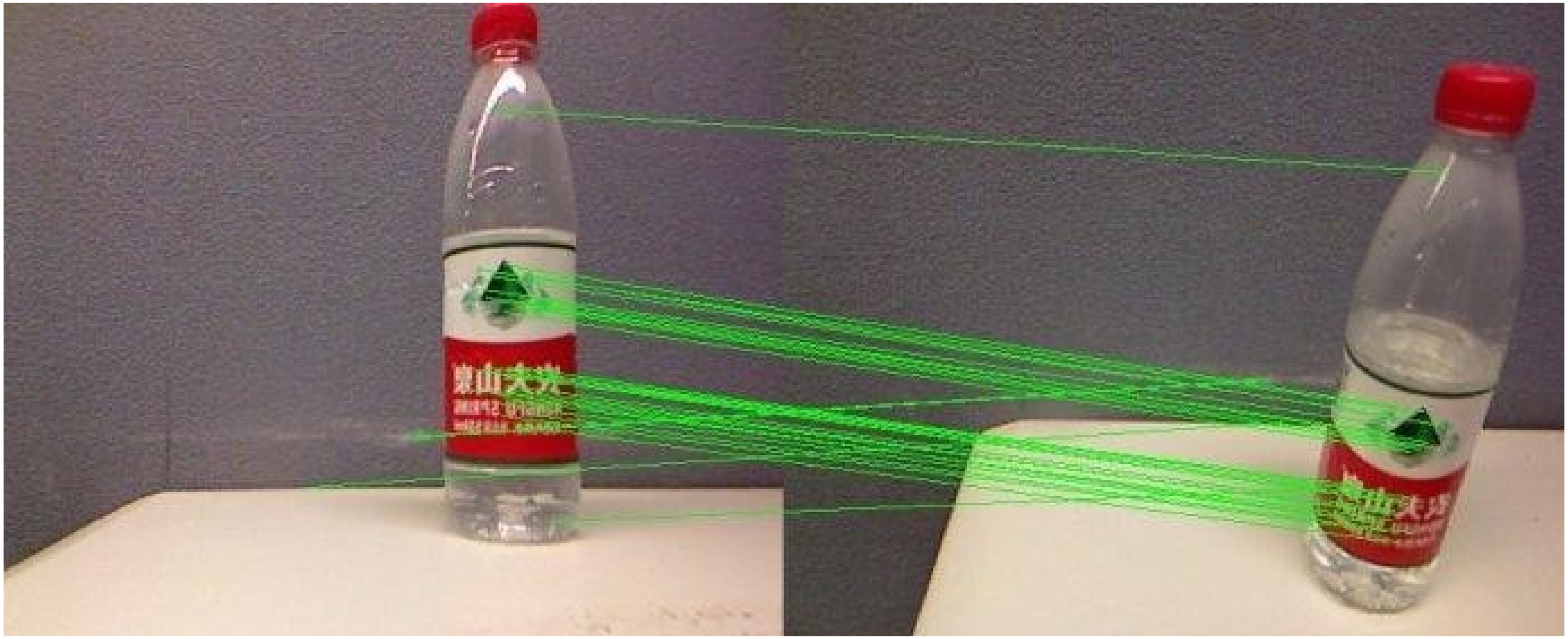


Bill T, et al. Bundle adjustment—a modern synthesis. International workshop on vision algorithms. 1999.

Descriptors



How do we know p and p' are corresponding points?



We need a descriptor

Scale Invariant Feature Transform(SIFT)

David G. Lowe. "Distinctive image features from scale-invariant keypoints." IJCV 60 (2), 04

Scale Invariant Feature Transform(SIFT)

Detector

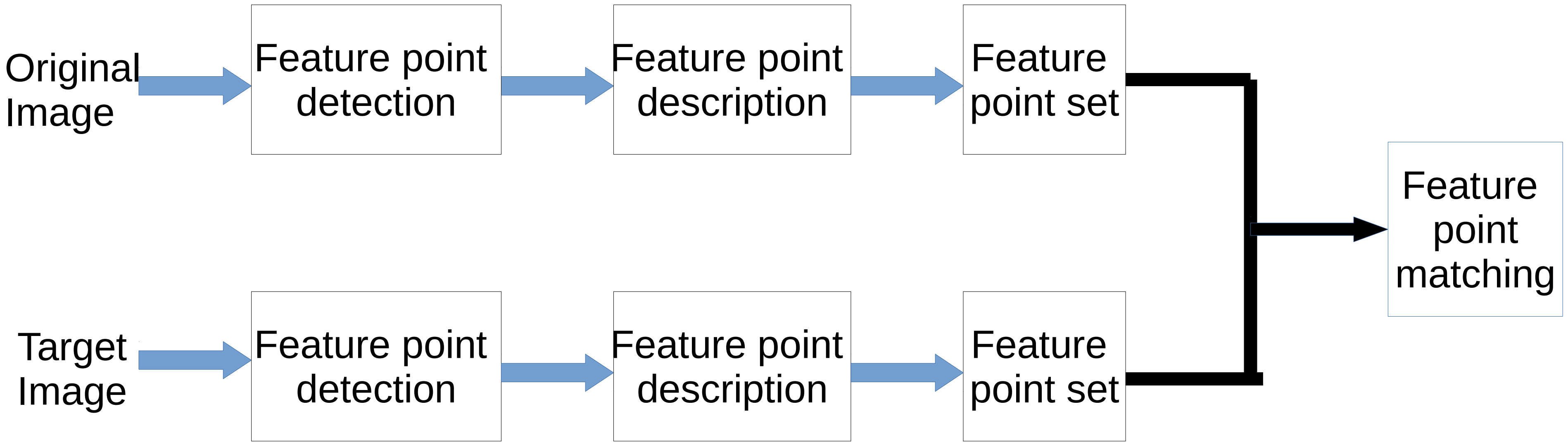
1. Find Scale-Space Extrema
2. Keypoint Localization & Filtering
   - Improve keypoints and throw out bad ones

3. **Orientation Assignment**
   - Remove effects of rotation and scale
4. Create descriptor
   - Using histograms of orientations

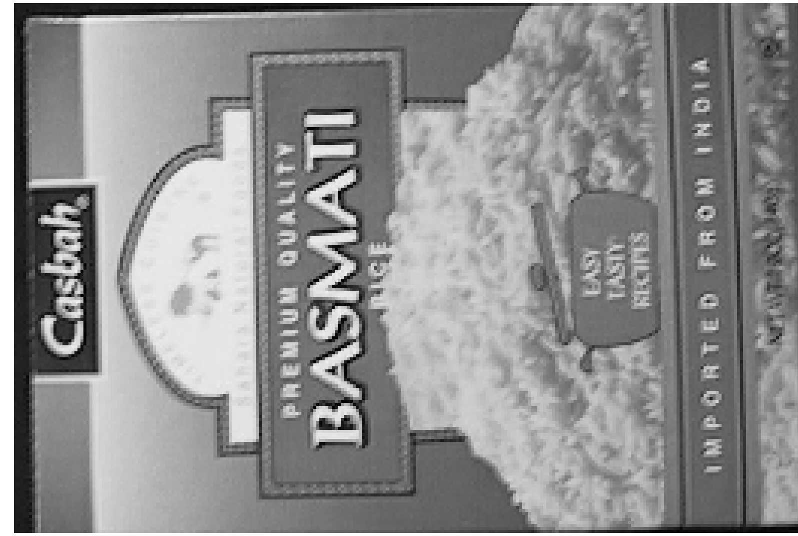
Descriptor

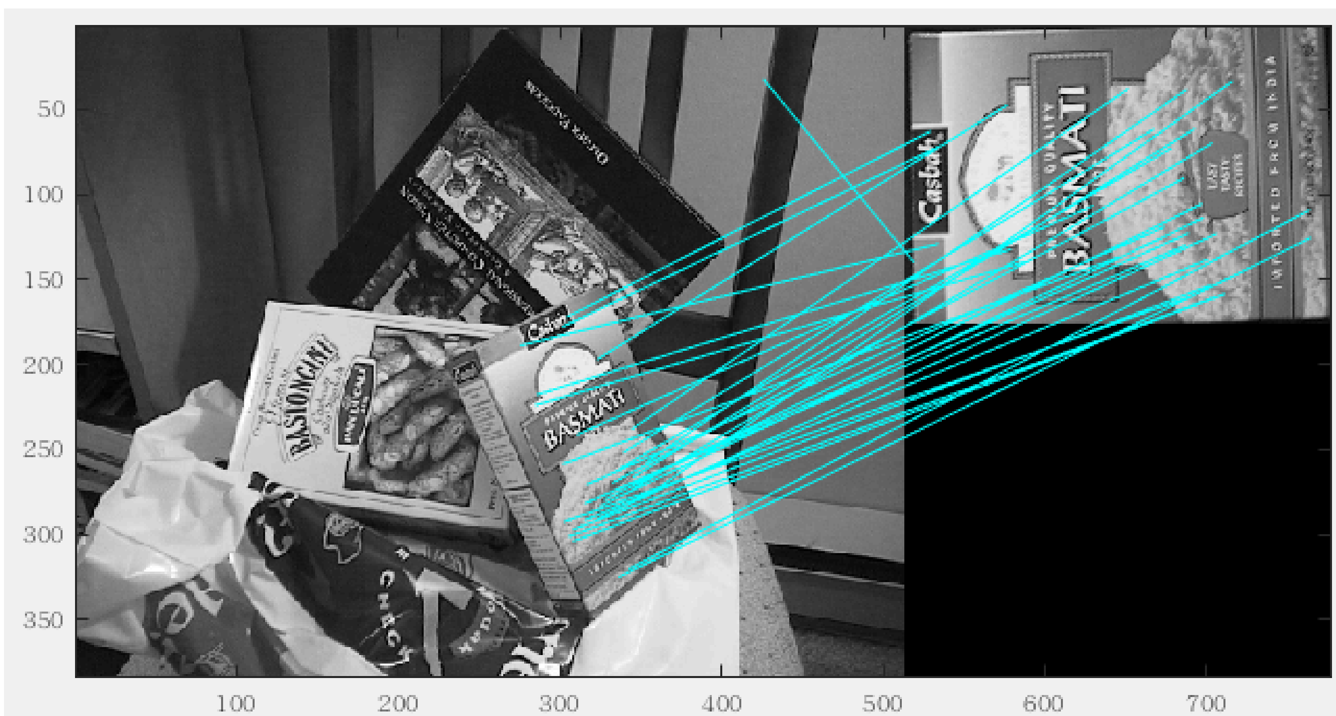
David G. Lowe. "Distinctive image features from scale-invariant keypoints." IJCV 60 (2), 04

SIFT(Scale Invariant Feature Transform)




BASTONCINI
PAN DUCALE
Casbah
BASMATI


Casbah
PREMIUM QUALITY
BASMATI
RICE
EASY
TASTY
RECIPES
IMPORTED FROM INDIA


50
100
150
200
250
300
350
100
200
300
400
500
600
700

Dense Reconstruction

GOAL:
Sparse point clouds---→Dense point clouds

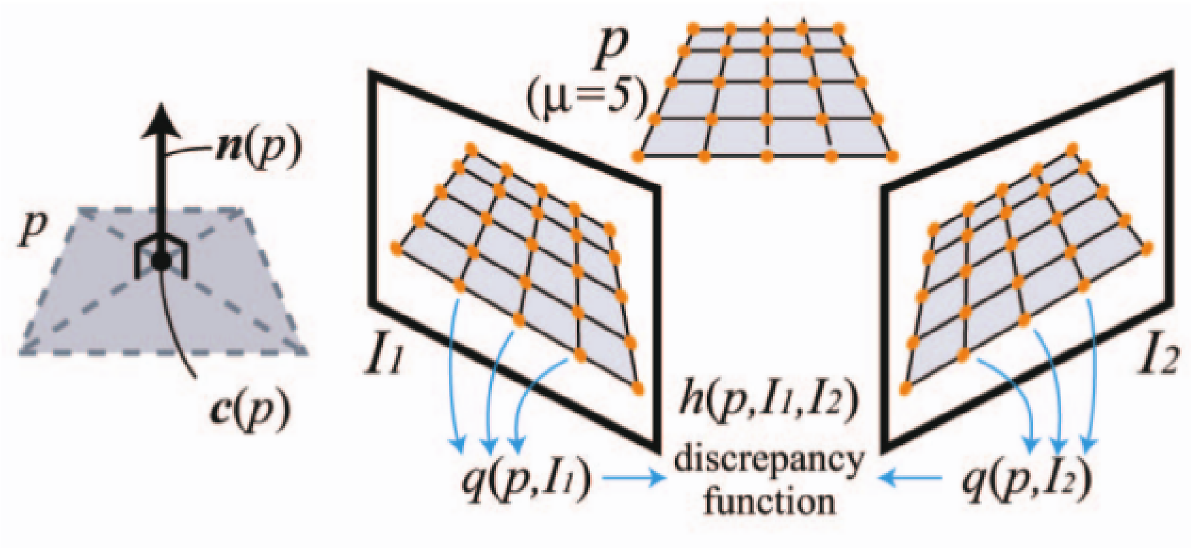
Patch-based Multi-view Stereo

Sparse point clouds
Camera parameters

$\Rightarrow$

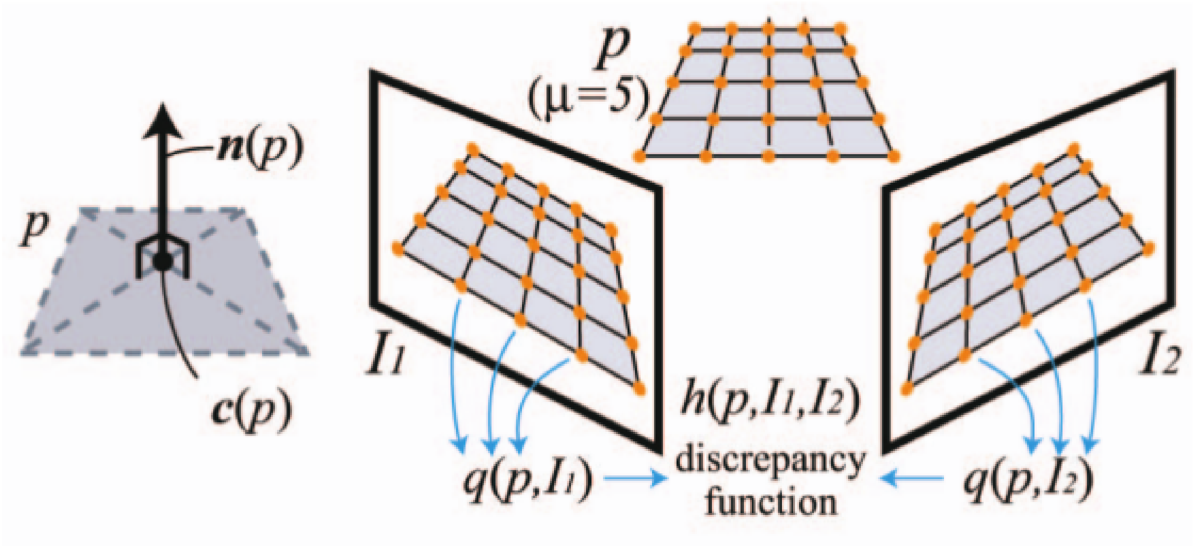
Dense point clouds

Patch-based Multi-view Stereo



Patch
- c(p)
- n(p)
- R(p)
- V(p)

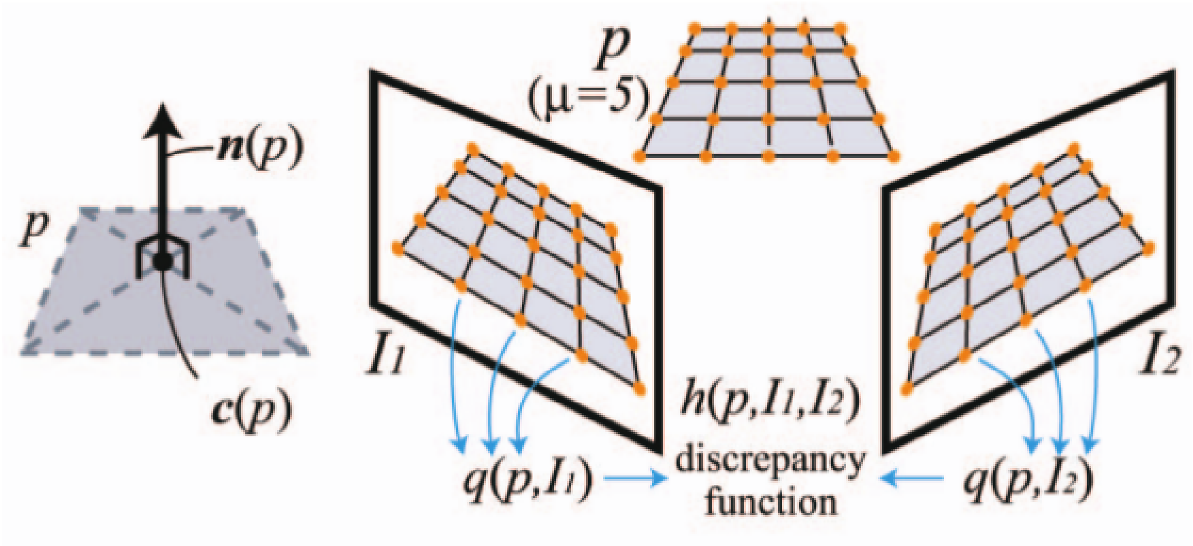
Yasutaka Furukawa and Jean Ponce, Accurate, Dense, and Robust Multiview Stereopsis, IEEE 2010



$$h(p, I_1, I_2) = 1 - NCC(q(p, I_1), q(p, I_2))$$

$$g(p) = \frac{1}{|V(p) \setminus R(p)|} \sum_{I \in V(p) \setminus R(p)} h(p, I, R(p))$$

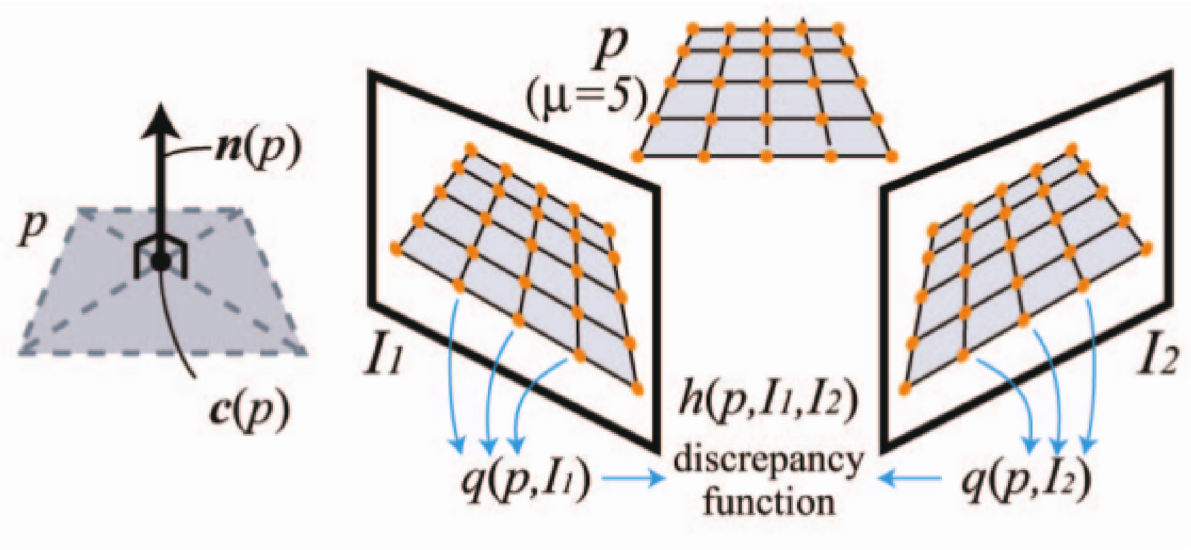
Yasutaka Furukawa and Jean Ponce, Accurate, Dense, and Robust Multiview Stereopsis, IEEE 2010



$$V^*(p) = \{I | I \in V(p), h(p, I, R(p)) \leq \alpha\},$$

$$g^*(p) = \frac{1}{|V^*(p) \setminus R(p)|} \sum_{I \in V^*(p) \setminus R(p)} h(p, I, R(p))$$

Yasutaka Furukawa and Jean Ponce, Accurate, Dense, and Robust Multiview Stereopsis, IEEE 2010

