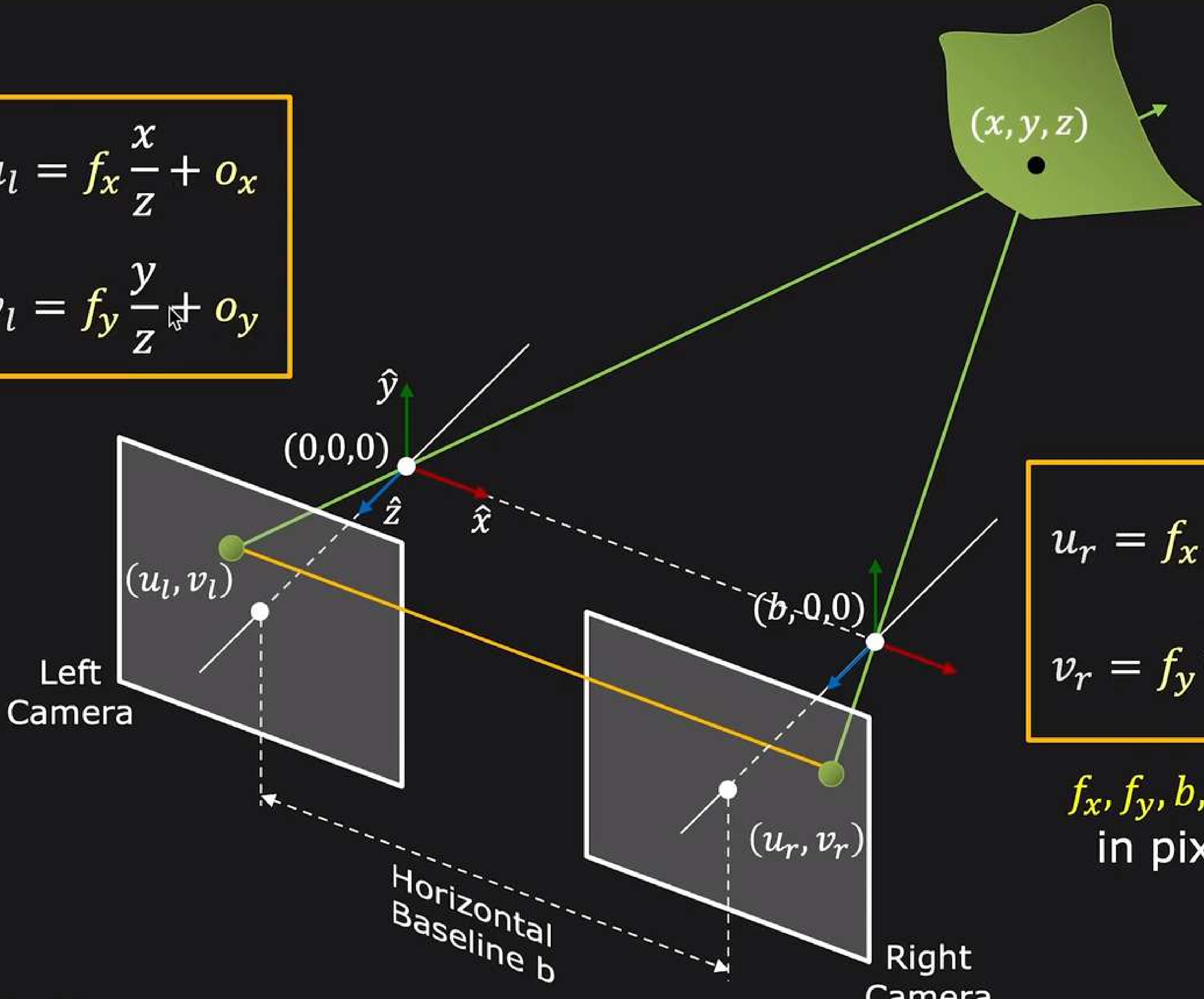
Simple (Calibrated) Stereo

$$u_l = f_x \frac{x}{z} + o_x$$

$$v_l = f_y \frac{y}{z} + o_y$$



$$u_r = f_x \frac{x - b}{z} + o_x$$

$$v_r = f_y \frac{y}{z} + o_y$$

f_x, f_y, b, o_x, o_y are in pixel units.

Depth and Disparity

Solving for (x, y, z):

$$x = \frac{b(u_l - o_x)}{(u_l - u_r)} \qquad y = \frac{bf_x(v_l - o_y)}{f_y(u_l - u_r)} \qquad \boxed{z = \frac{bf_x}{(u_l - u_r)}}$$

where $(u_l - u_r)$ is called the Disparity.

Uncalibrated Stereo

Method to estimate 3D structure of a static scene from two arbitrary views.