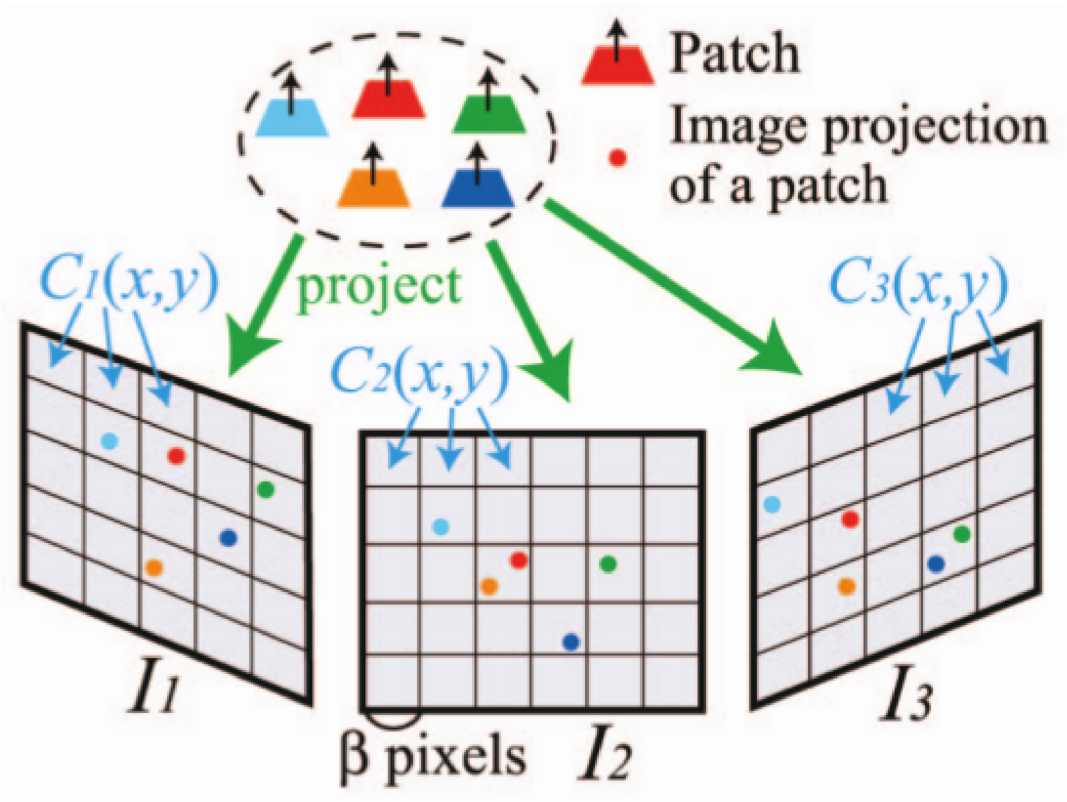


Patch Optimization：

- initialization of the corresponding parameters, namely, its center c(p), normal n(p), visible images V*(p), and the reference image R(p)
- optimization of its geometric component, c(p) and n(p)→ min g(p)

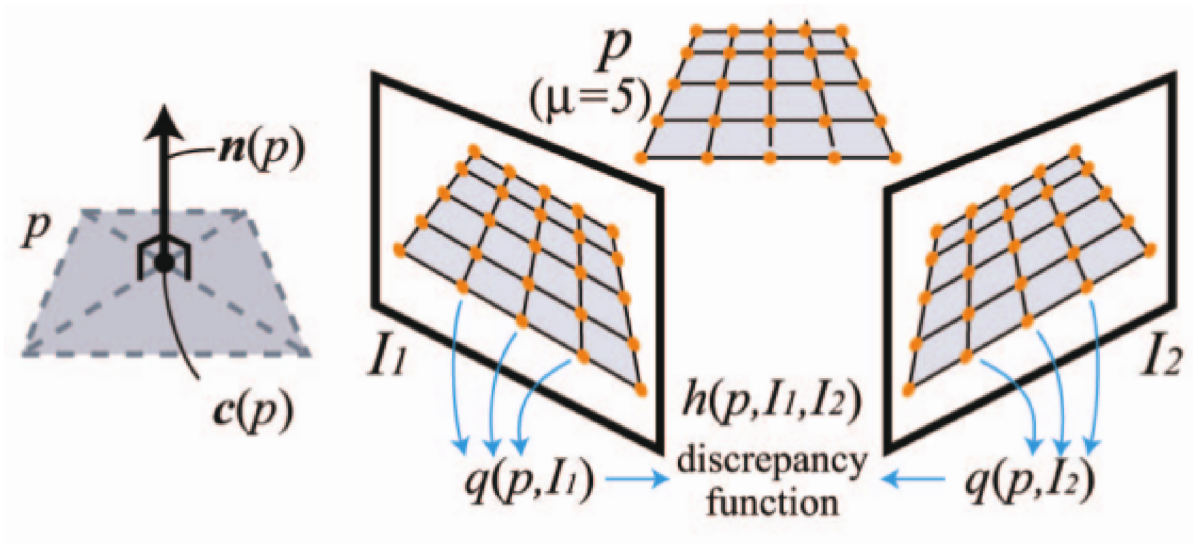
Yasutaka Furukawa and Jean Ponce, Accurate, Dense, and Robust Multiview Stereopsis, IEEE 2010

Image Model



Yasutaka Furukawa and Jean Ponce, Accurate, Dense, and Robust Multiview Stereopsis, IEEE 2010

Patch Reconstruction



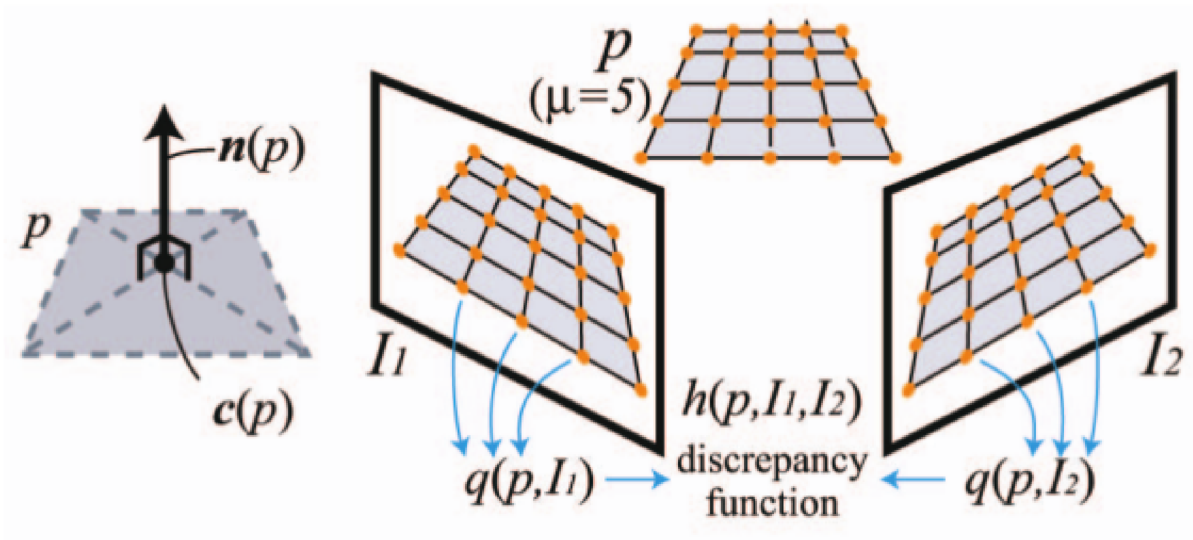
$$\mathbf{c}(p) \leftarrow \{\text{Triangulation from } f \text{ and } f'\}$$

$$\mathbf{n}(p) \leftarrow \overrightarrow{\mathbf{c}(p)O(I_i)} / |\overrightarrow{\mathbf{c}(p)O(I_i)}|,$$

$$R(p) \leftarrow I_i.$$

Yasutaka Furukawa and Jean Ponce, Accurate, Dense, and Robust Multiview Stereopsis, IEEE 2010

Patch Reconstruction



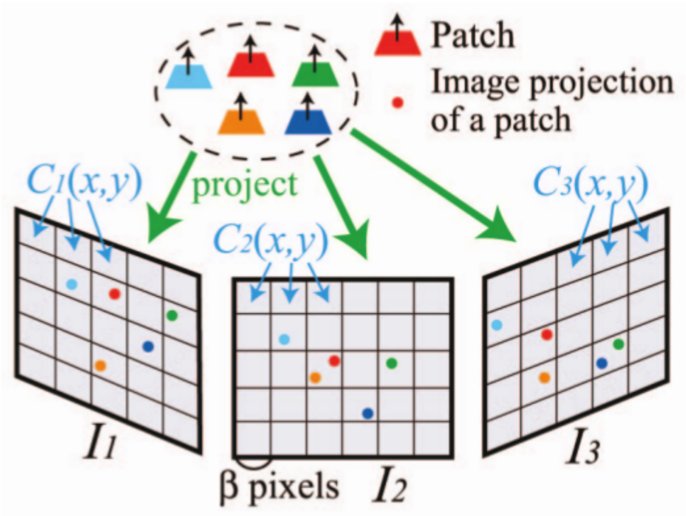
$$V(p) \leftarrow \left\{ I | \mathbf{n}(p) \cdot \overrightarrow{\mathbf{c}(p)O(I)} / |\overrightarrow{\mathbf{c}(p)O(I)}| > \cos(\iota) \right\}$$

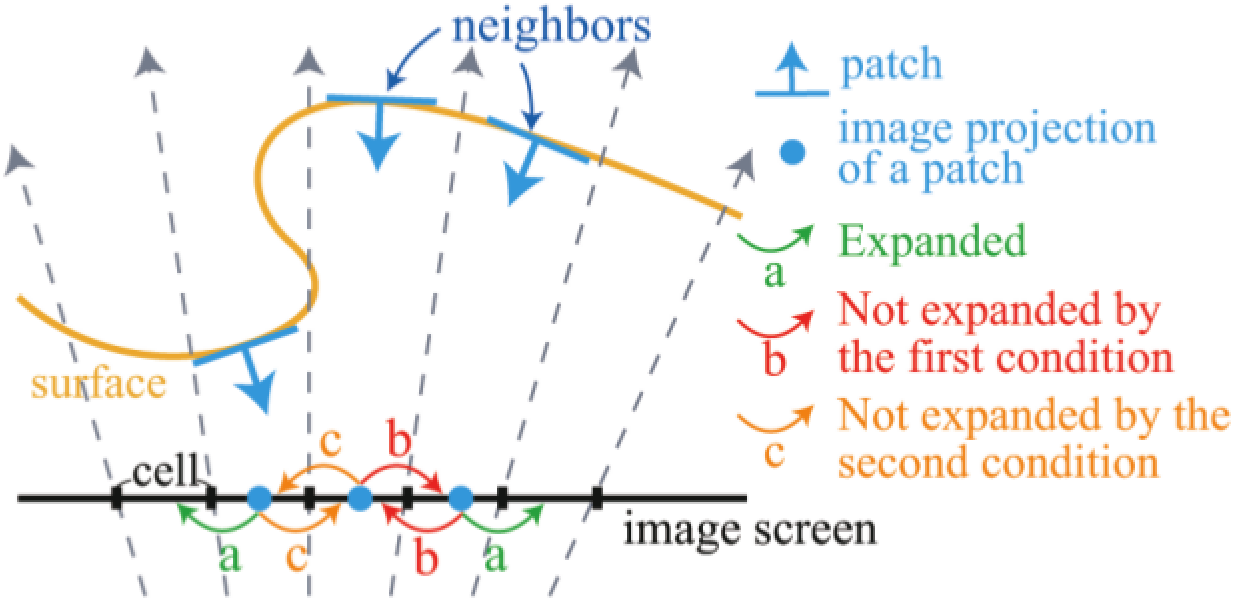
$$V^*(p) = \{I | I \in V(p), h(p, I, R(p)) \le \alpha\},$$

$$g^*(p) = \frac{1}{|V^*(p) \setminus R(p)|} \sum_{I \in V^*(p) \setminus R(p)} h(p, I, R(p))$$

Yasutaka Furukawa and Jean Ponce, Accurate, Dense, and Robust Multiview Stereopsis, IEEE 2010

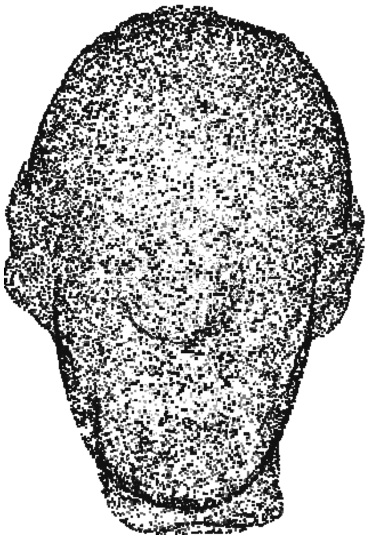
Patch Expansion

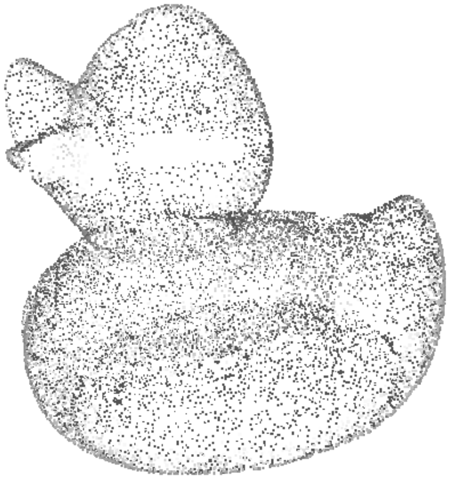


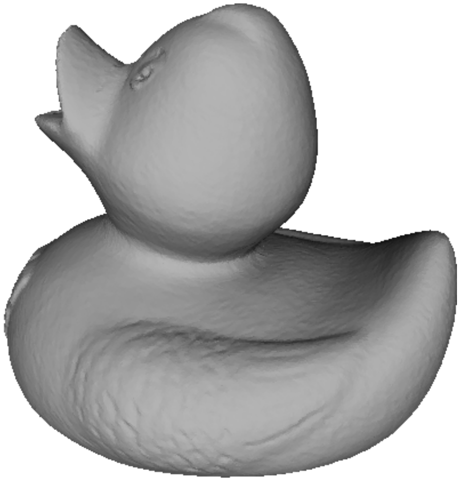


Yasutaka Furukawa and Jean Ponce, Accurate, Dense, and Robust Multiview Stereopsis, IEEE 2010