Topics:

(1) Problem of Uncalibrated Stereo

(2) Epipolar Geometry

(3) Estimating Fundamental Matrix

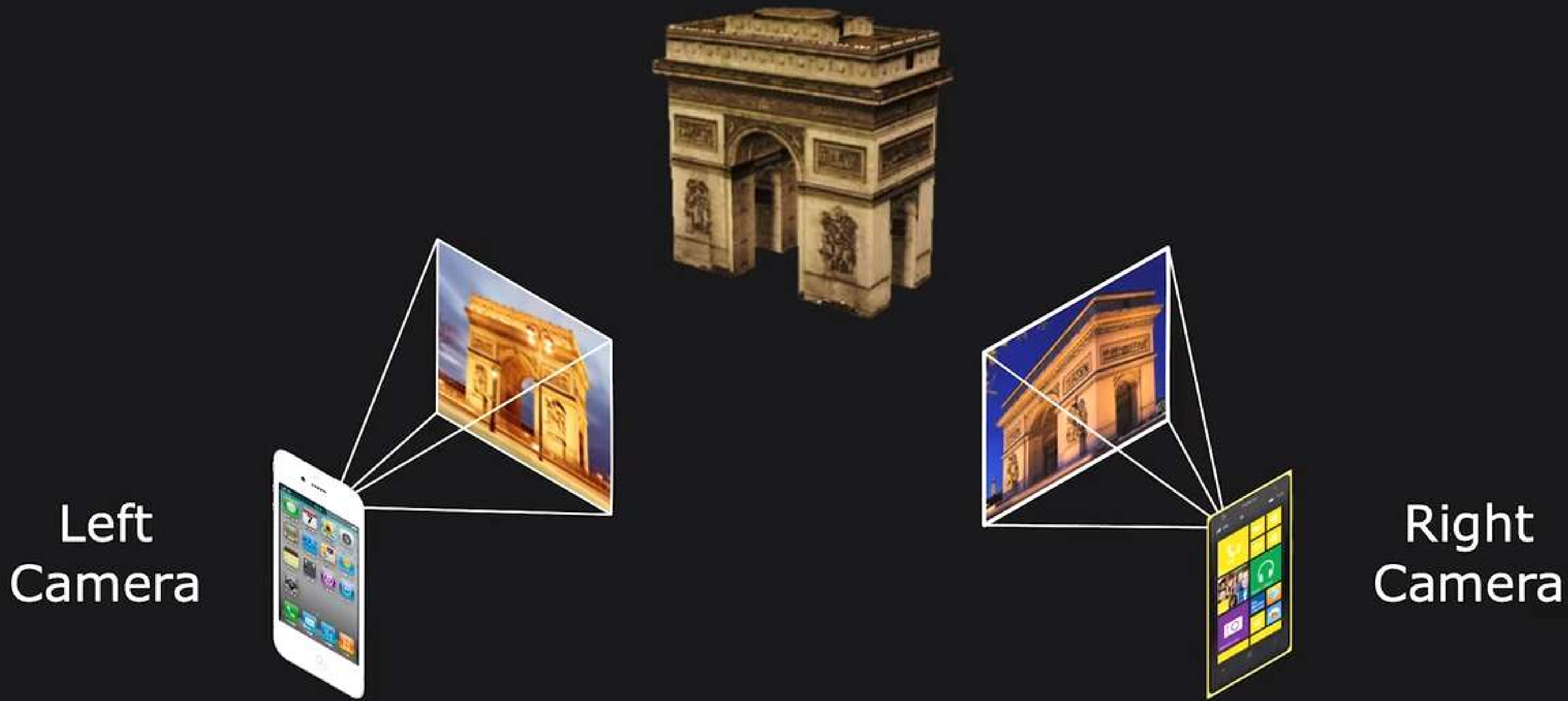
(4) Finding Dense Correspondences

(5) Computing Depth

(6) Stereopsis: Stereo in Nature

Uncalibrated Stereo

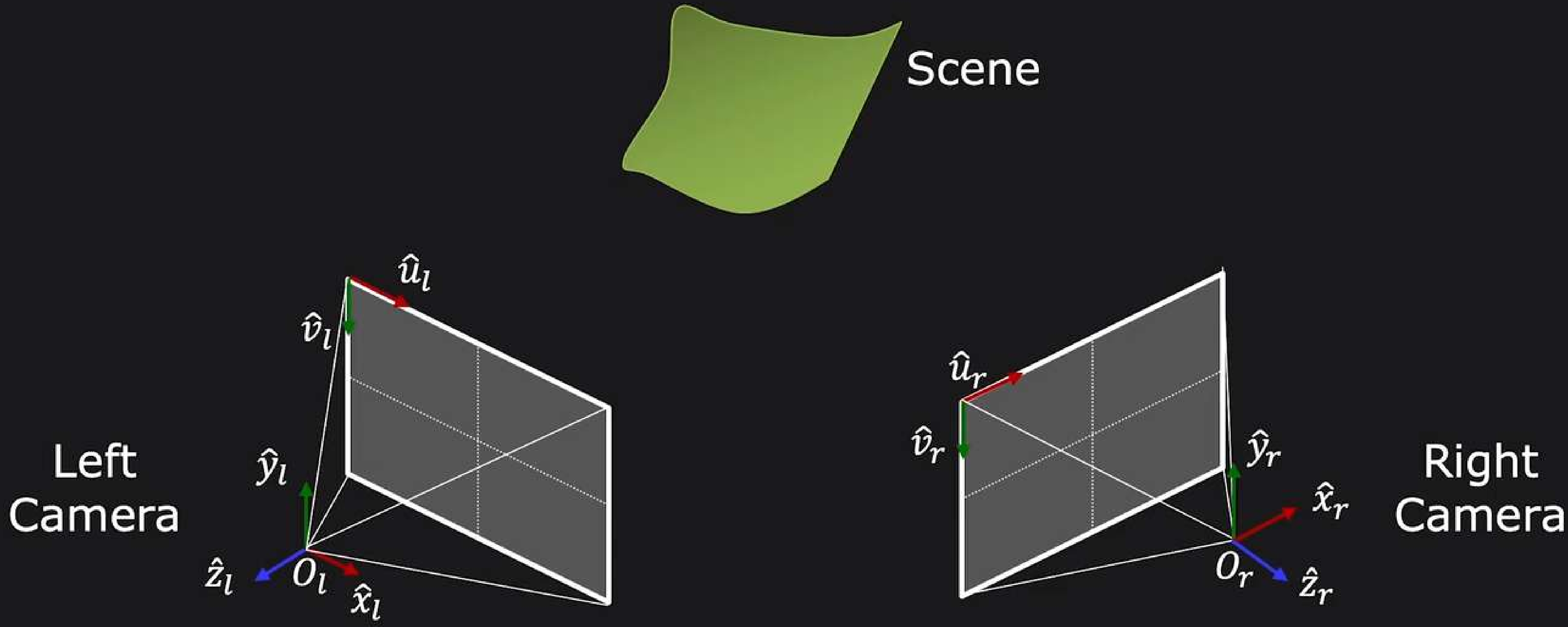
Compute 3D structure of static scene from two arbitrary views