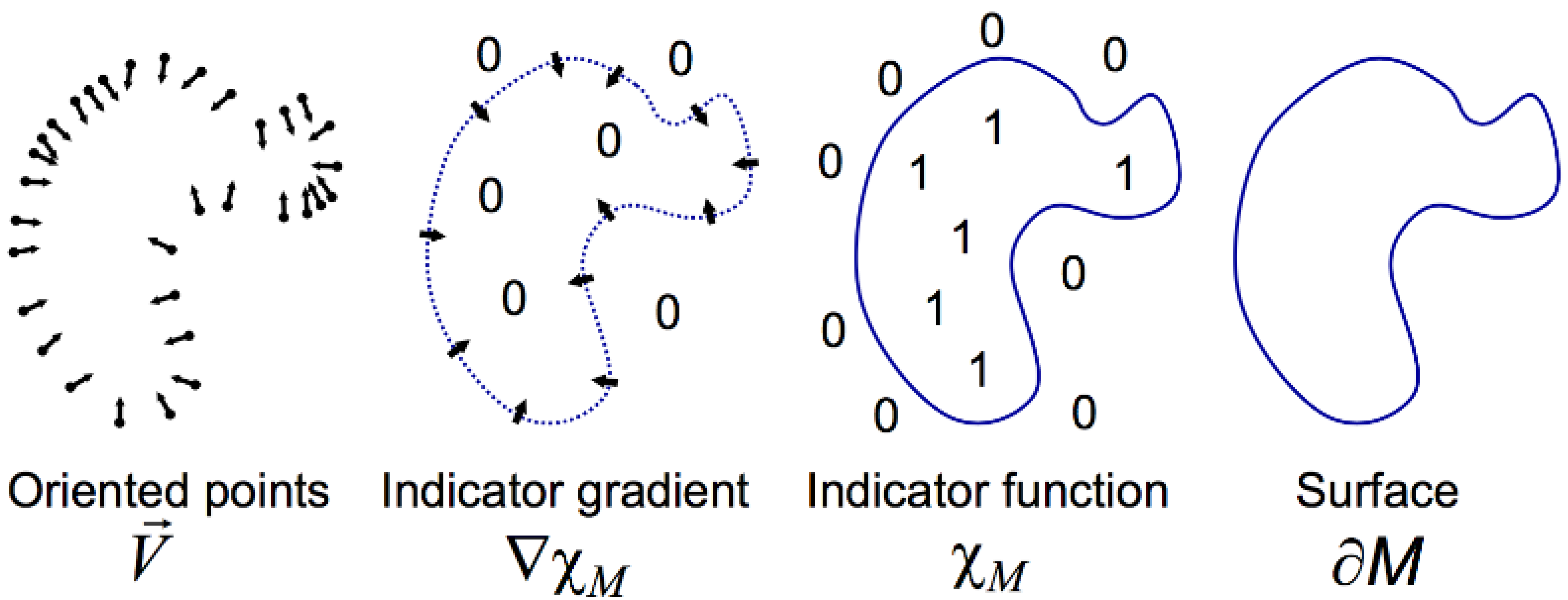
Surface Reconstruction





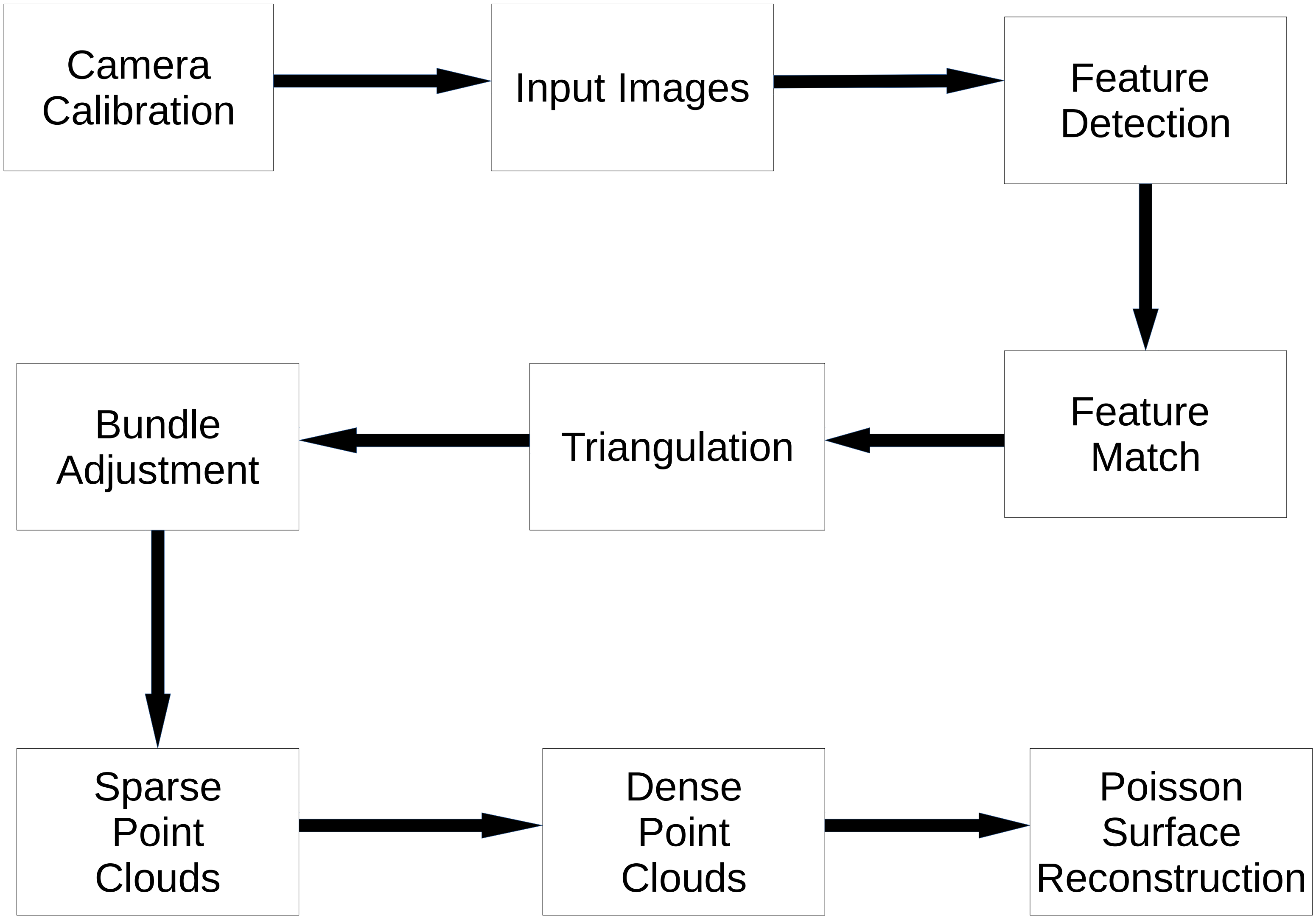


Poisson Surface Reconstruction

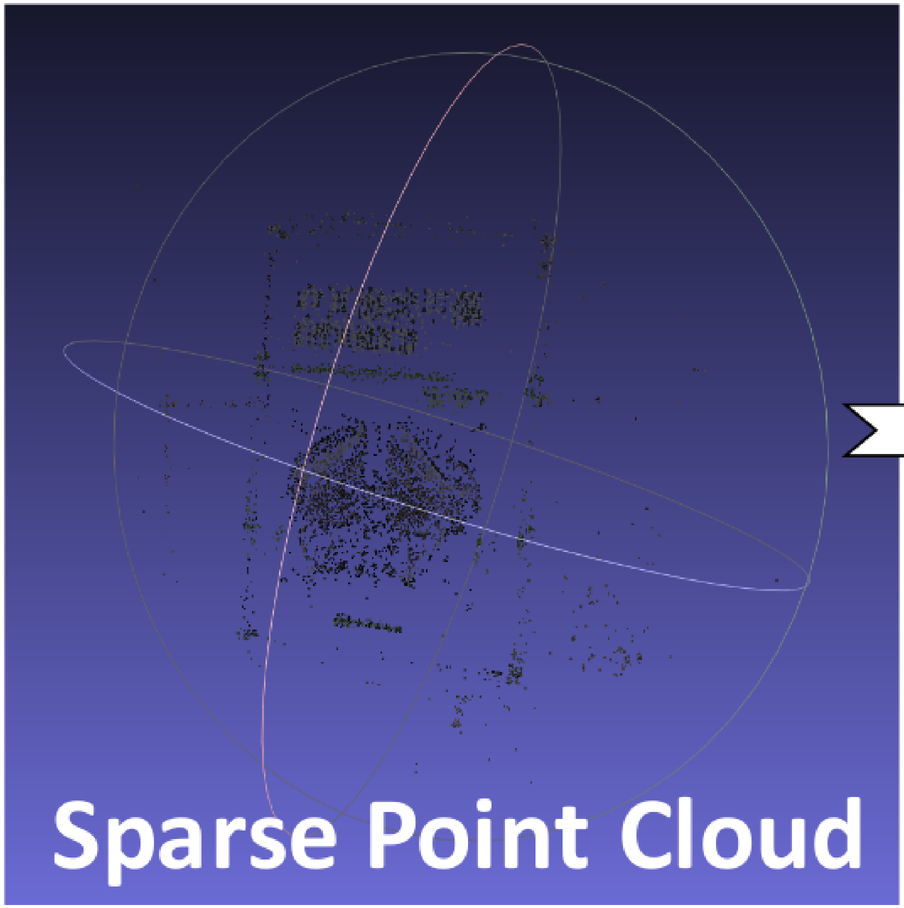


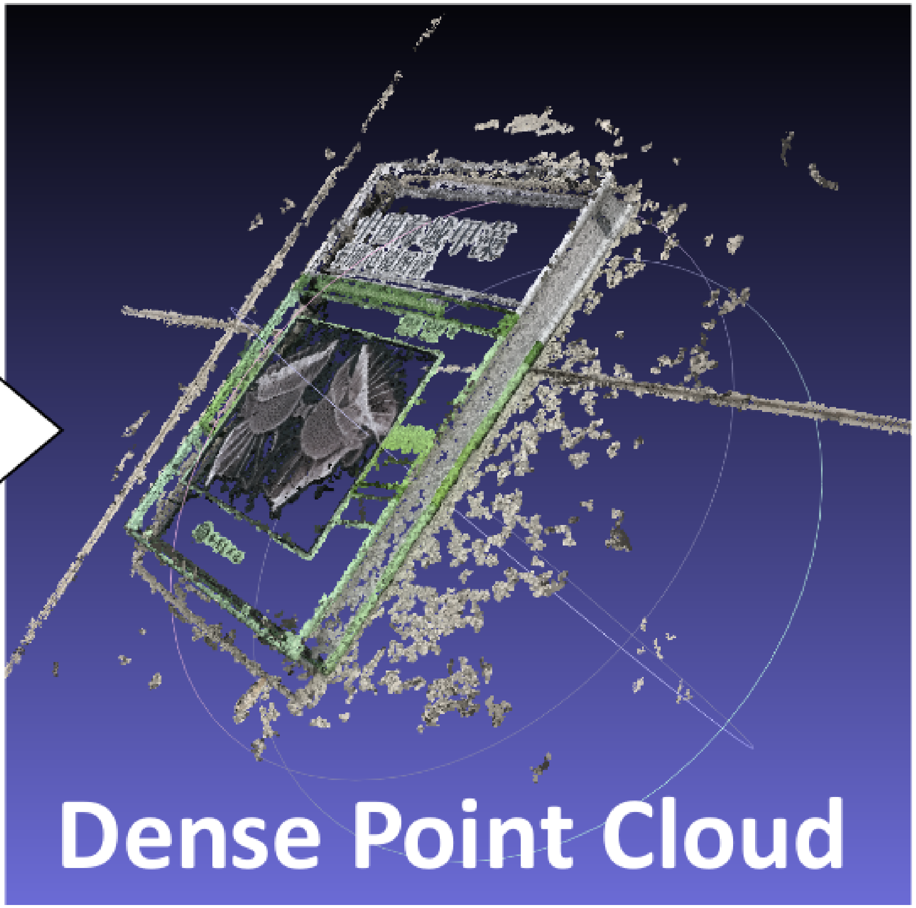
M. Kazhdan, M. Bolitho, and H. Hoppe, “Poisson Surface Reconstruction,” Proc. Symp. Geometry Processing, 2006.

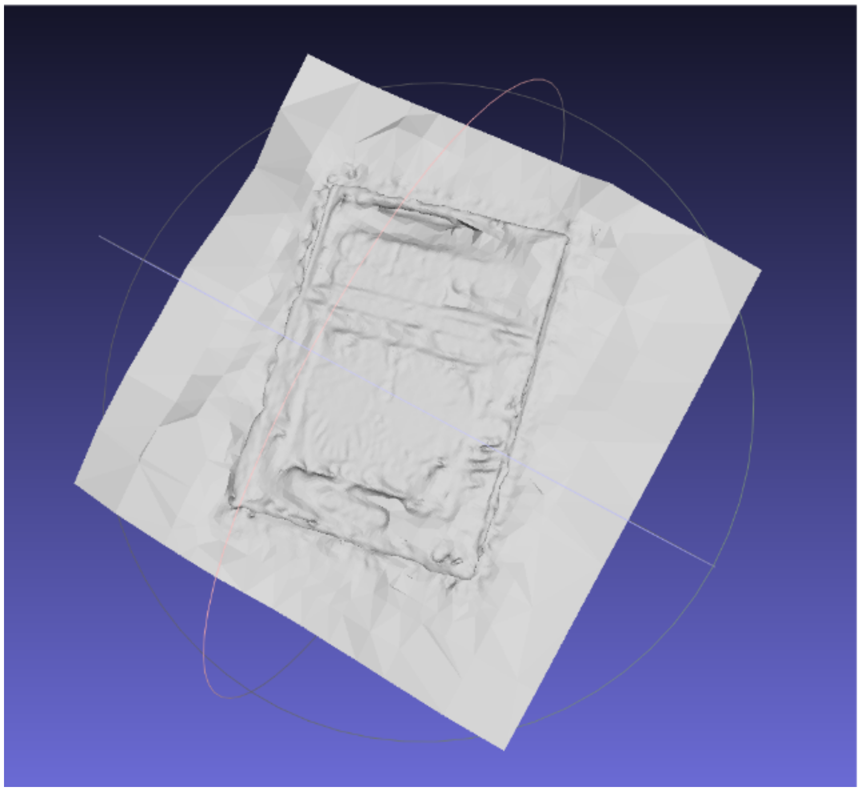
Conclusion

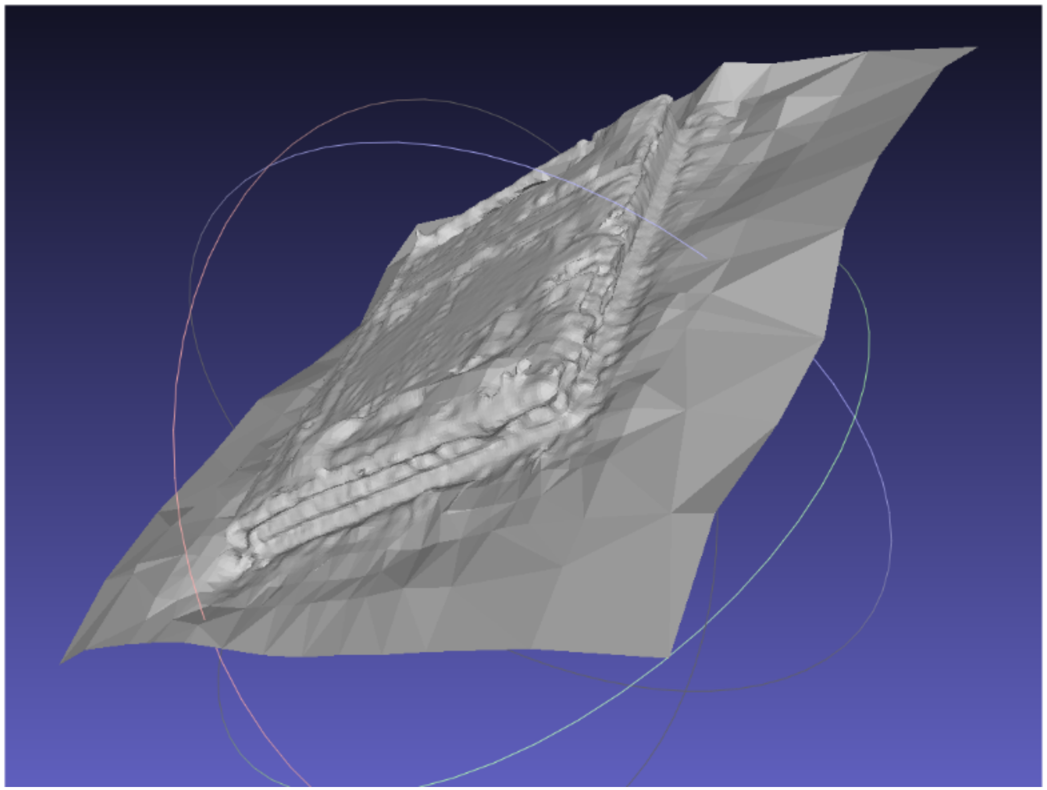

Camera Calibration
Input Images
Feature Detection
Feature Match
Triangulation
Bundle Adjustment
Sparse Point Clouds
Dense Point Clouds
Poisson Surface Reconstruction

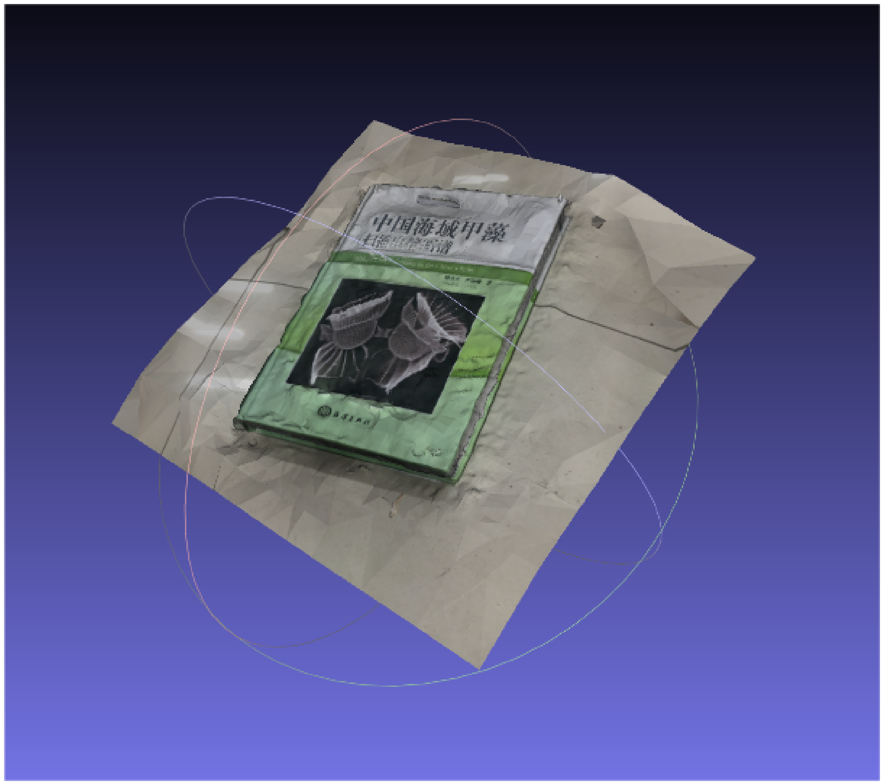

中国海域甲藻
扫描电镜图谱
Atlas of Dinoflagellates in the China's Seas

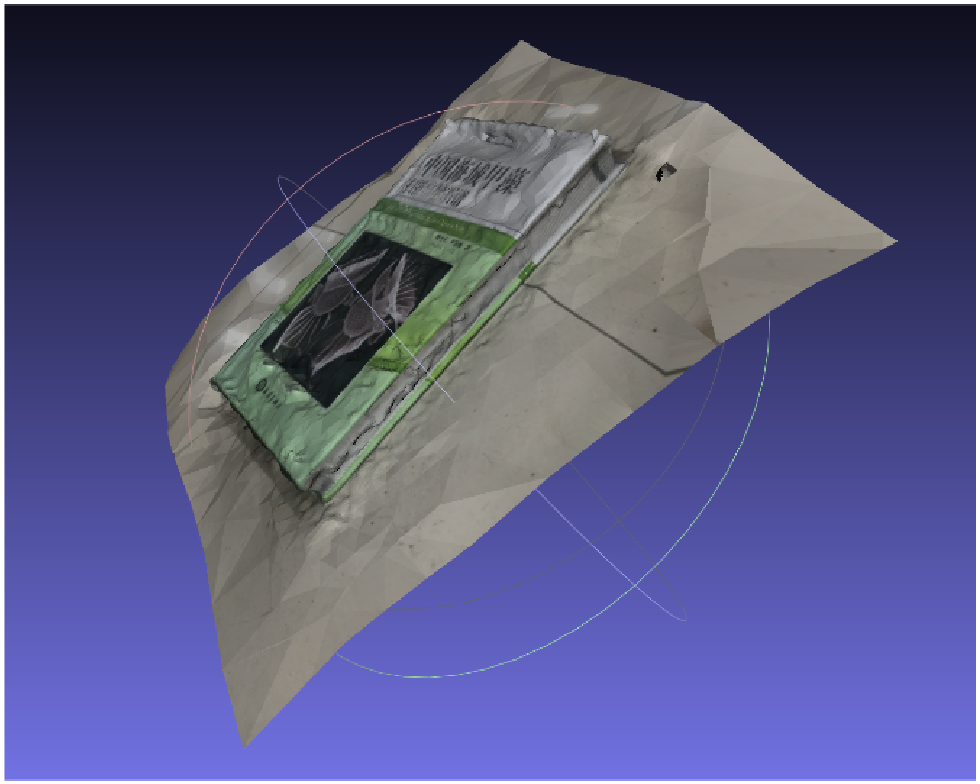

Sparse Point Cloud


Dense Point Cloud

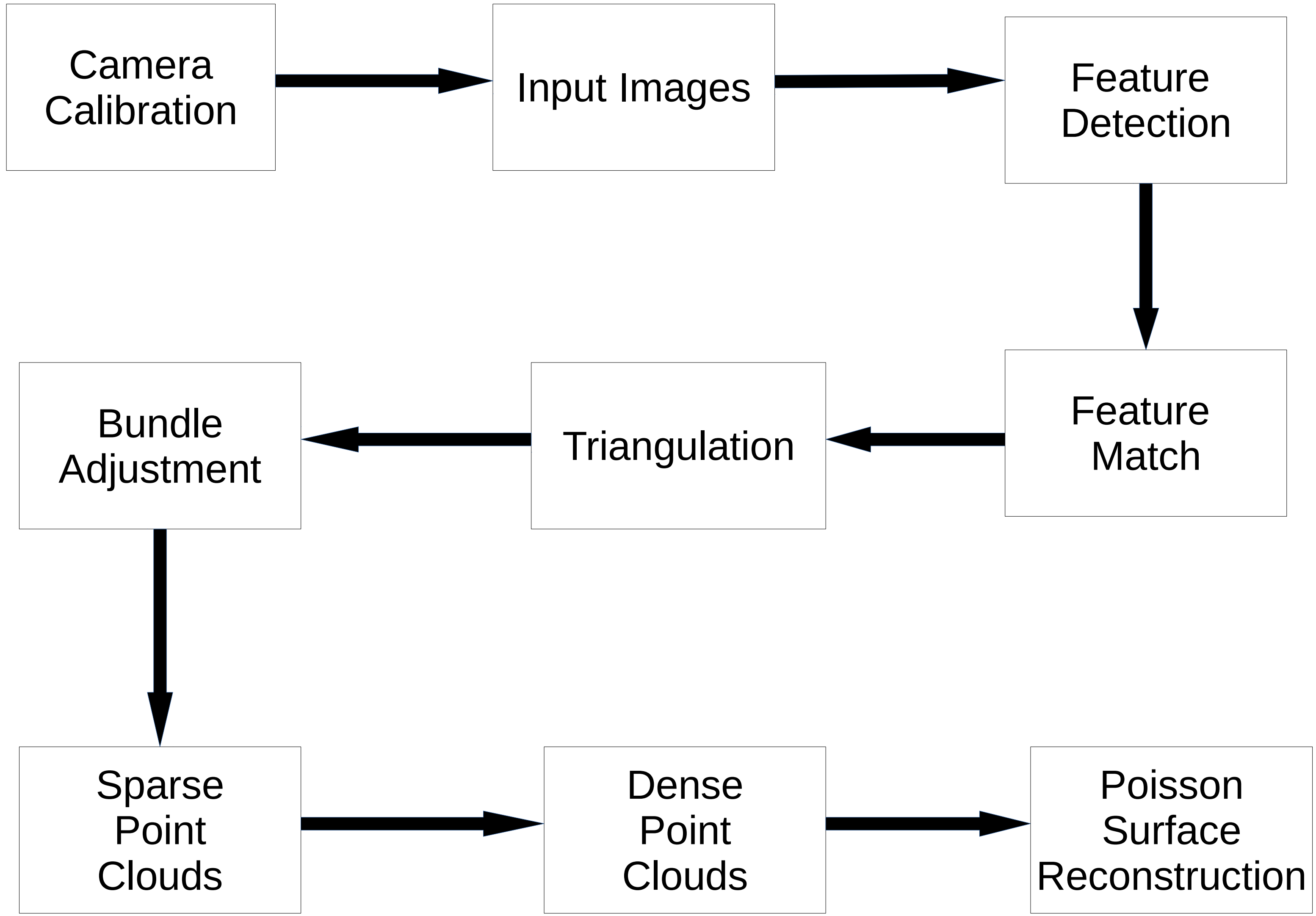




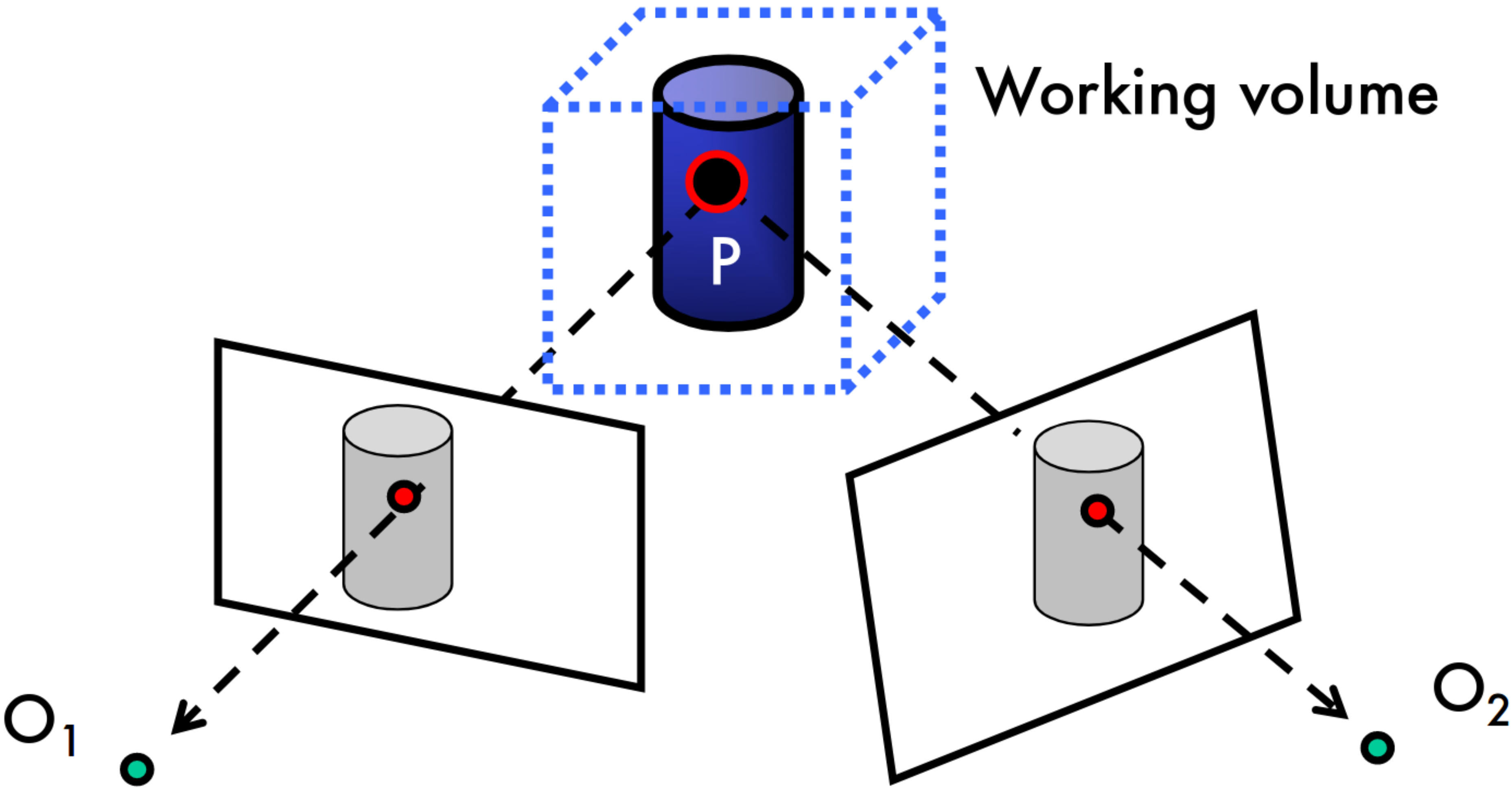

中国海域甲藻

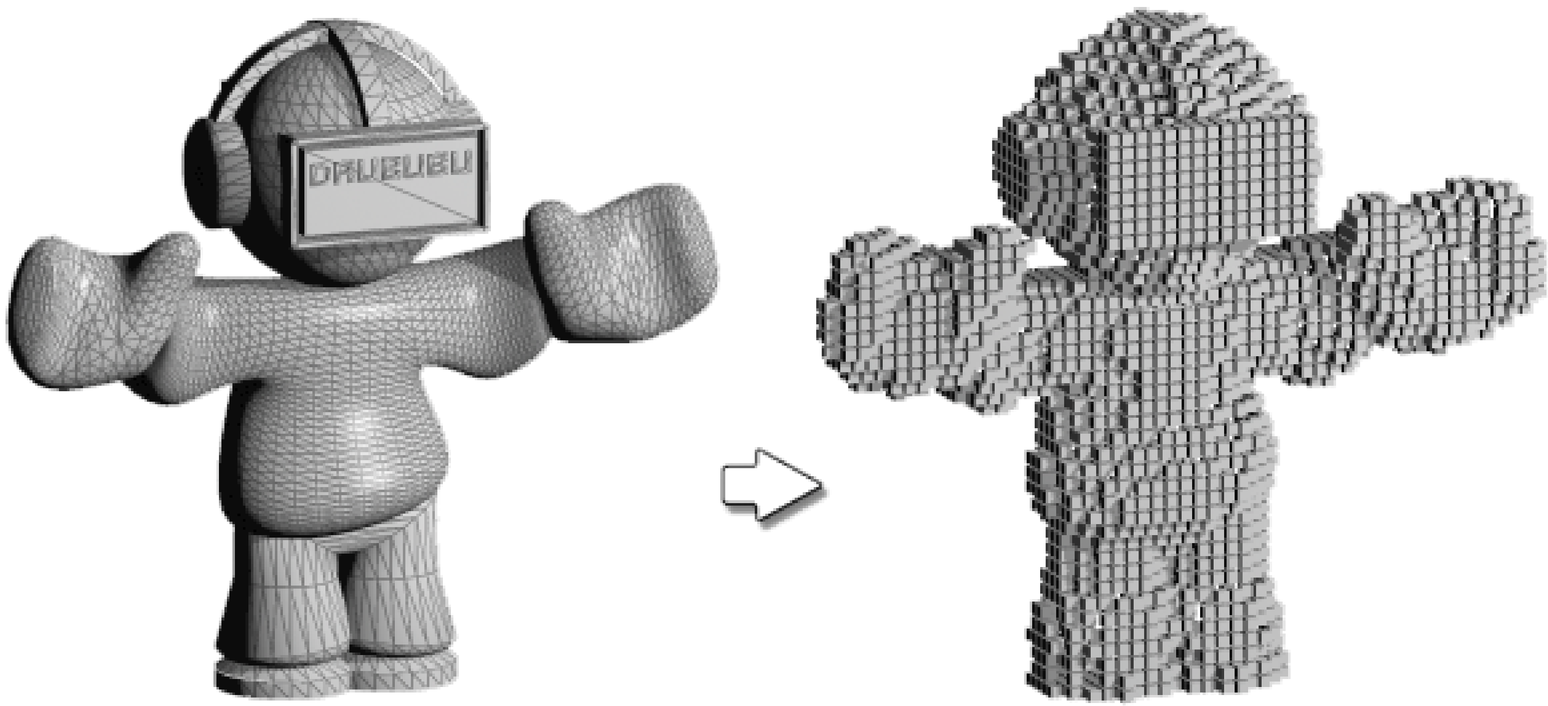

中国海域甲藻

Conclusion

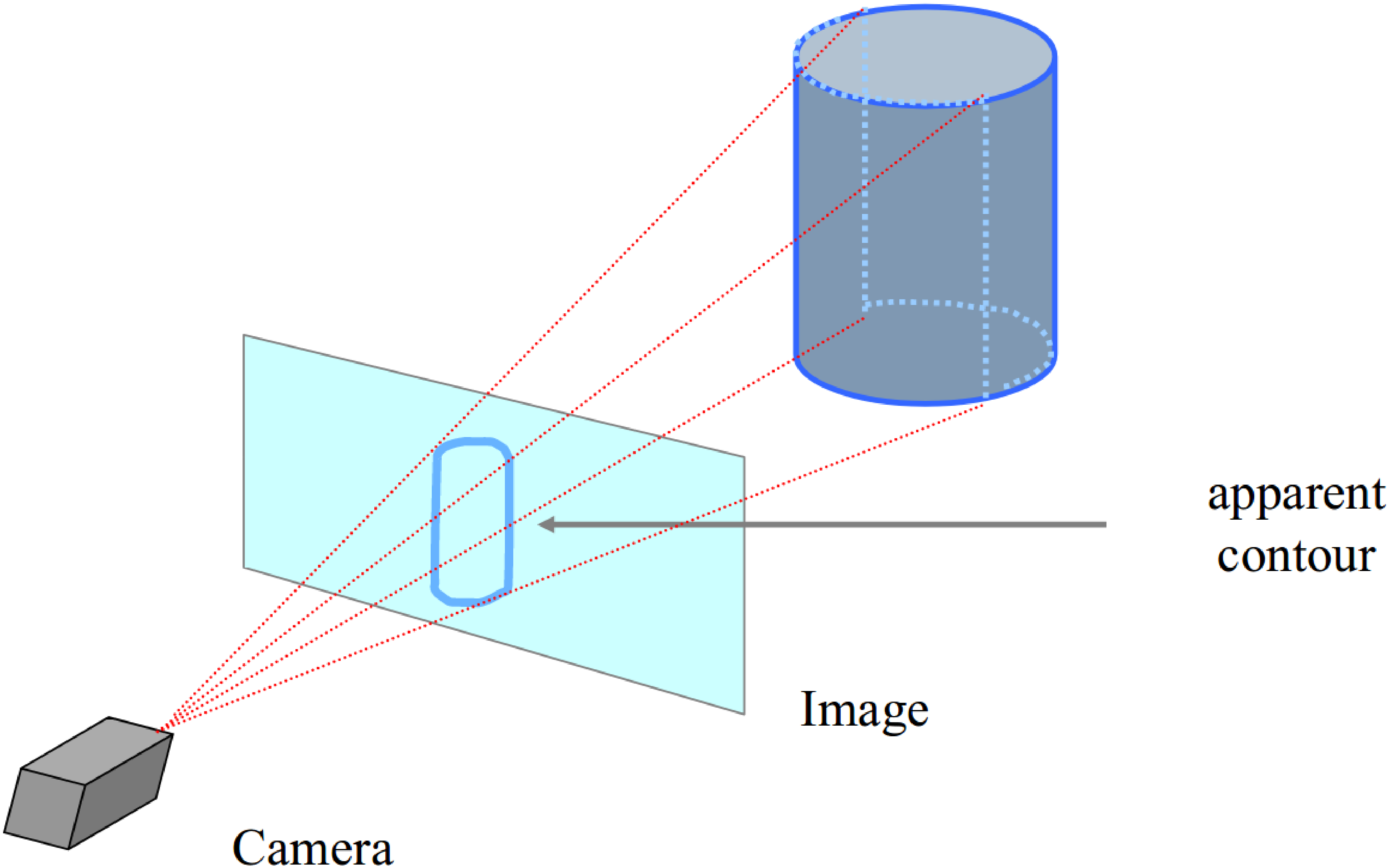

Camera Calibration
Input Images
Feature Detection
Feature Match
Triangulation
Bundle Adjustment
Sparse Point Clouds
Dense Point Clouds
Poisson Surface Reconstruction