Instrinsics (f_x, f_y, o_x, o_y) are known for both views/cameras.

Extrinsics (relative position/orientation of cameras) are unknov

Uncalibrated Stereo

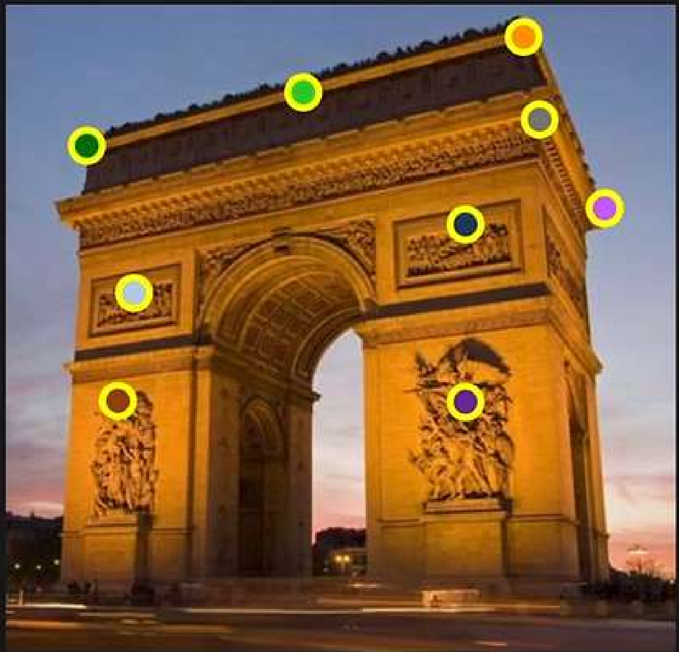


1. Assume Camera Matrix K is known for each camera
2. Find a few Reliable Corresponding Points

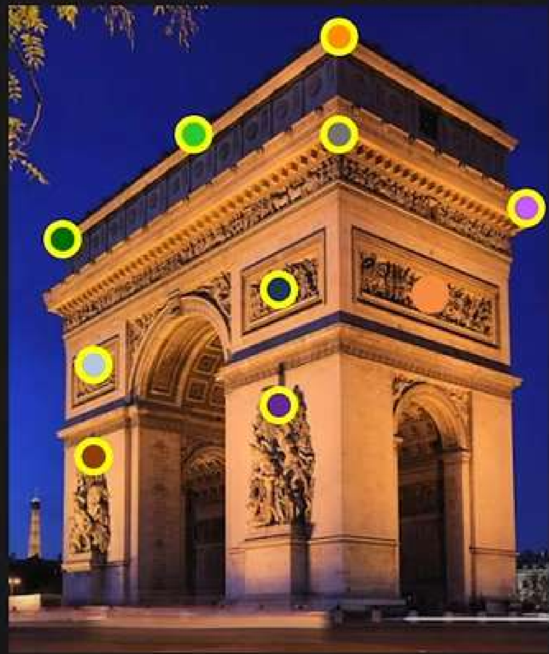
Initial Correspondence

Find a set of corresponding features (at least 8) in left and right images (e.g. using SIFT or hand-picked).

Left image



Right image

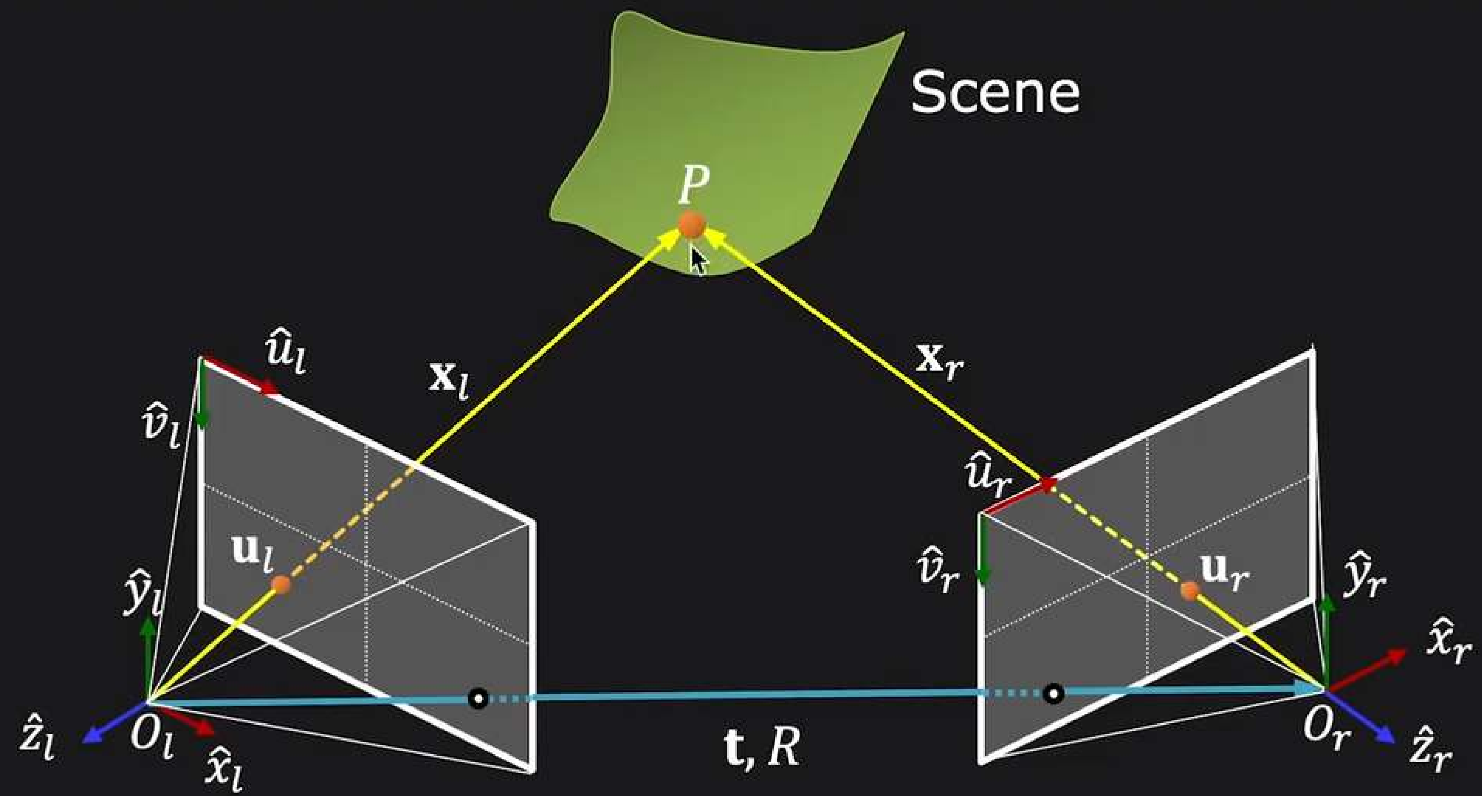


$(u_l^{(1)}, v_l^{(1)})$ $(u_r^{(1)}, v_r^{(1)})$

$\vdots$ $\vdots$

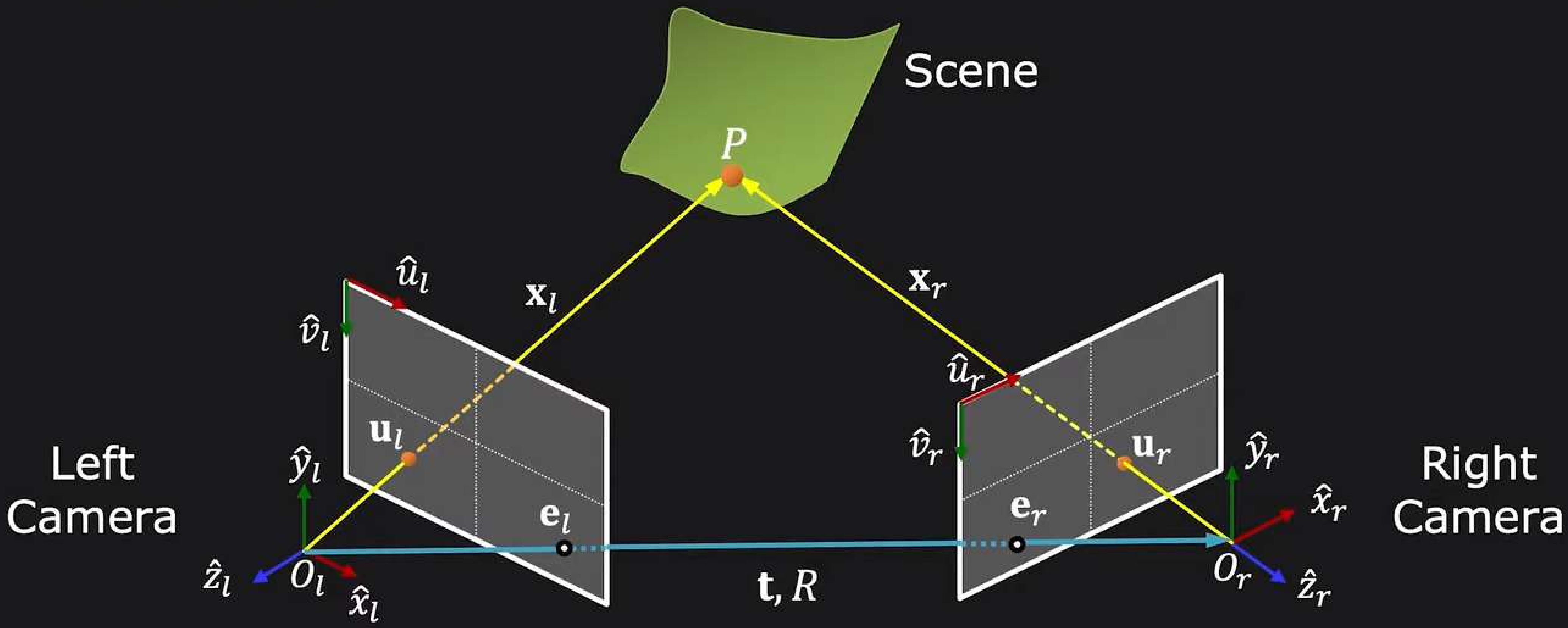
$(u_l^{(m)}, v_l^{(m)})$ $(u_r^{(m)}, v_r^{(m)})$