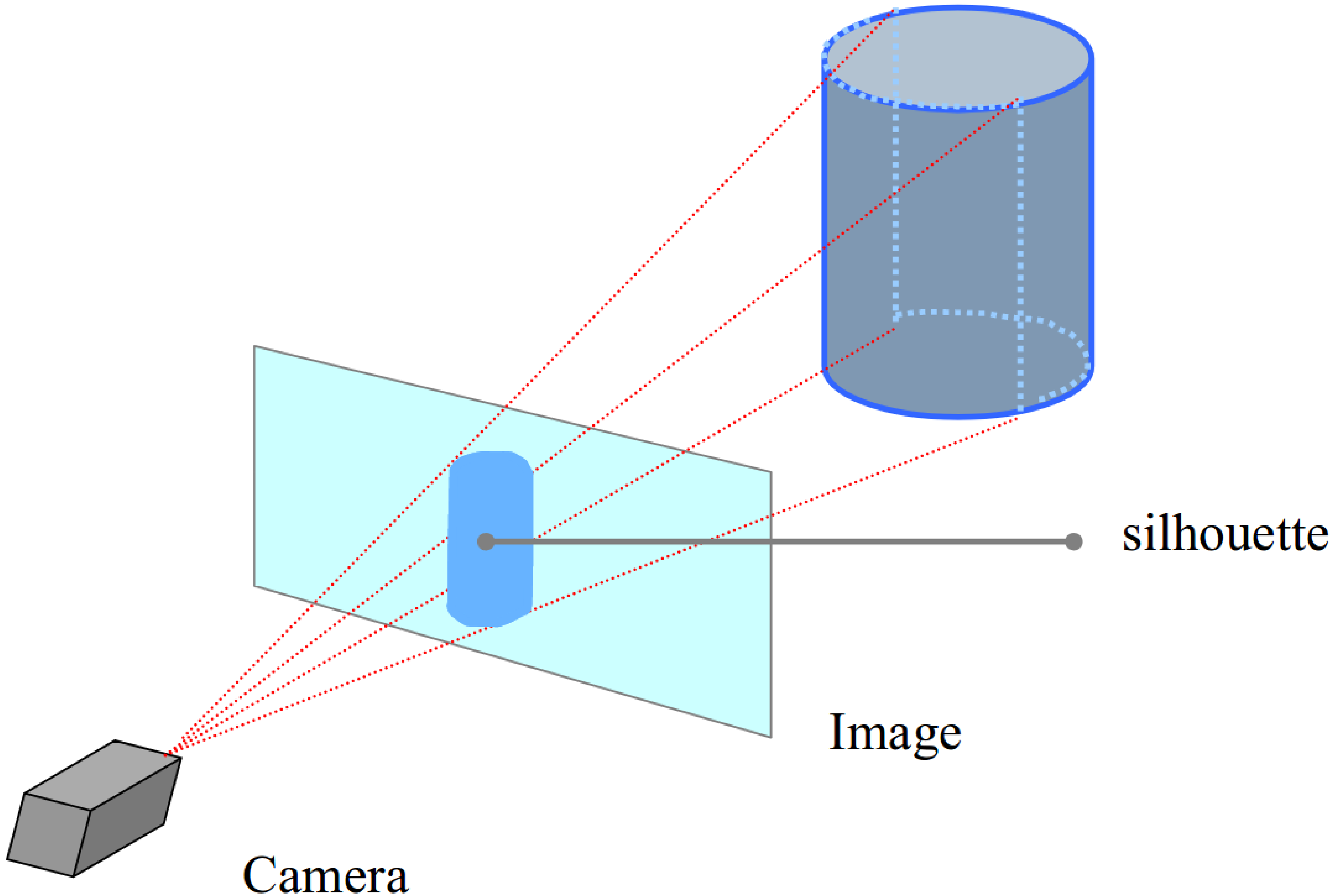
Volumetric Stereo



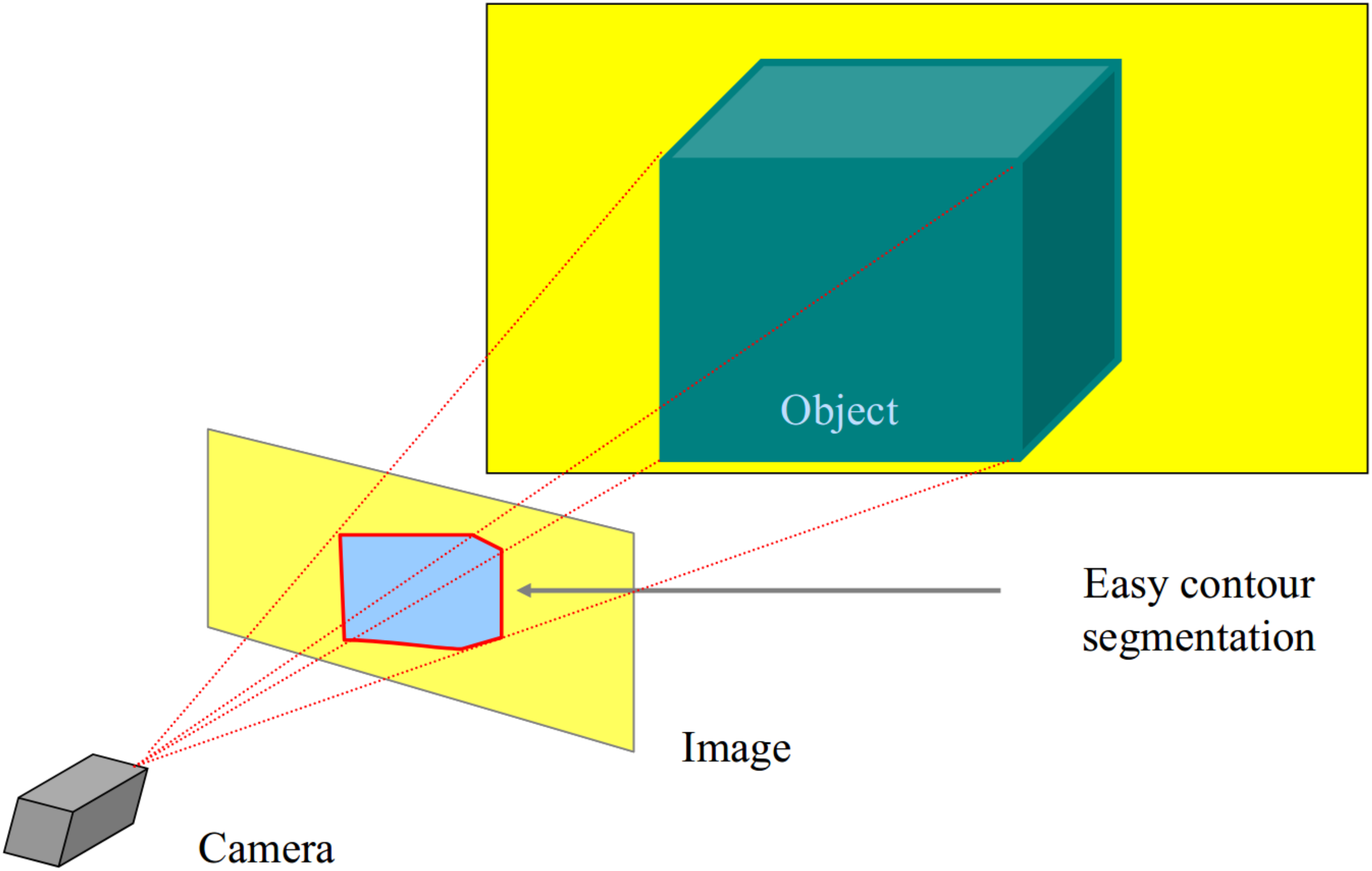


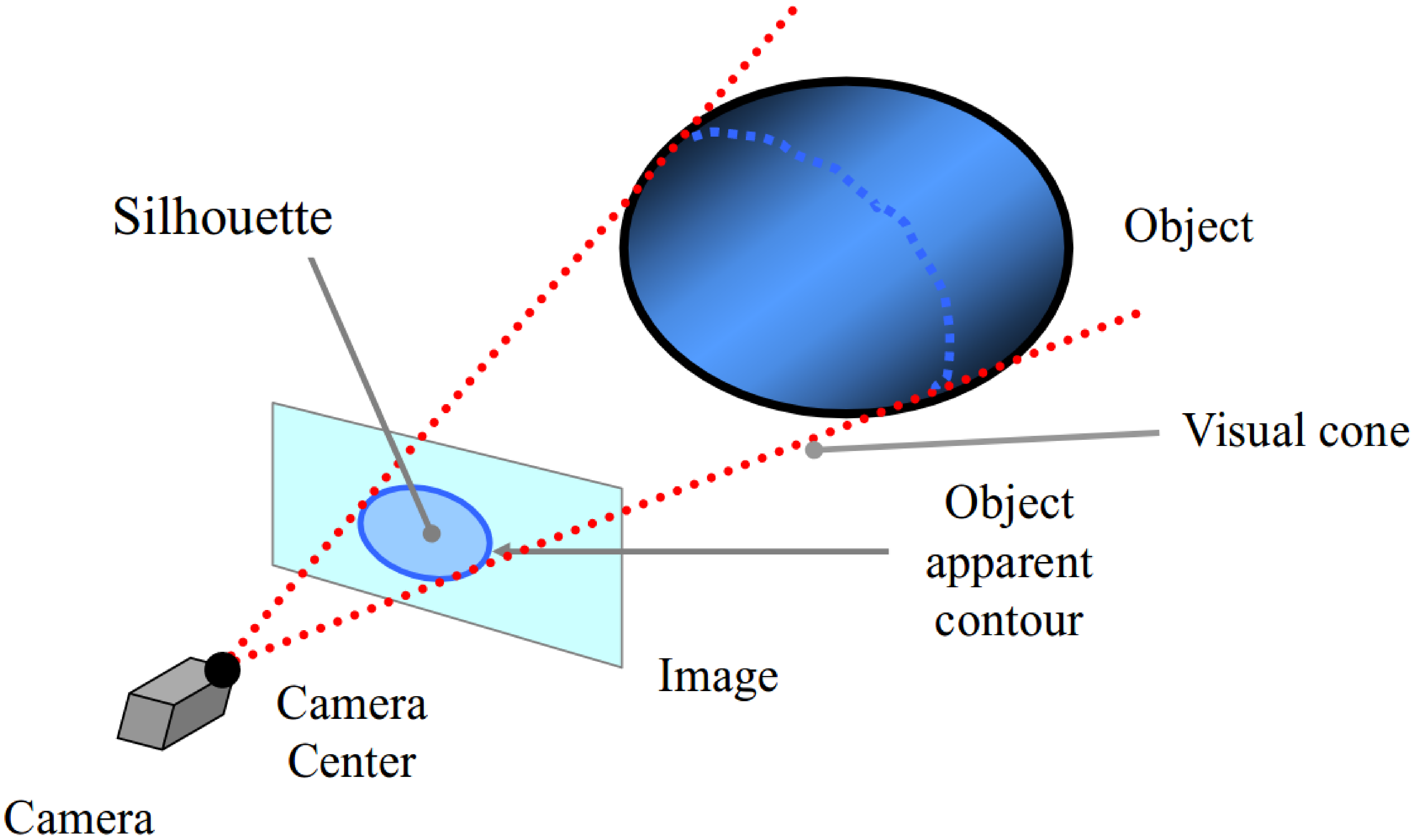
Contours/silhouettes


apparent
contour
Image
Camera

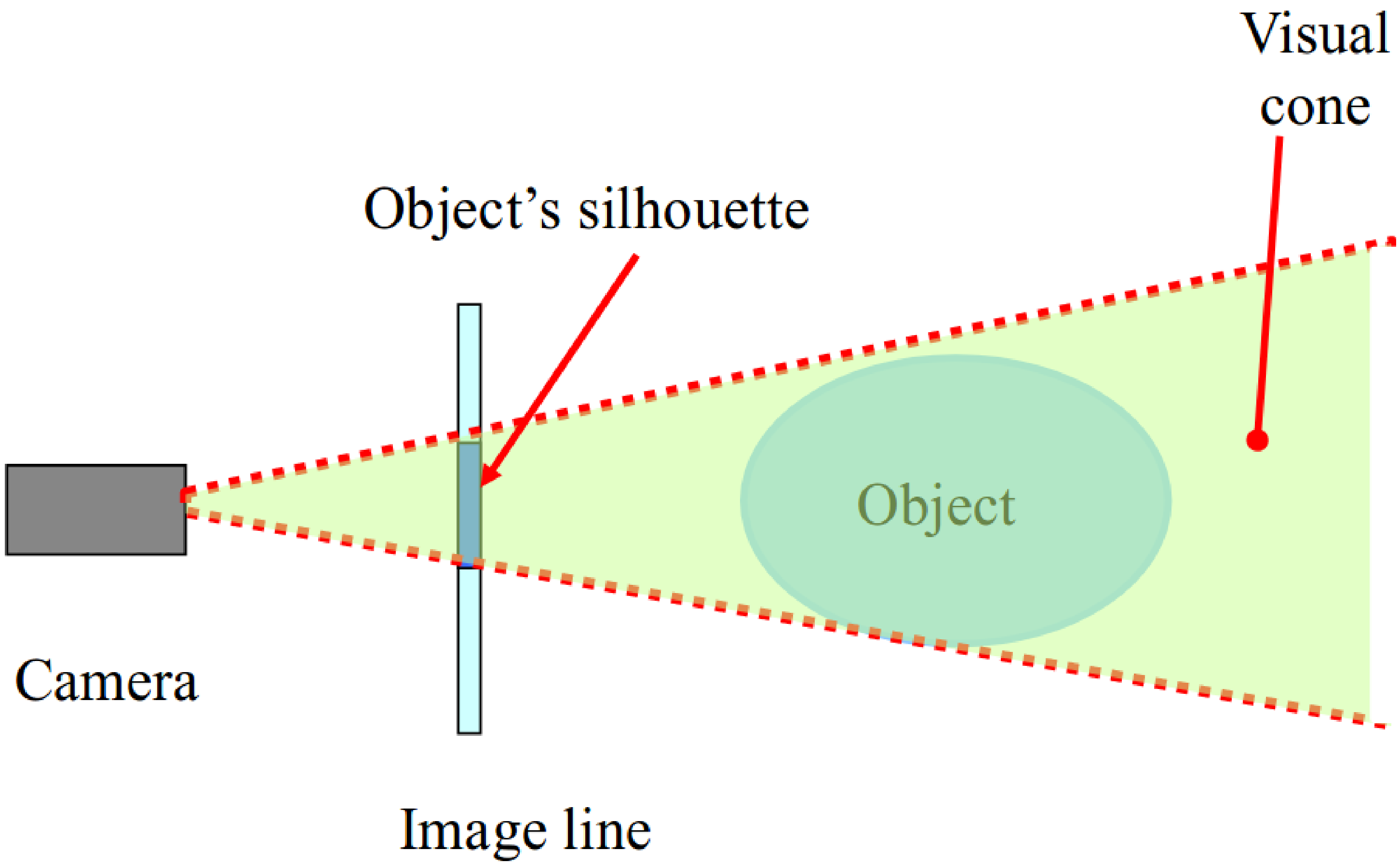

silhouette
Image
Camera

Detecting silhouettes

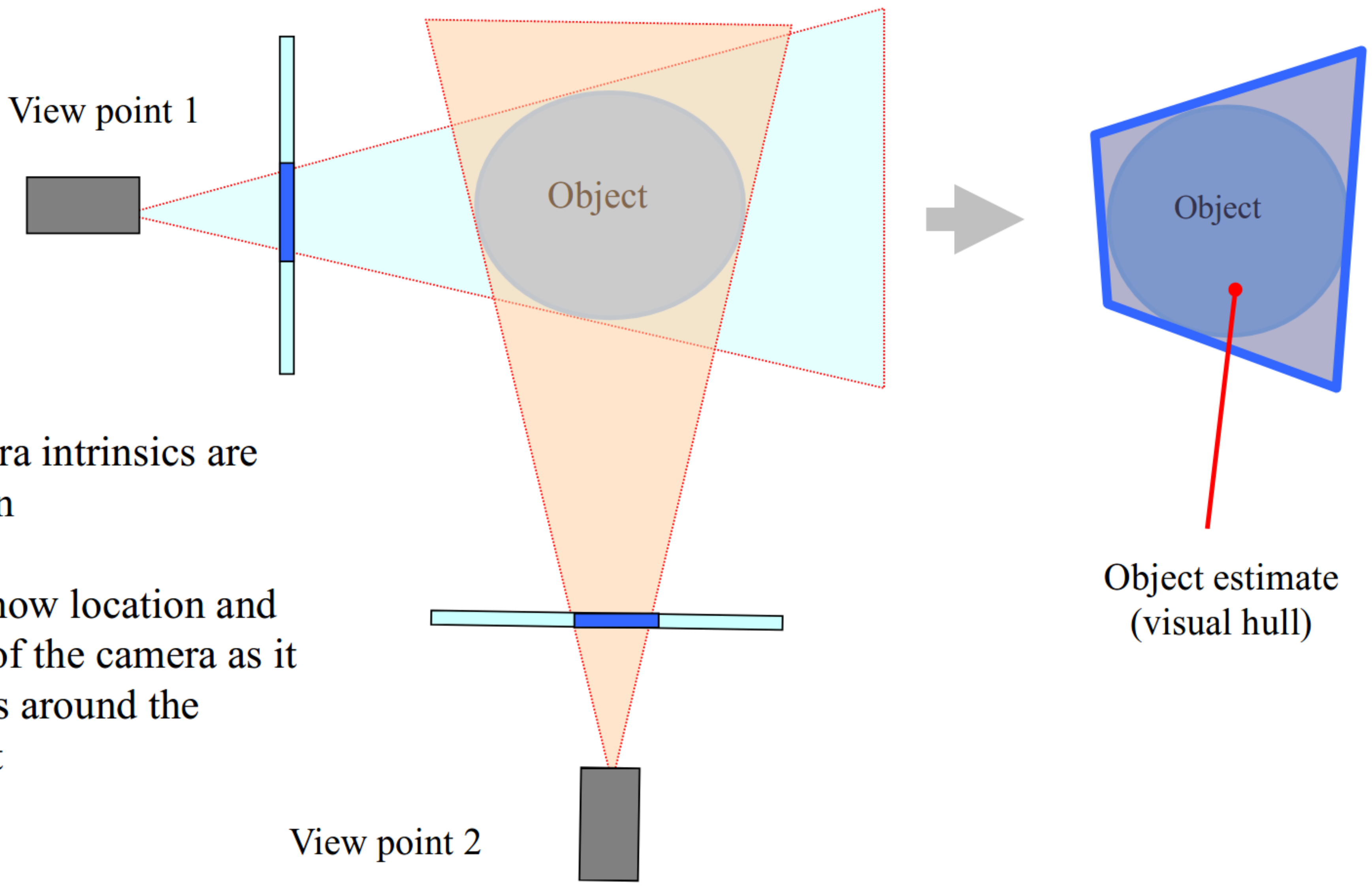



Silhouette
Object
Visual cone
Object
apparent
contour
Image
Camera
Center
Camera

2D slice:



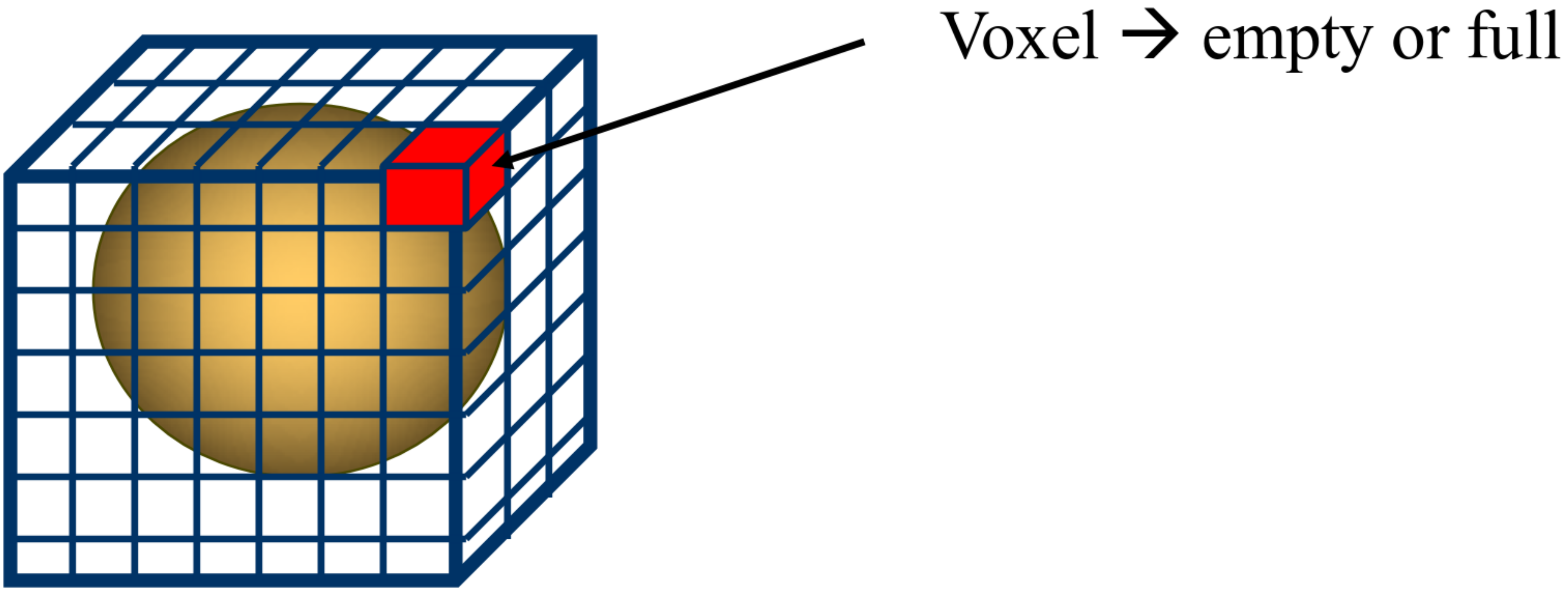
Space Carving



- Camera intrinsics are known
- We know location and pose of the camera as it moves around the object

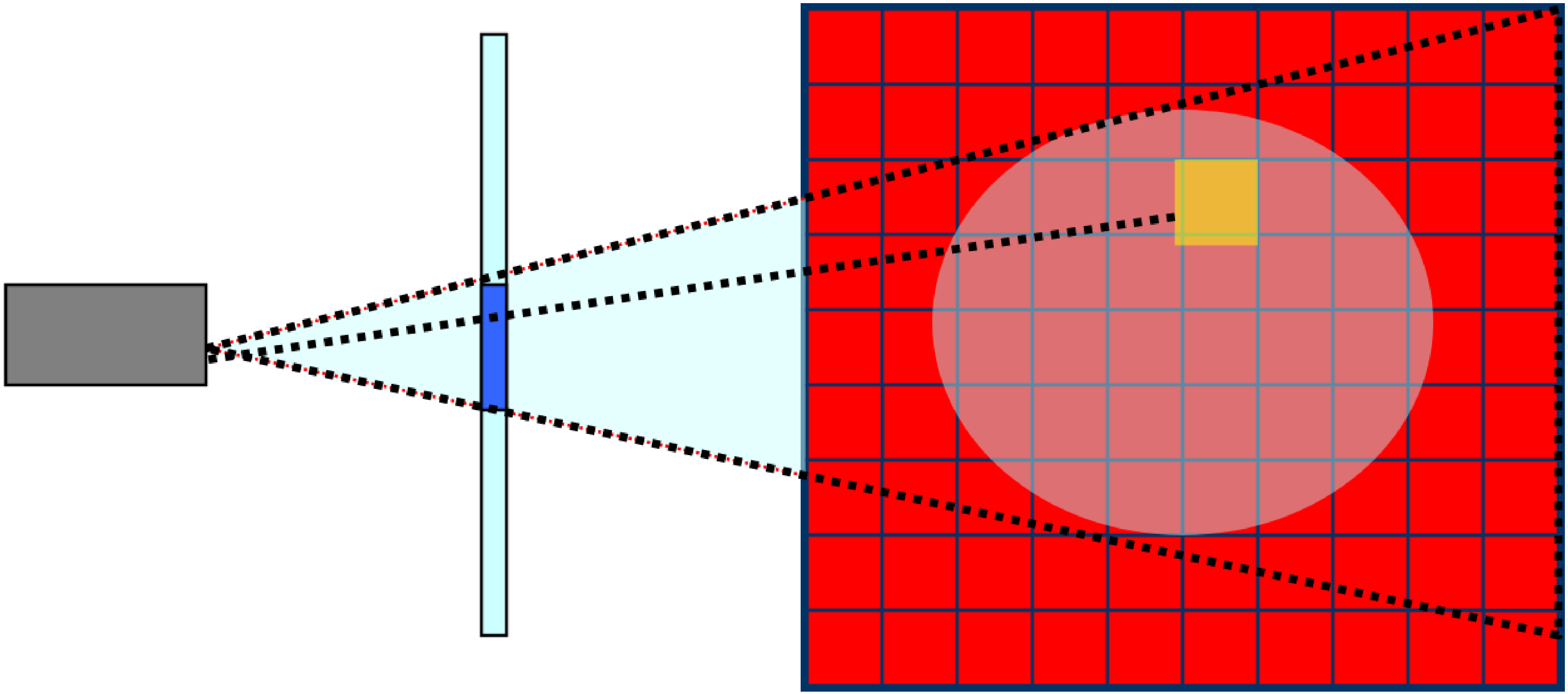
Martin and Aggarwal (1983)

Space Carving

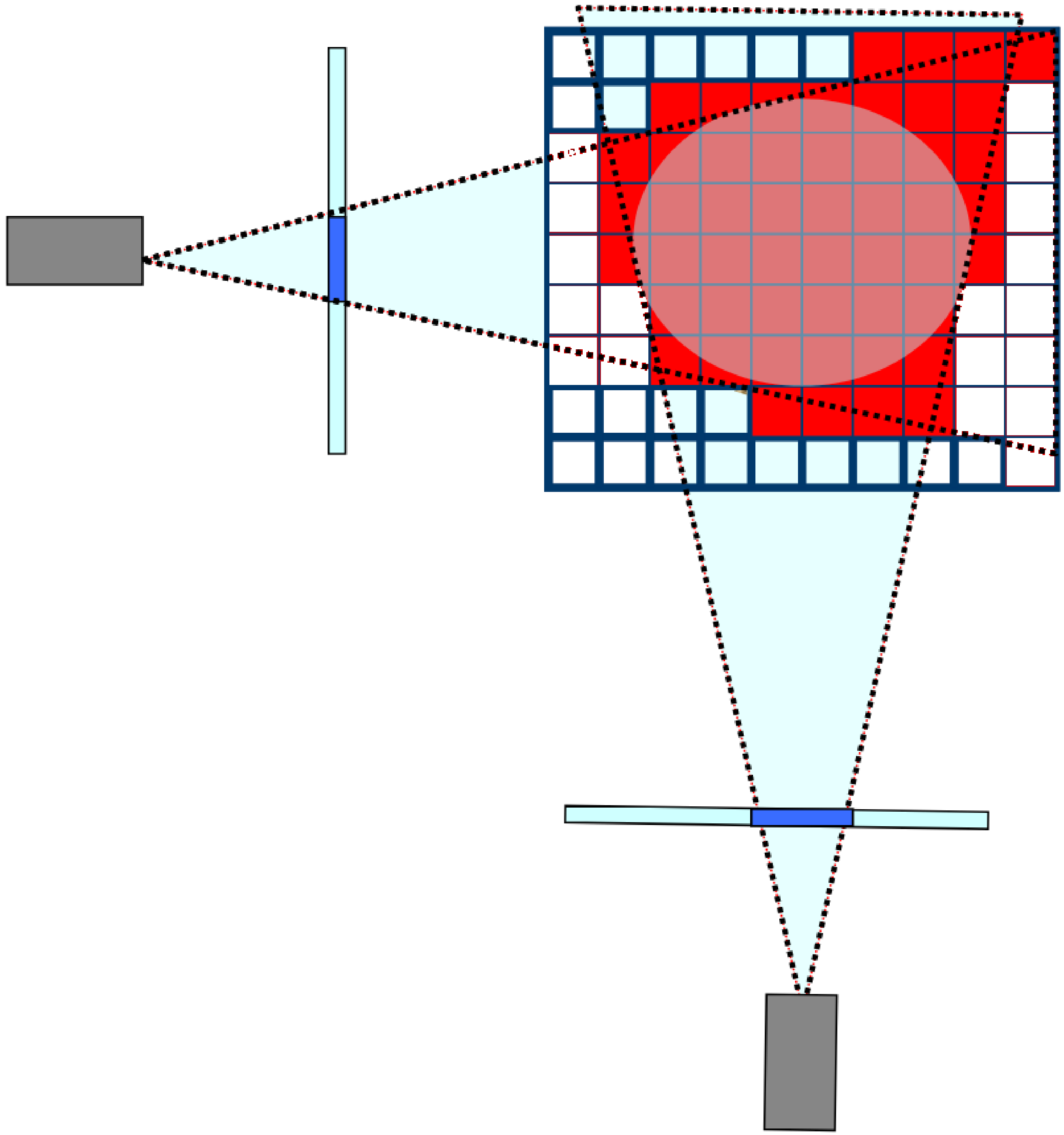


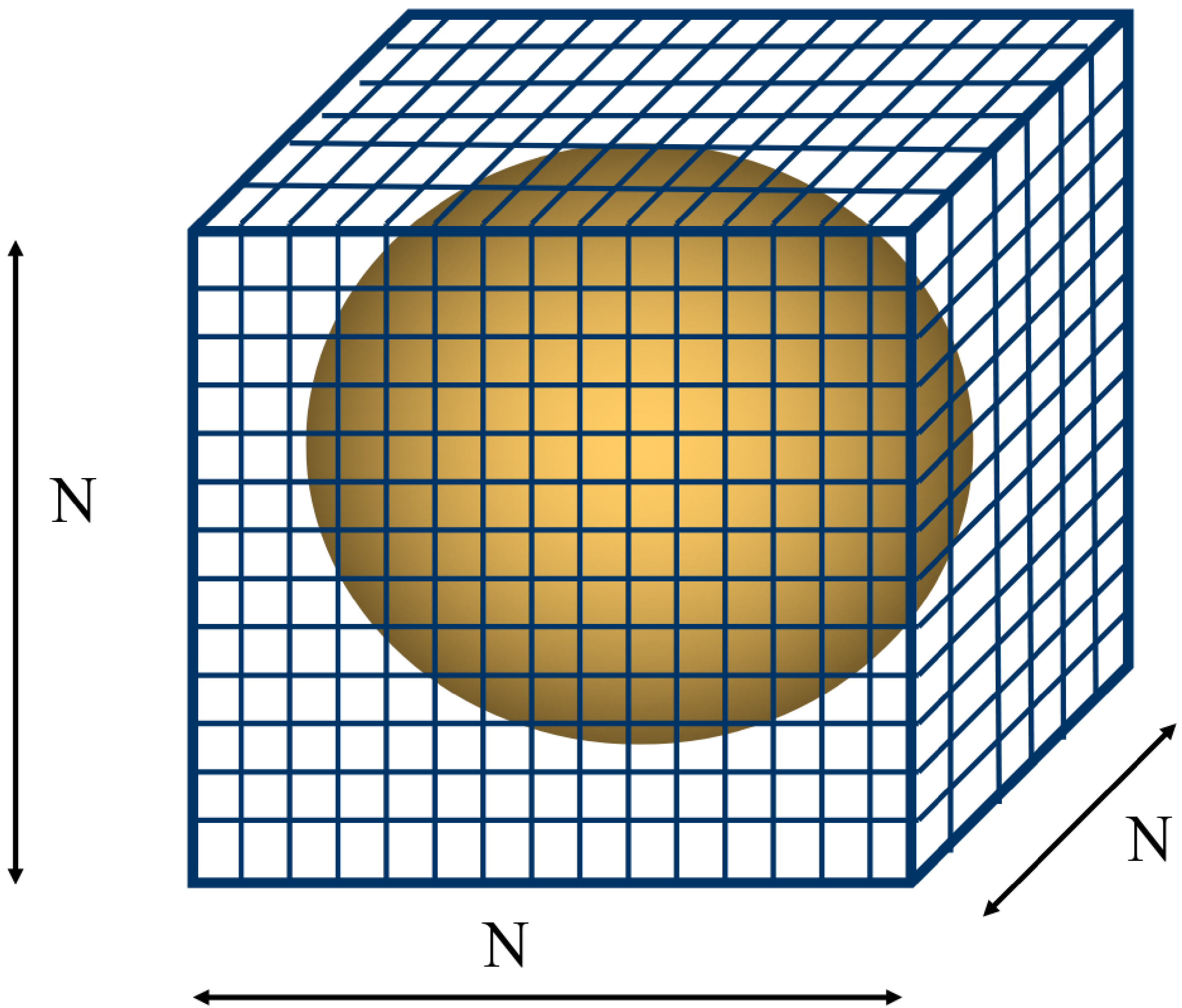
Martin and Aggarwal (1983)