Uncalibrated Stereo



1. Assume Camera Matrix K is known for each camera
2. Find a few Reliable Corresponding Points
3. Find Relative Camera Position $\mathbf{t}$ and Orientation R
4. Find Dense Correspondence
5. Compute Depth using Triangulation

Epipolar Geometry: Epipoles

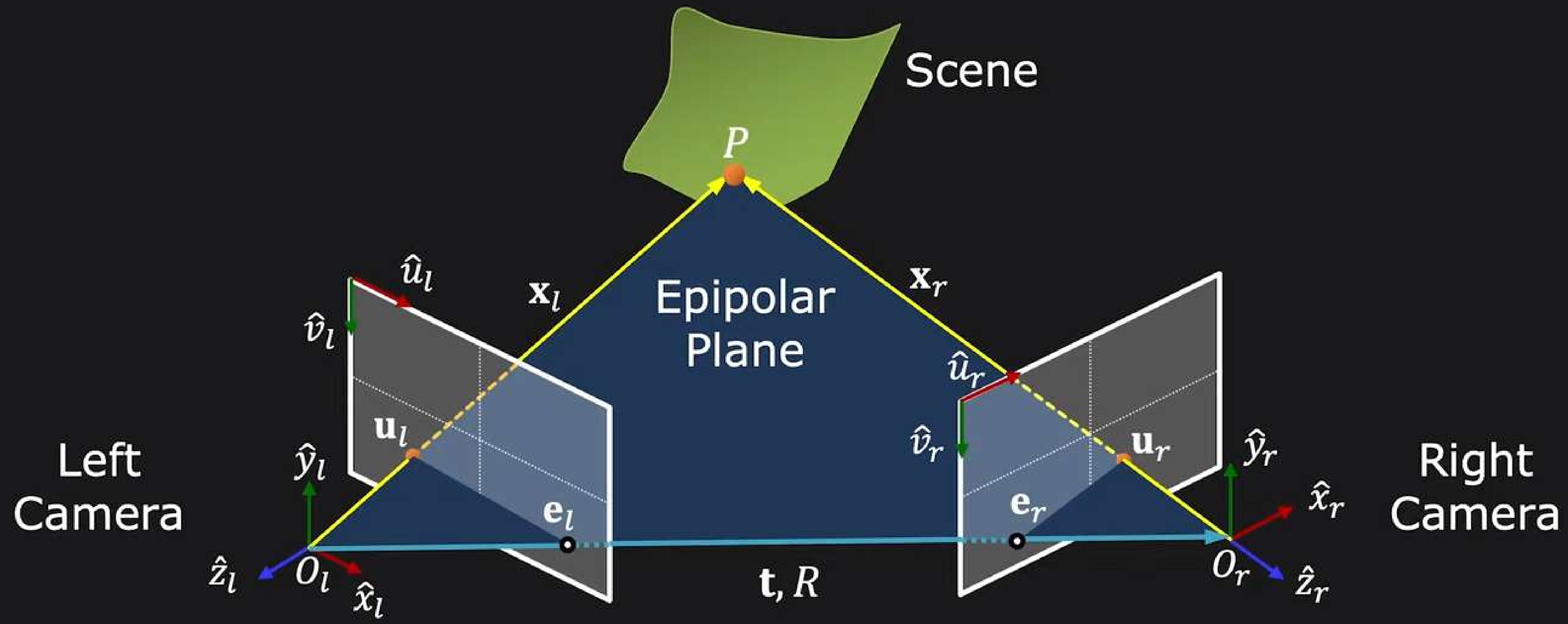


Epipole: Image point of origin/pinhole of one camera as viewed by the other camera.

$\mathbf{e}_l$ and $\mathbf{e}_r$ are the epipoles.

$\mathbf{e}_l$ and $\mathbf{e}_r$ are unique for a given stereo pair.

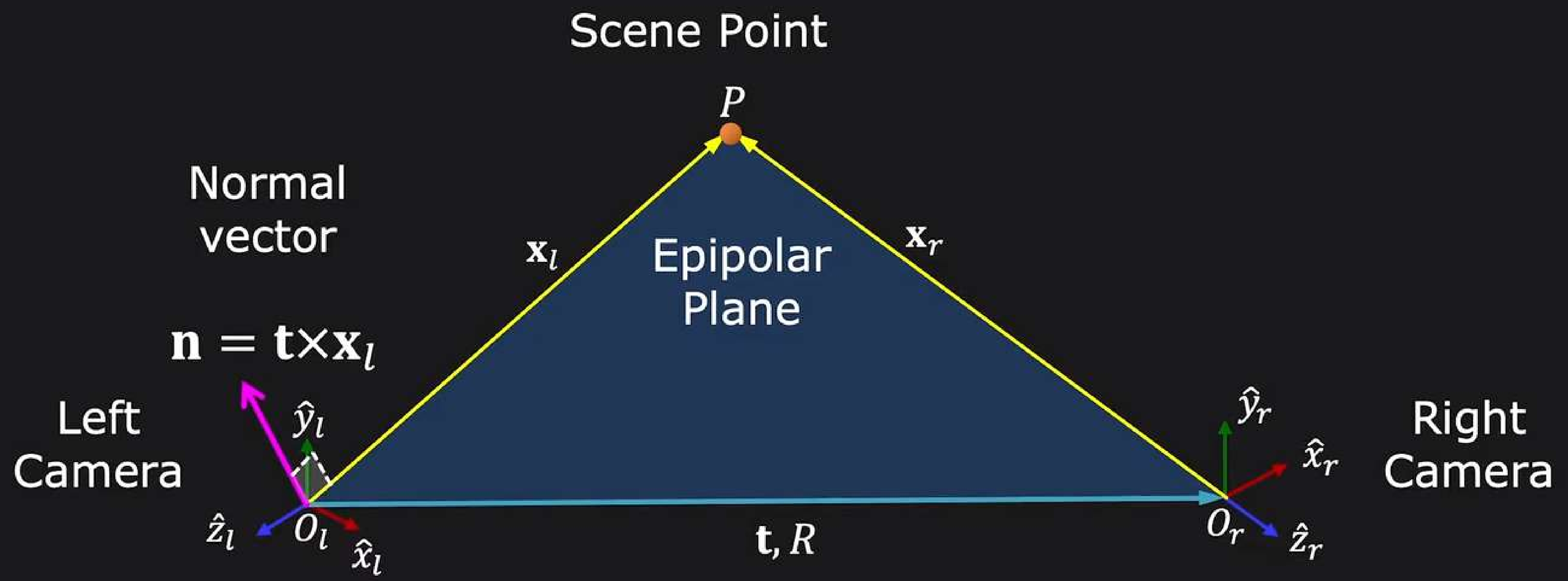
Epipolar Geometry: Epipolar Plane



Epipolar Plane of Scene Point P: The plane formed by camera origins (O_l and O_r), epipoles ($\mathbf{e}_l$ and $\mathbf{e}_r$) and scene point P.

Every scene point lies on a unique epipolar plane.

Epipolar Constraint



Vector normal to the epipolar plane: $\mathbf{n} = \mathbf{t} \times \mathbf{x}_l$

Dot product of $\mathbf{n}$ and $\mathbf{x}_l$ (perpendicular vectors) is zero:

$$\boxed{\mathbf{x}_l \cdot (\mathbf{t} \times \mathbf{x}_l) = 0}$$

Epipolar Constraint

Writing the epipolar constraint in matrix form:

$$\mathbf{x}_l \cdot (\mathbf{t} \times \mathbf{x}_l) = 0$$

$$\begin{bmatrix} x_l & y_l & z_l \end{bmatrix} \begin{bmatrix} t_y z_l - t_z y_l \\ t_z x_l - t_x z_l \\ t_x y_l - t_y x_l \end{bmatrix} = 0$$

Cross-product definition

$$\boxed{\begin{bmatrix} x_l & y_l & z_l \end{bmatrix} \underbrace{\begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}}_{T_\times} \begin{bmatrix} x_l \\ y_l \\ z_l \end{bmatrix} = 0}$$

Matrix-vector form

$\mathbf{t}_{3\times 1}$: Position of Right Camera in Left Camera's Frame

$R_{3\times 3}$: Orientation of Left Camera in Right Camera's Frame

$$\boxed{\mathbf{x}_l = R\mathbf{x}_r + \mathbf{t}}$$

$$\boxed{\begin{bmatrix} x_l \\ y_l \\ z_l \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}}$$

Epipolar Constraint

Substituting into the epipolar constraint gives:

$$\begin{bmatrix} x_l & y_l & z_l \end{bmatrix} \left(\boxed{\begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}} \begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} + \boxed{\begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}} \right) = 0$$

$$\boldsymbol{t} \times \boldsymbol{t} = \boldsymbol{0}$$

$$\begin{bmatrix} x_l & y_l & z_l \end{bmatrix} \boxed{\begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix}} \begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} = 0$$

Essential Matrix E

$$\boxed{E = T_{\times} R}$$

[Longuet-Higgins 1981]

Essential Matrix E: Decomposition

$$\boxed{E = T_\times R}$$

$$\begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \boxed{\begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}}$$

Given that $T_\times$ is a Skew-Symmetric matrix (a_{ij} = - a_{ji}) and R is an Orthonormal matrix, it is possible to "decouple" $T_\times$ and R from their product using "Singular Value Decomposition".

Take Away: If E is known, we can calculate $\boldsymbol{t}$ and R.

How do we find E?