Space Carving

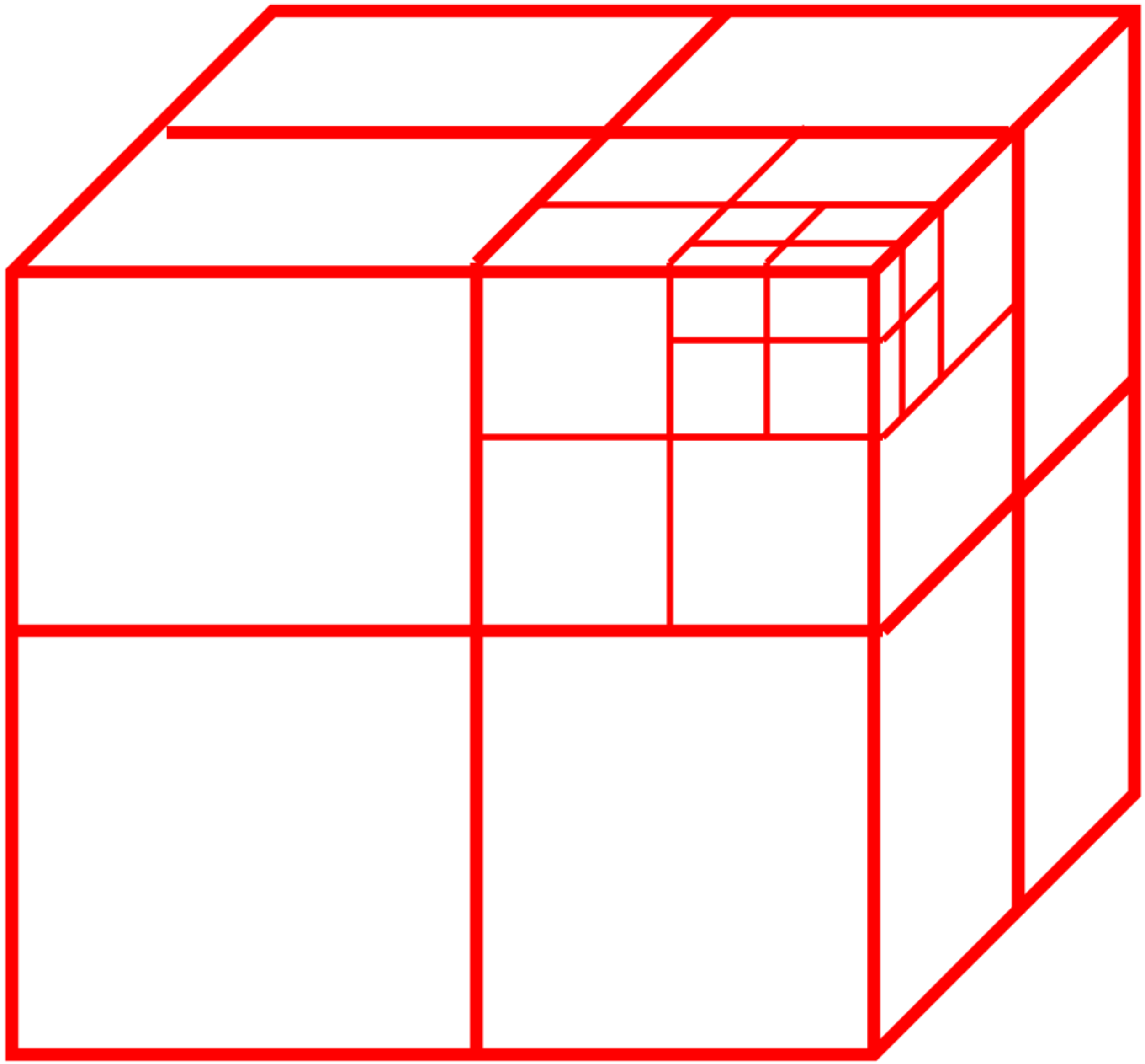


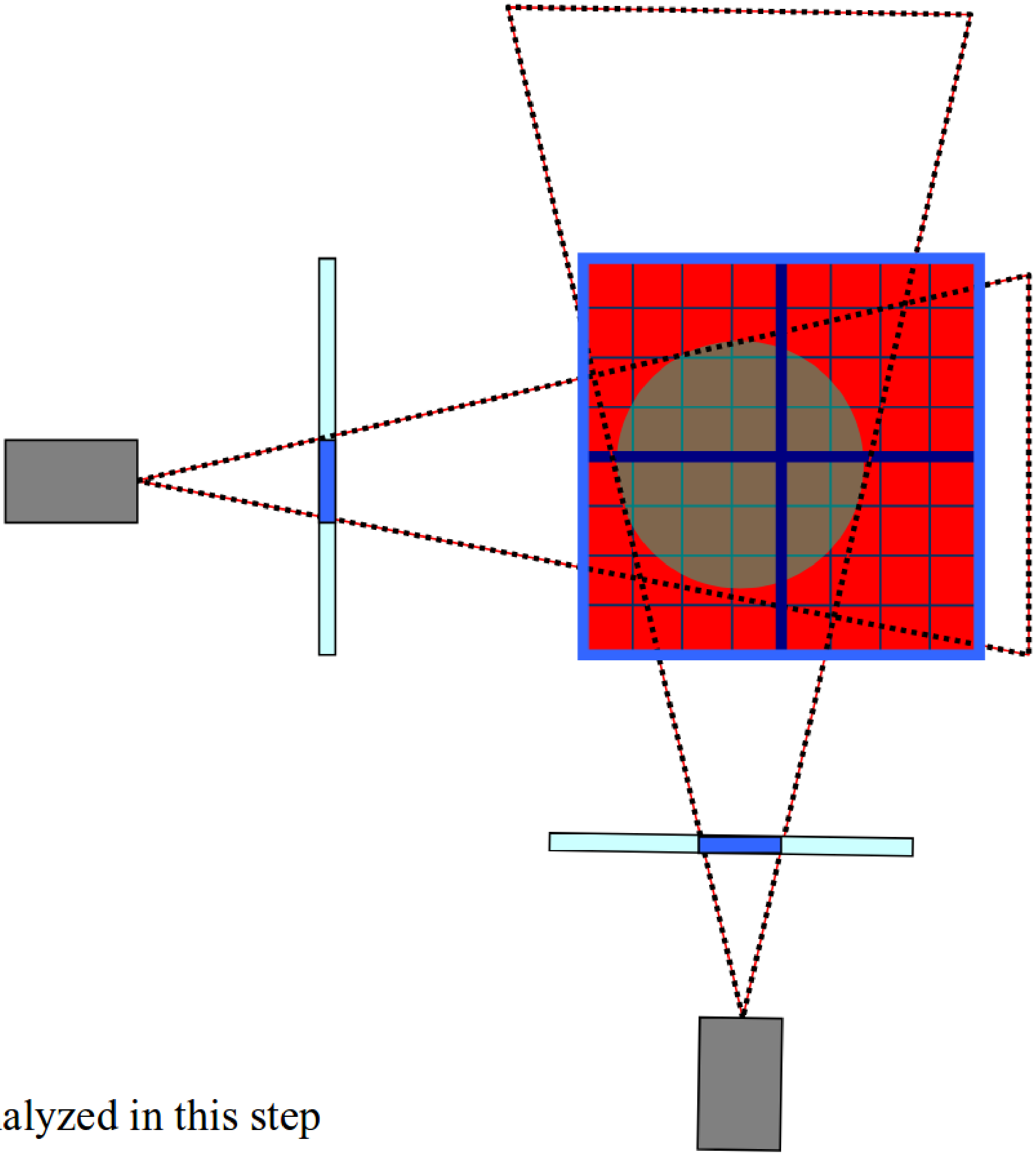
Martin and Aggarwal (1983)