Relates 3D position (x_l, y_l, z_l) of scene point w.r.t left camera to its 3D position (x_r, y_r, z_r) w.r.t. right camera

$$\boxed{\mathbf{x}_l^T E \mathbf{x}_r = 0}$$

$$[x_l \quad y_l \quad z_l] \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} \begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} = 0$$

3D position in left camera coordinates

3x3 Essential Matrix

3D position in right camera coordinates

Unfortunately, we don't have $\mathbf{x}_l$ and $\mathbf{x}_r$.

But we do know corresponding points in image coordinates.

Incorporating the Image Coordinates

Perspective projection equations for left camera:

$$u_l = f_x^{(l)} \frac{x_l}{z_l} + o_x^{(l)} \qquad v_l = f_y^{(l)} \frac{y_l}{z_l} + o_y^{(l)}$$

$$z_l u_l = f_x^{(l)} x_l + z_l o_x^{(l)} \qquad z_l v_l = f_y^{(l)} y_l + z_l o_y^{(l)}$$

Representing in matrix form:

$$z_l \begin{bmatrix} u_l \\ v_l \\ 1 \end{bmatrix} = \begin{bmatrix} z_l u_l \\ z_l v_l \\ z_l \end{bmatrix} = \begin{bmatrix} f_x^{(l)} x_l + z_l o_x^{(l)} \\ f_y^{(l)} y_l + z_l o_y^{(l)} \\ z_l \end{bmatrix} = \underbrace{\begin{bmatrix} f_x^{(l)} & 0 & o_x^{(l)} \\ 0 & f_y^{(l)} & o_y^{(l)} \\ 0 & 0 & 1 \end{bmatrix}}_{\substack{\text{Known} \\ \text{Camera Matrix } K_l}} \begin{bmatrix} x_l \\ y_l \\ z_l \end{bmatrix}$$

REVIEW

Incorporating the Image Coordinates

Left camera

$$z_l \begin{bmatrix} u_l \\ v_l \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} f_x^{(l)} & 0 & o_x^{(l)} \\ 0 & f_y^{(l)} & o_y^{(l)} \\ 0 & 0 & 1 \end{bmatrix}}_{K_l} \begin{bmatrix} x_l \\ y_l \\ z_l \end{bmatrix}$$

Right camera

$$z_r \begin{bmatrix} u_r \\ v_r \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} f_x^{(r)} & 0 & o_x^{(r)} \\ 0 & f_y^{(r)} & o_y^{(r)} \\ 0 & 0 & 1 \end{bmatrix}}_{K_r} \begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix}$$

$$\boxed{\mathbf{x}_l{}^T = [u_l \quad v_l \quad 1] z_l \, {K_l^{-1}}^T}$$

$$\boxed{\mathbf{x}_r = K_r^{-1} z_r \begin{bmatrix} u_r \\ v_r \\ 1 \end{bmatrix}}$$

Incorporating the Image Coordinates

Epipolar constraint:

$$\begin{bmatrix} x_l & y_l & z_l \end{bmatrix} \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} \begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} = 0$$

Rewriting in terms of image coordinates:

$$\begin{bmatrix} u_l & v_l & 1 \end{bmatrix} \cancel{z_l} \, {K_l^{-1}}^T \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} K_r^{-1} \, \cancel{z_r} \begin{bmatrix} u_r \\ v_r \\ 1 \end{bmatrix} = 0$$

$$\boxed{\begin{matrix} z_l \neq 0 \\ z_r \neq 0 \end{matrix}}$$

Incorporating the Image Coordinates

Epipolar constraint:

$$[x_l \quad y_l \quad z_l] \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} \begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} = 0$$

Rewriting in terms of image coordinates:

$$\boxed{[u_l \quad v_l \quad 1]\, {K_l^{-1}}^T \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} K_r^{-1} \begin{bmatrix} u_r \\ v_r \\ 1 \end{bmatrix} = 0}$$

Fundamental Matrix F

Epipolar constraint:

$$[x_l \quad y_l \quad z_l] \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} \begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} = 0$$

Rewriting in terms of image coordinates:

$$[u_l \quad v_l \quad 1] \underbrace{\boxed{\begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}}}_{\text{Fundamental Matrix } F} \begin{bmatrix} u_r \\ v_r \\ 1 \end{bmatrix} = 0$$

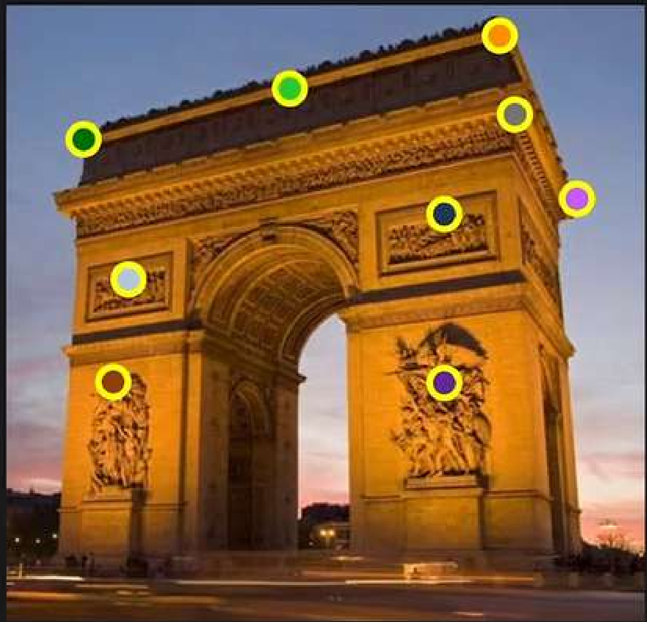
$$\boxed{E = K_l^T F K_r} \qquad \boxed{E = T_\times R}$$