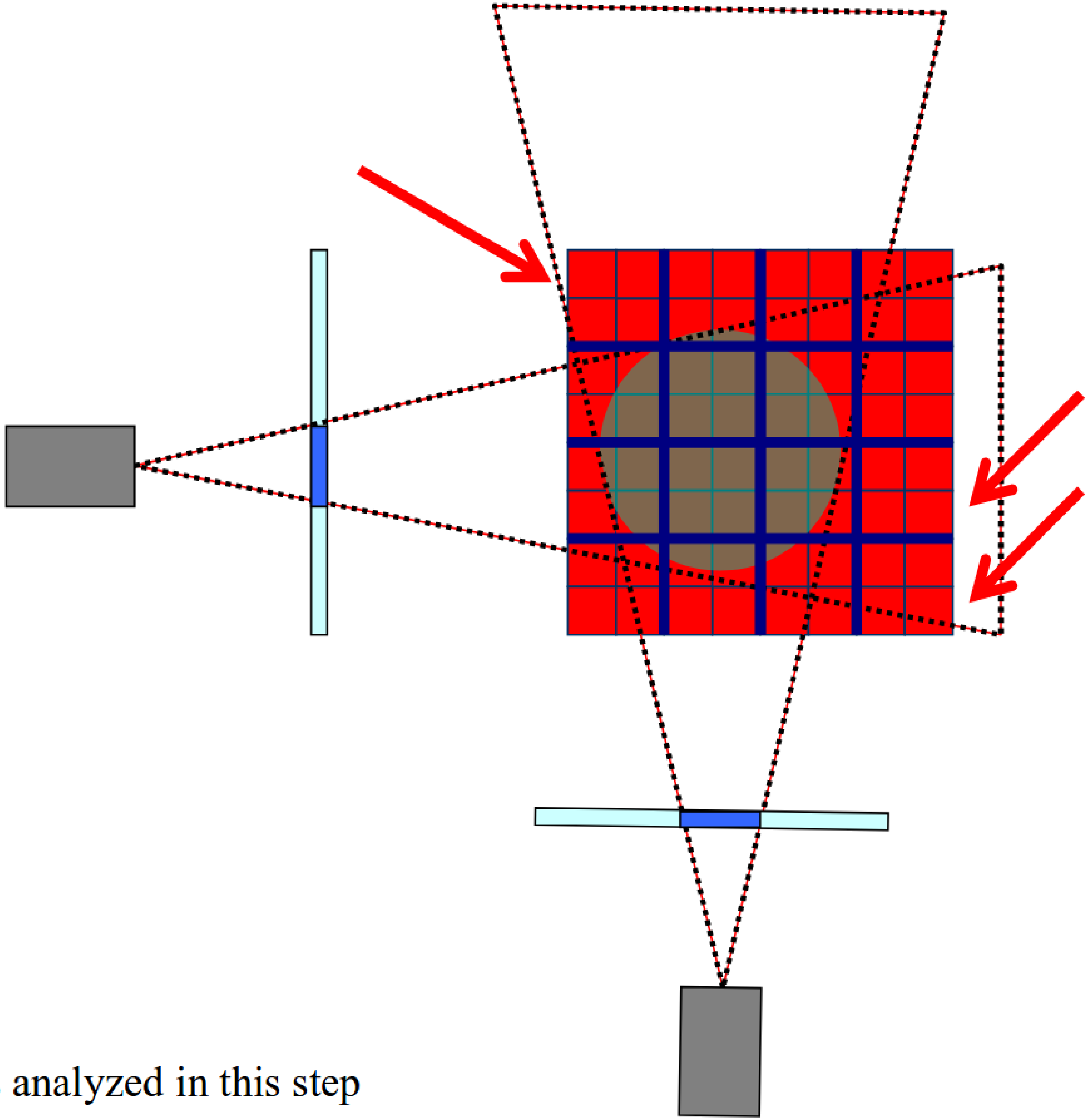

N
N
N

Complexity Reduction: Octrees

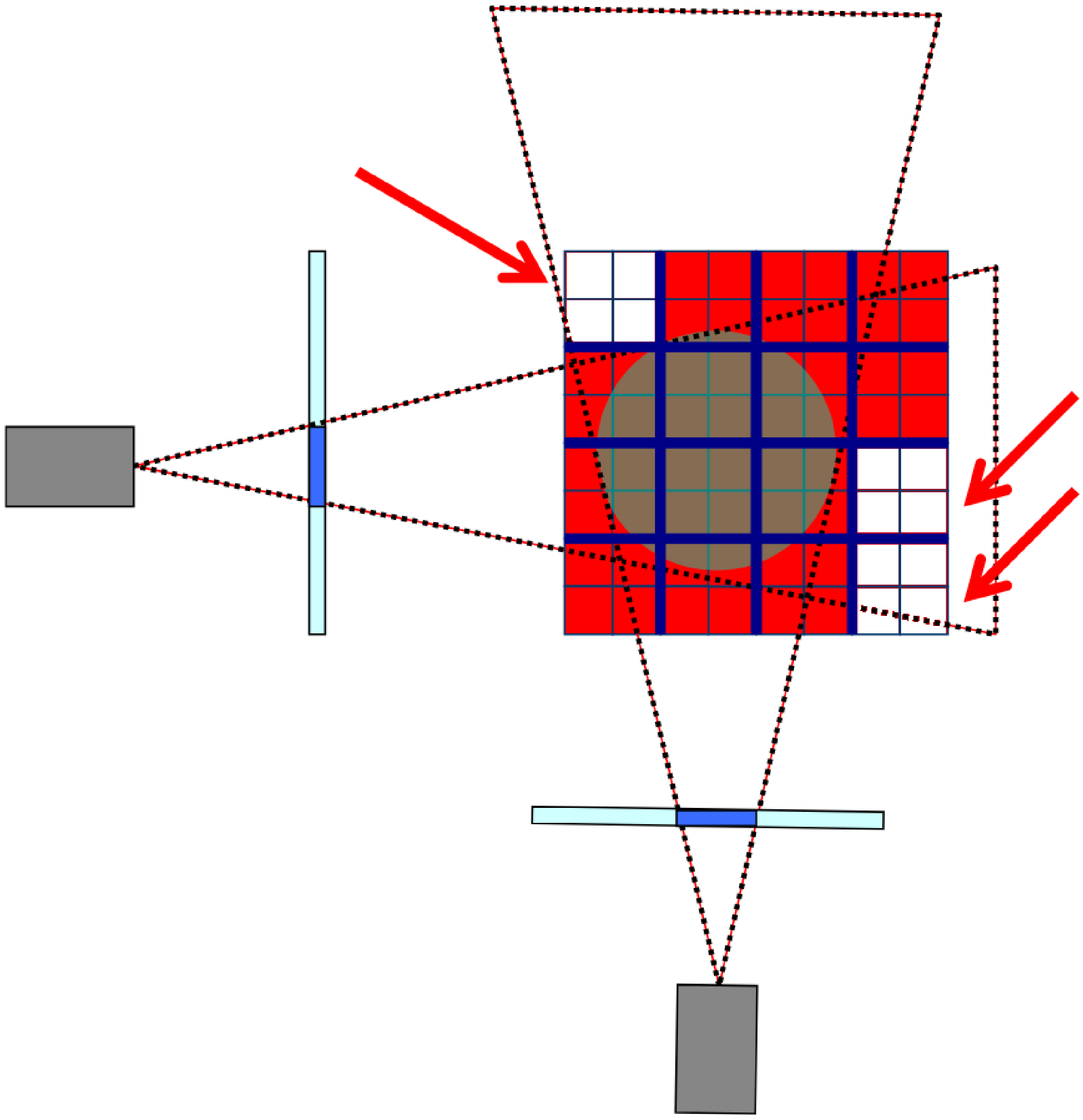


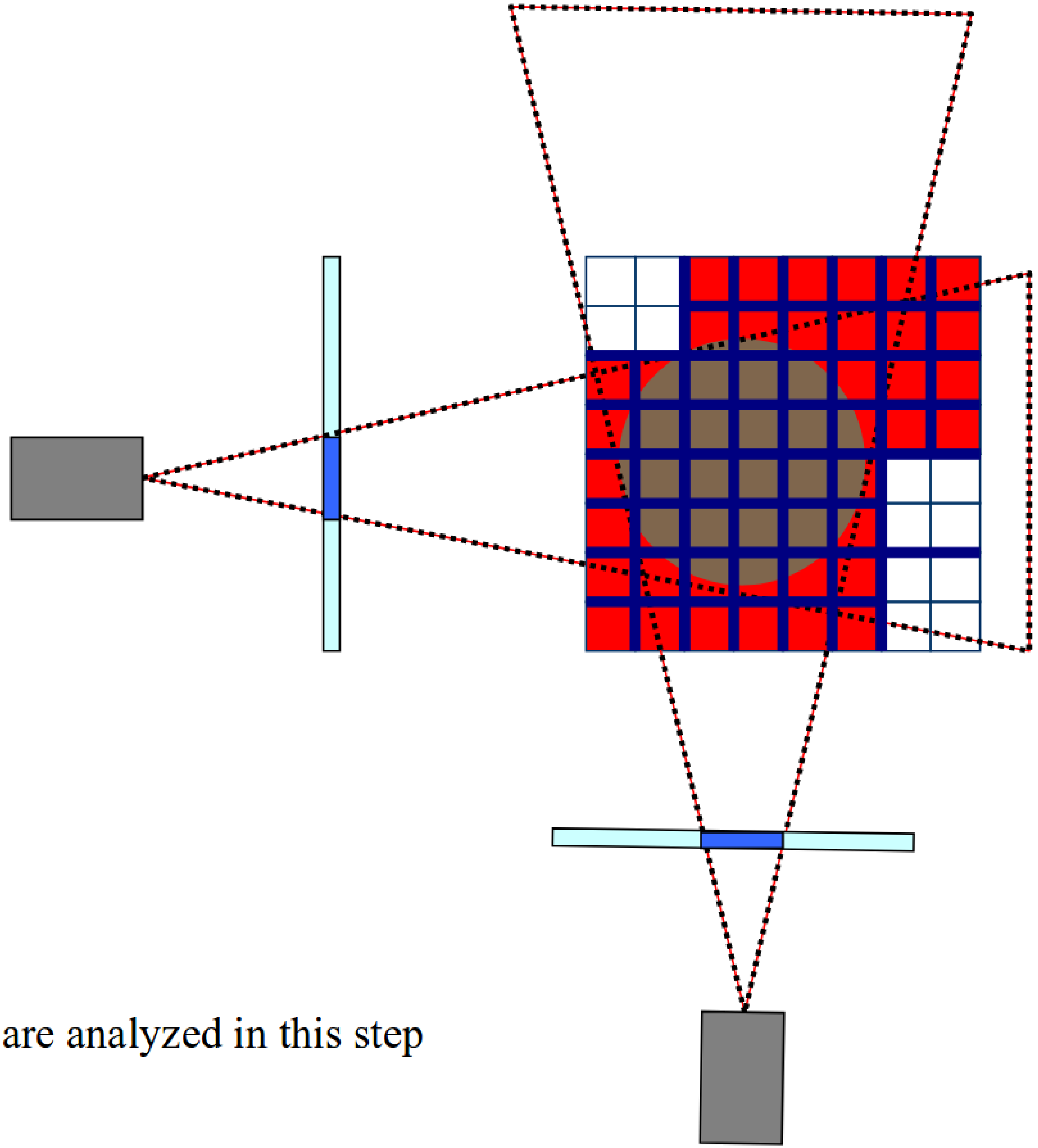


4 elements analyzed in this step

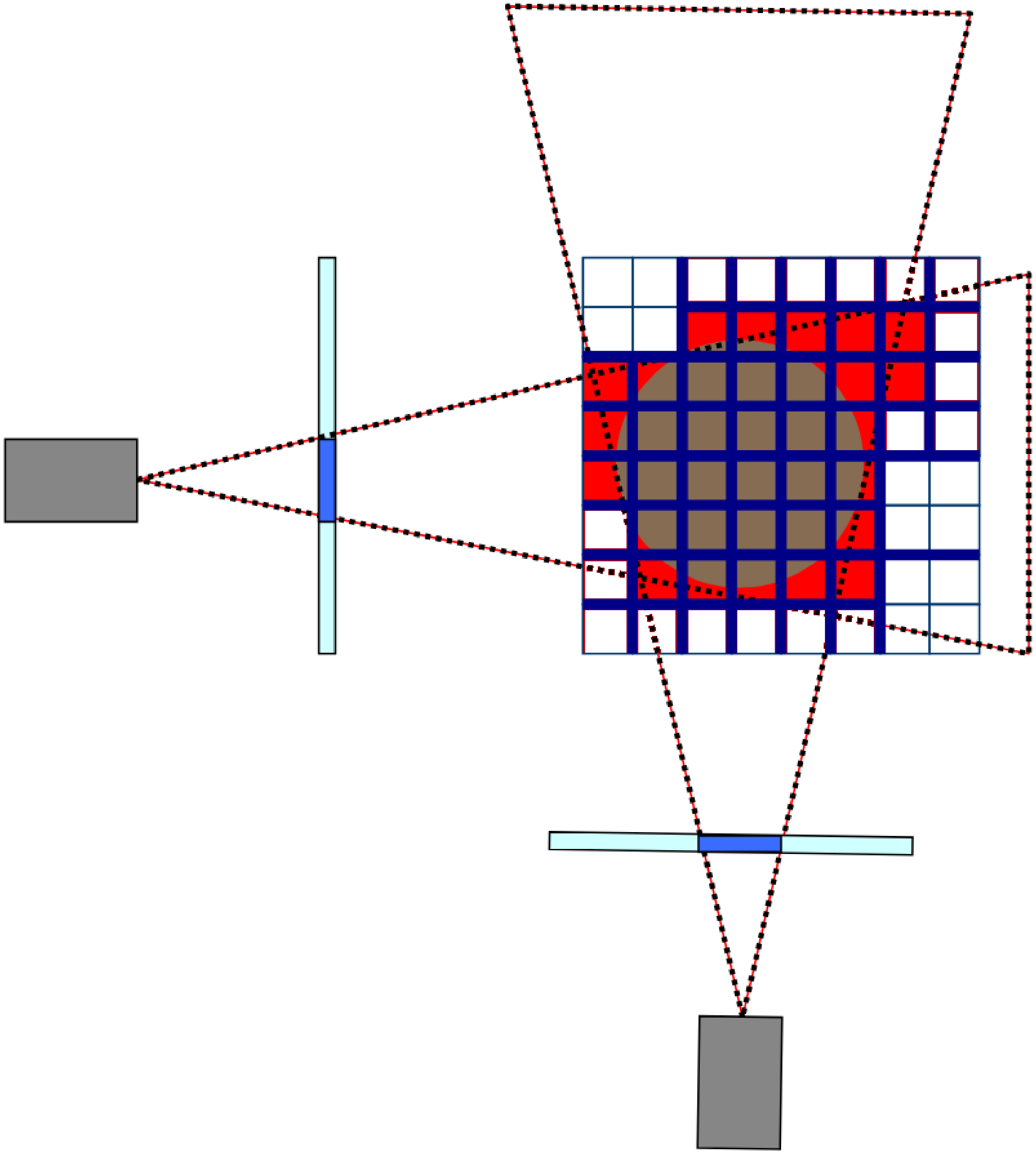


16 elements analyzed in this step





52 elements are analyzed in this step

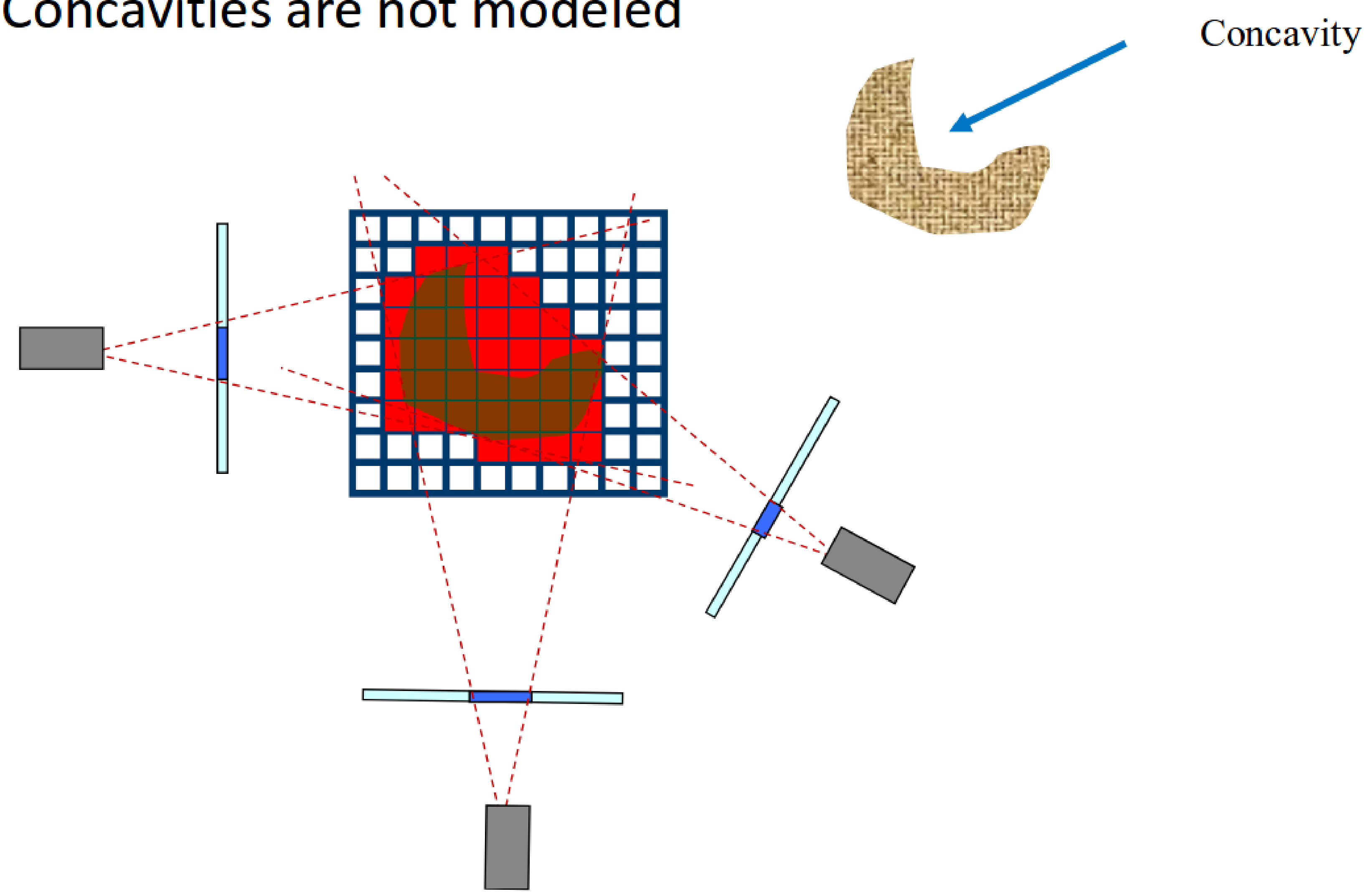


Advantages of Space carving

- Robust and simple
- No need to solve for correspondences

Limitations of Space carving

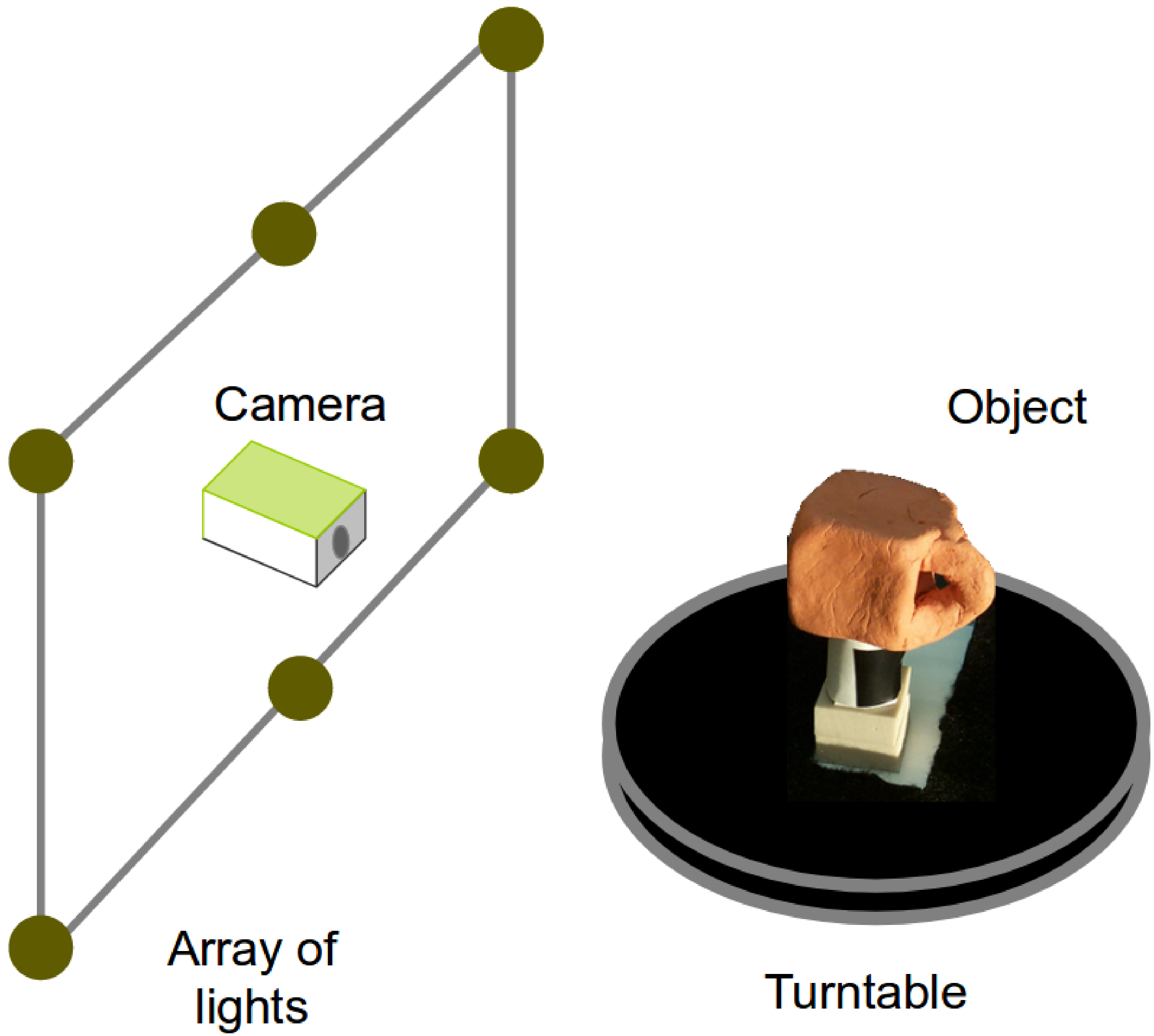
- Concavities are not modeled



Shape from Shadows

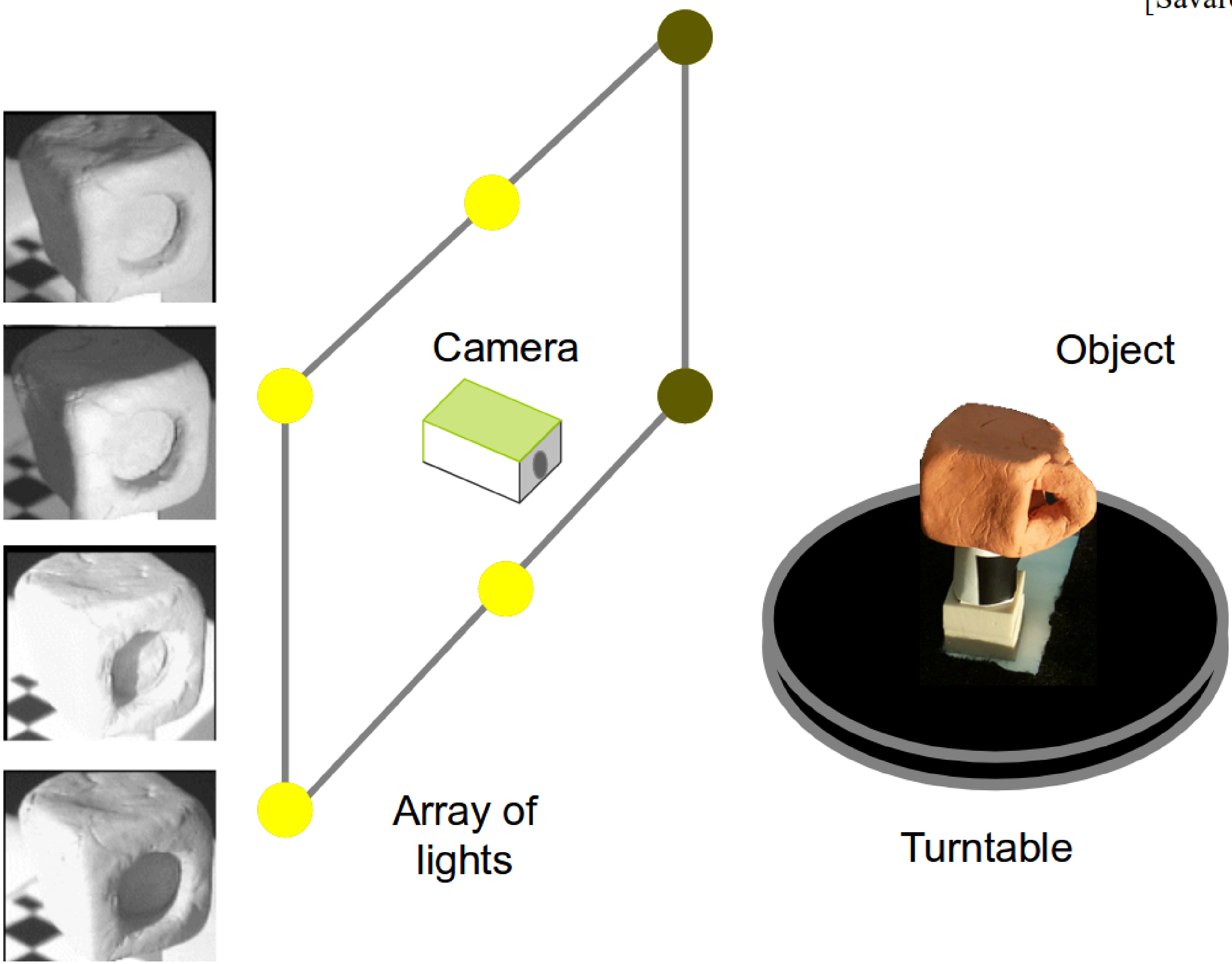
Self-shadows indicate concavities
(no modeled by contours)

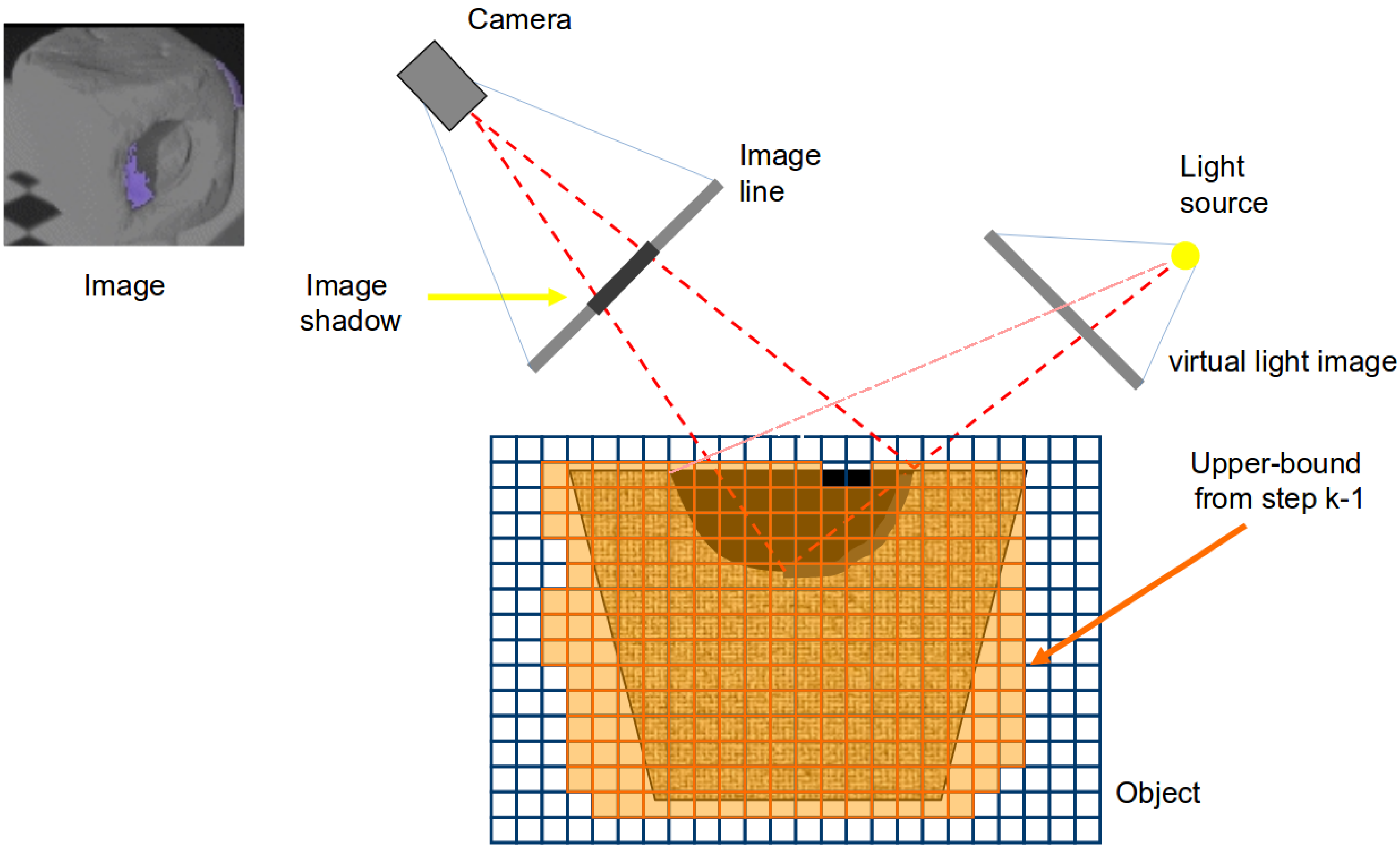
Shadow Carving

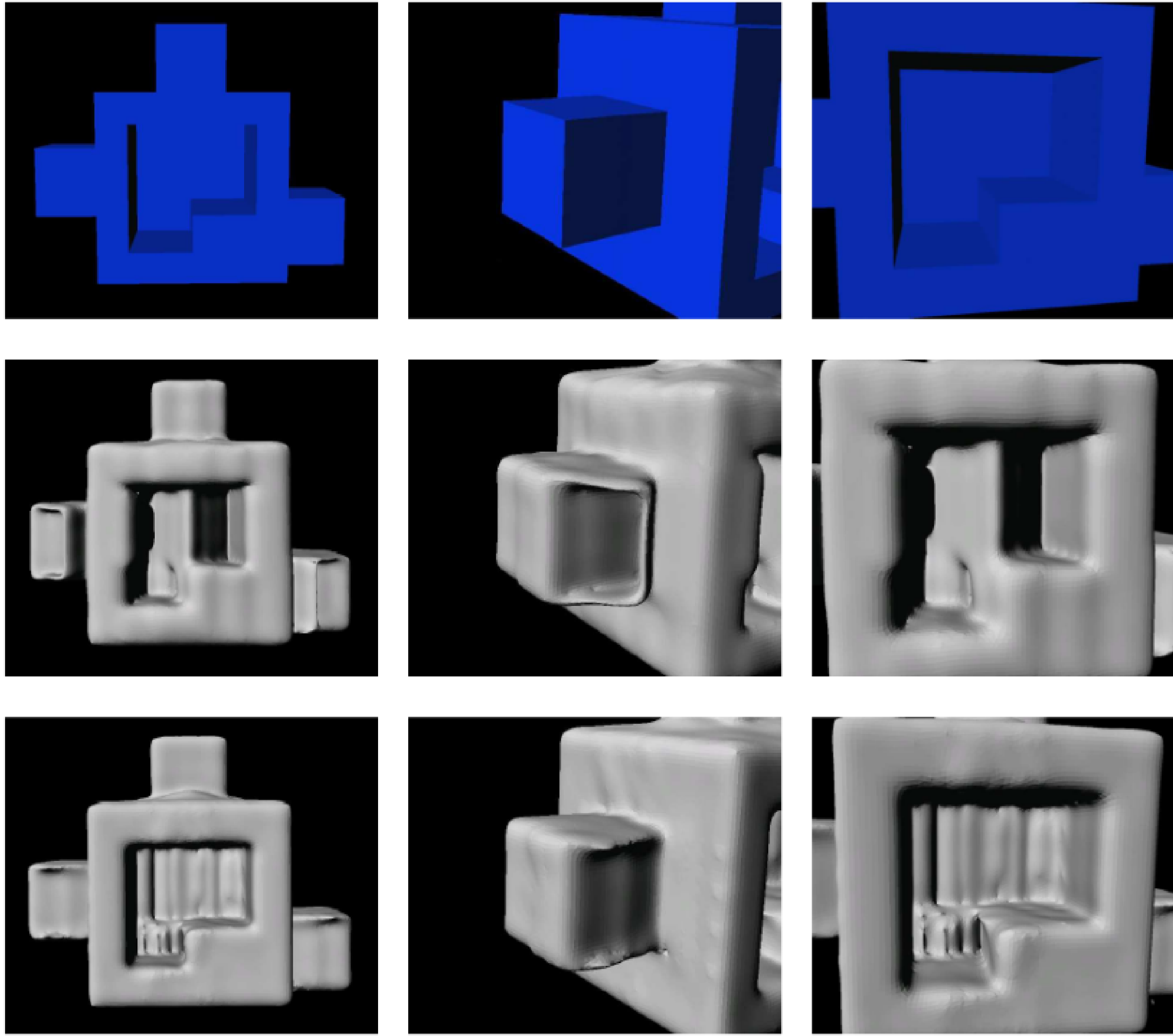


Shadow Carving

[Savarese et al '01]



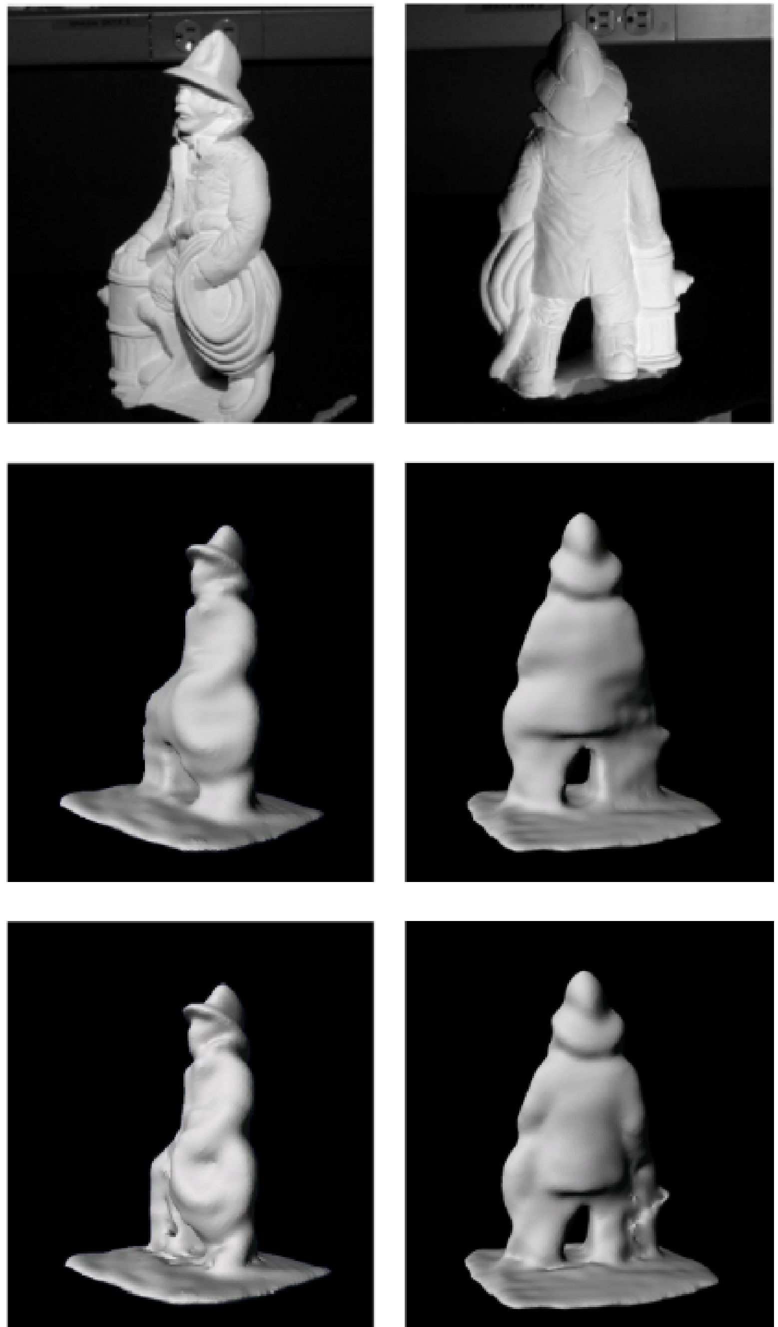

Camera
Image line
Light source
Image
Image shadow
virtual light image
Upper-bound from step k-1
Object



- 24 positions
- 4 lights

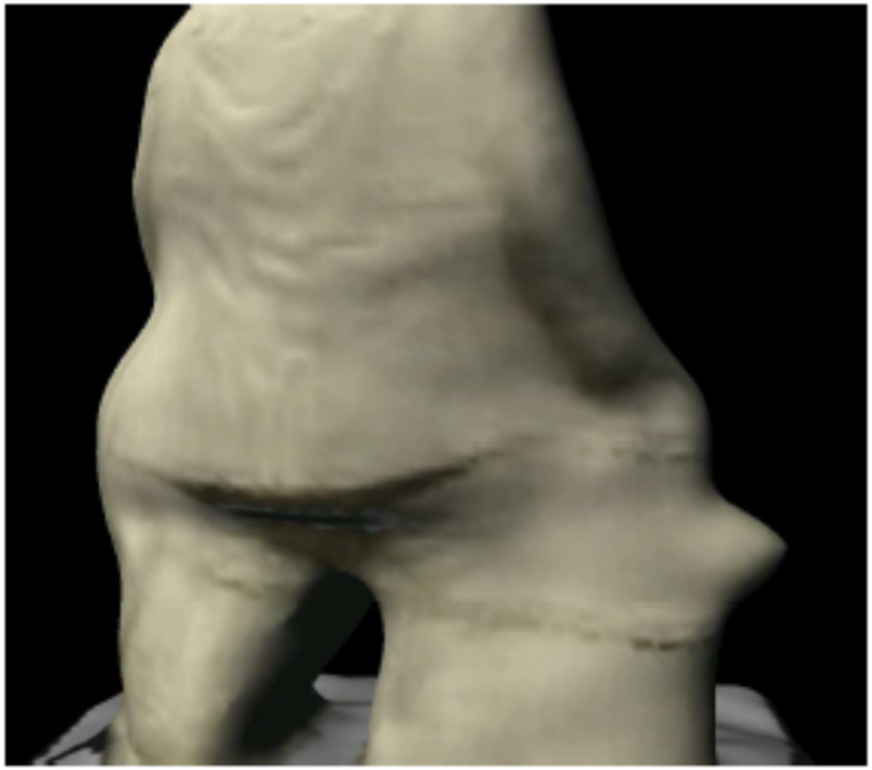
- 72 positions
- 8 lights

- 16 positions
- 4 lights

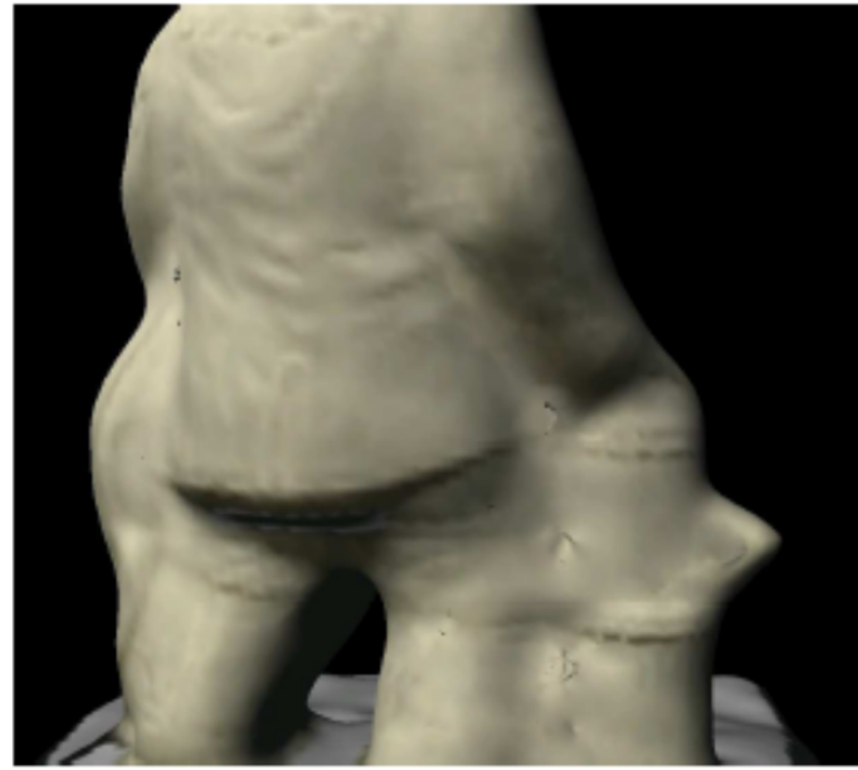


Space carving

Shadow carving



Space carving

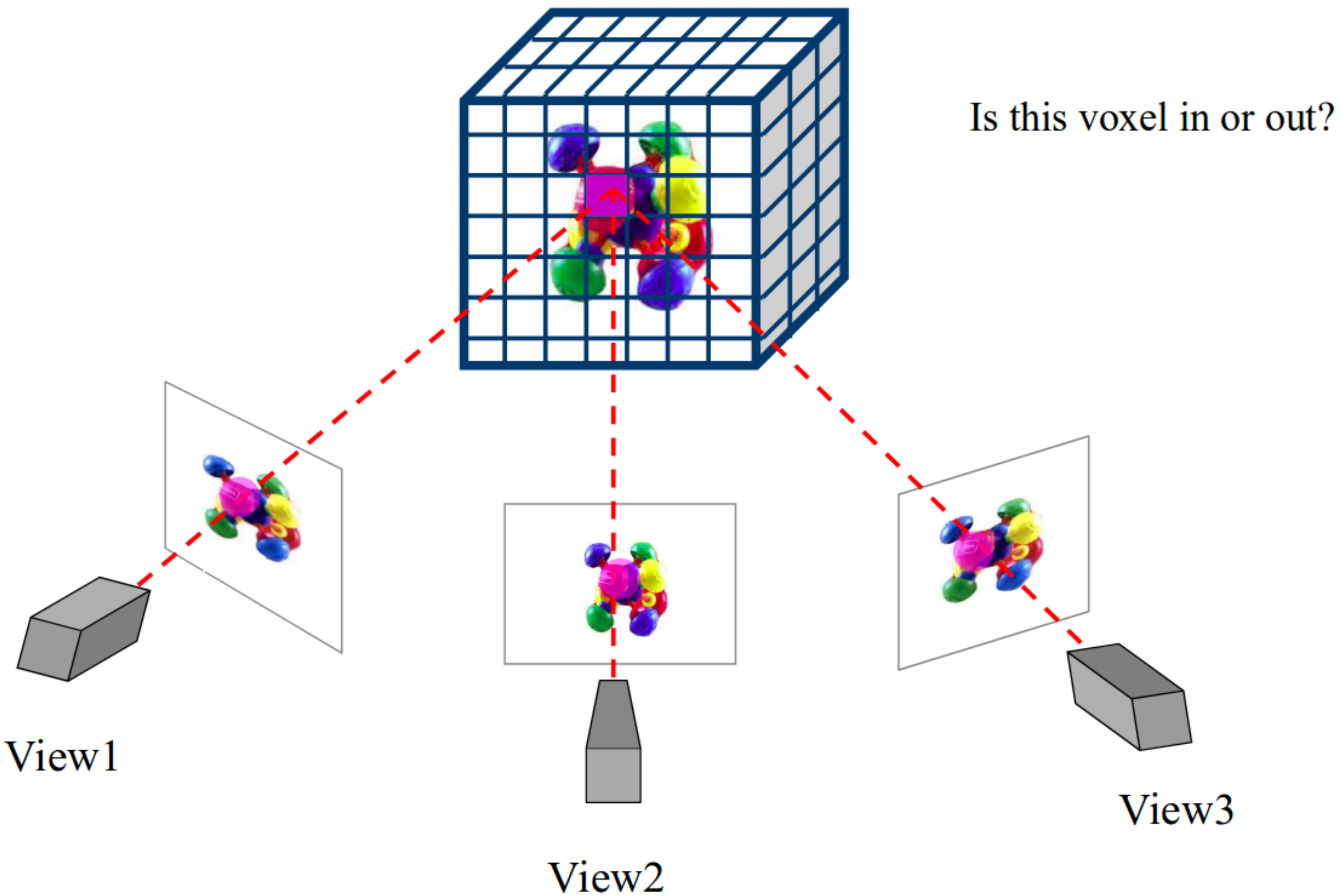


Shadow carving

Limitation of Shadow Carving

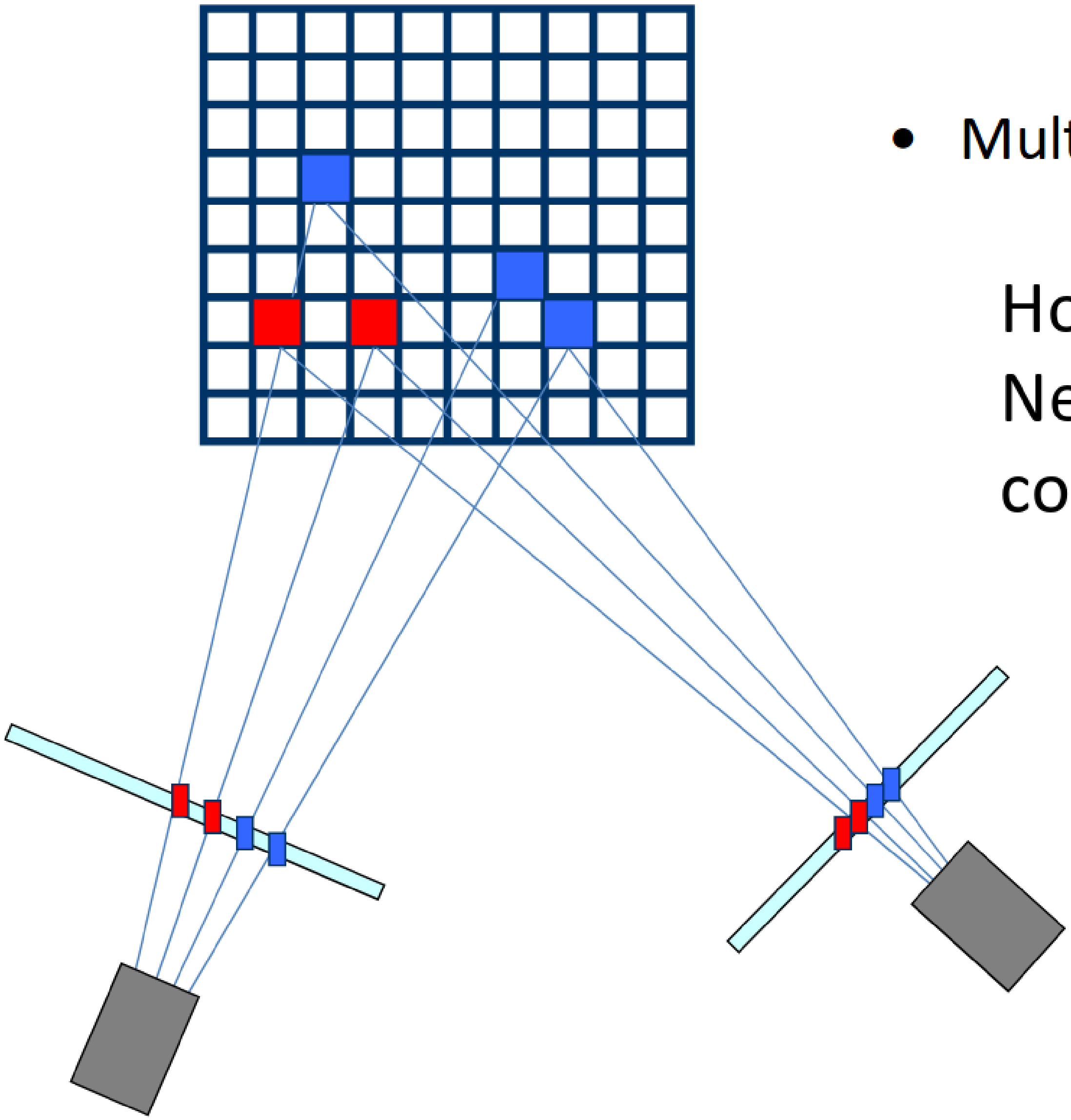
- Cannot handle cases where object contains reflective or low albedo regions.
- Shadows cannot be detected accurately in such conditions

Voxel Coloring



[Seitz & Dyer (‘97)]
[R. Collins (Space Sweep, ’96)]

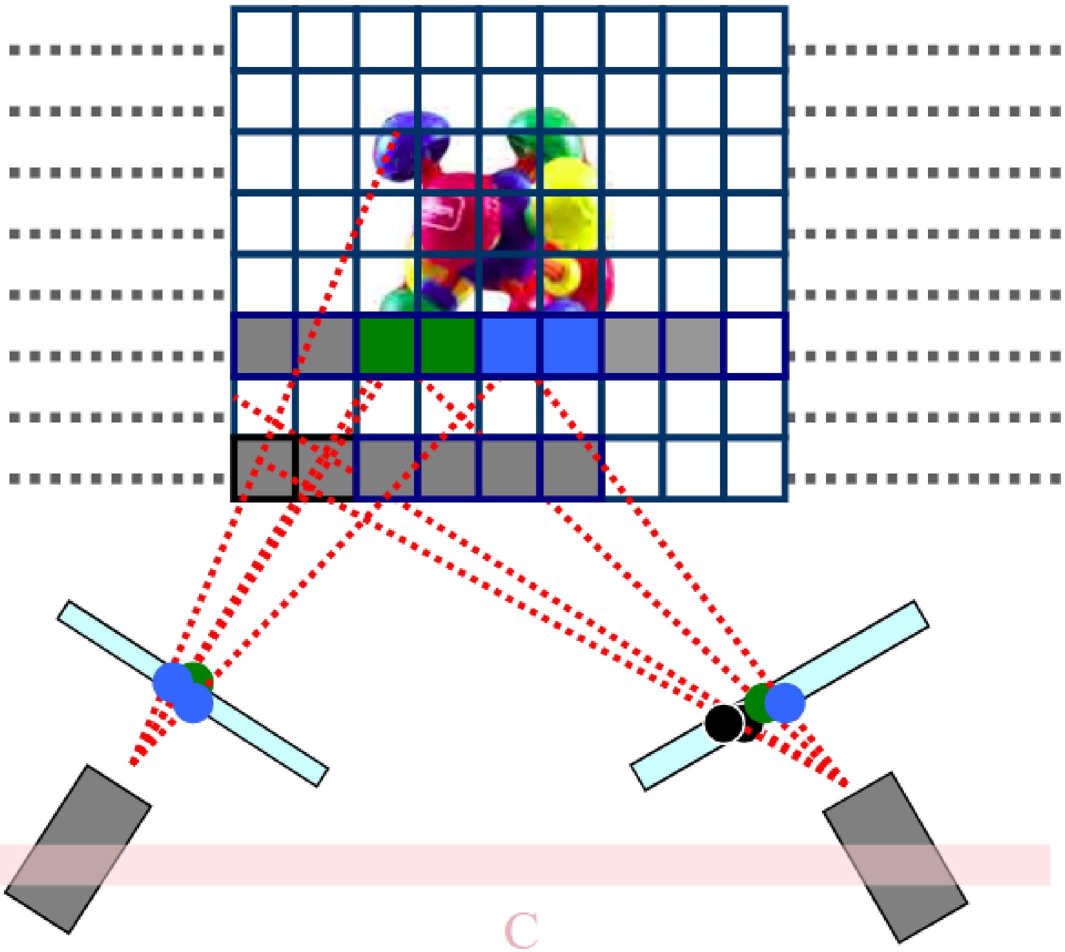
Uniqueness



- Multiple consistent scenes

How to fix this?
Need to use a visibility constraint

Algorithm for enforcing visibility constraints



Voxel Coloring

Advantage:
Clolored voxel

Disadvantage:
Lambertian Surface

Non Lambertian Surface