[Fagueras 1992, Luong 1992]

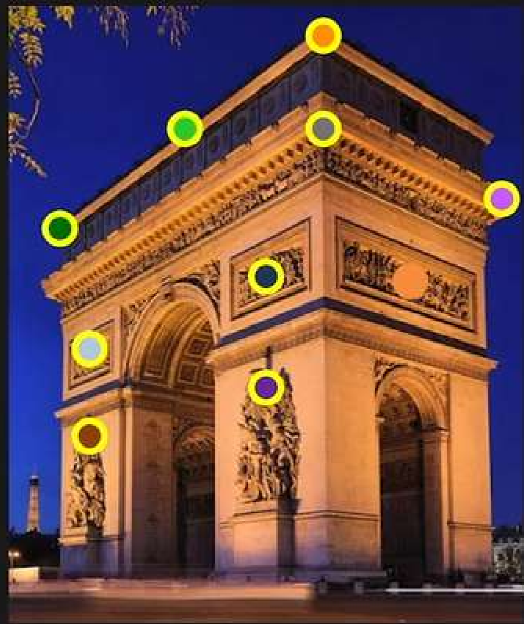
Stereo Calibration Procedure

Find a set of corresponding features in left and right images (e.g. using SIFT or hand-picked)

Left image



Right image



$(u_l^{(1)}, v_l^{(1)})$ $(u_r^{(1)}, v_r^{(1)})$

$\vdots$ $\vdots$

$(u_l^{(m)}, v_l^{(m)})$ $(u_r^{(m)}, v_r^{(m)})$

Stereo Calibration Procedure

Step A: For each correspondence i, write out epipolar constraint.

$$\underbrace{\begin{bmatrix} u_l^{(i)} & v_l^{(i)} & 1 \end{bmatrix}}_{\text{Known}} \underbrace{\begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}}_{\text{Unknown}} \underbrace{\begin{bmatrix} u_r^{(i)} \\ v_r^{(i)} \\ 1 \end{bmatrix}}_{\text{Known}} = 0$$

Expand the matrix to get linear equation:

$$\left(f_{11}u_r^{(i)} + f_{12}v_r^{(i)} + f_{13}\right)u_l^{(i)} + \left(f_{21}u_r^{(i)} + f_{22}v_r^{(i)} + f_{23}\right)v_l^{(i)} + f_{31}u_r^{(i)} + f_{32}v_r^{(i)} + f_{33} =$$

Stereo Calibration Procedure

Step B: Rearrange terms to form a linear system.

$$\begin{bmatrix} u_l^{(1)}u_r^{(1)} & u_l^{(1)}v_r^{(1)} & u_l^{(1)} & v_l^{(1)}u_r^{(1)} & v_l^{(1)}v_r^{(1)} & v_l^{(1)} & u_r^{(1)} & v_r^{(1)} & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u_l^{(i)}u_r^{(i)} & u_l^{(i)}v_r^{(i)} & u_l^{(i)} & v_l^{(i)}u_r^{(i)} & v_l^{(i)}v_r^{(i)} & v_l^{(i)} & u_l^{(i)} & u_r^{(i)} & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u_l^{(m)}u_r^{(m)} & u_l^{(m)}v_r^{(m)} & u_l^{(m)} & v_l^{(m)}u_r^{(m)} & v_l^{(m)}v_r^{(m)} & v_l^{(m)} & u_l^{(m)} & u_r^{(m)} & 1 \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{21} \\ f_{31} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

A

(Known)

$\mathbf{f}$

(Unknown)

$$\boxed{A\,\mathbf{f} = \mathbf{0}}$$

The Tale of Missing Scale

Fundamental matrix acts on homogenous coordinates.

$$[u_l \quad v_l \quad 1]\begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}\begin{bmatrix} u_r \\ v_r \\ 1 \end{bmatrix} = 0 = [u_l \quad v_l \quad 1]\begin{bmatrix} kf_{11} & kf_{12} & kf_{13} \\ kf_{21} & kf_{22} & kf_{23} \\ kf_{31} & kf_{32} & kf_{33} \end{bmatrix}\begin{bmatrix} u_r \\ v_r \\ 1 \end{bmatrix}$$

Fundamental Matrix F and kF describe the same epipolar geometry. That is, F is defined only up to a scale.

Set Fundamental Matrix to some arbitrary scale.

$$\boxed{\|\mathbf{f}\|^2 = 1}$$

Solving for F

Step C: Find least squares solution for fundamental matrix F.

We want $A\mathbf{f}$ as close to 0 as possible and $\|\mathbf{f}\|^2 = 1$:

$$\boxed{\min_{\mathbf{f}} \|A\mathbf{f}\|^2 \quad \text{such that } \|\mathbf{f}\|^2 = 1}$$

Constrained linear least squares problem

Like solving Projection Matrix during Camera Calibration.

Or, Homography Matrix for Image Stitching.

Rearrange solution $\mathbf{f}$ to form the fundamental matrix F.

Extracting Rotation and Translation

Step D: Compute essential matrix E from known left and right intrinsic camera matrices and fundamental matrix F.

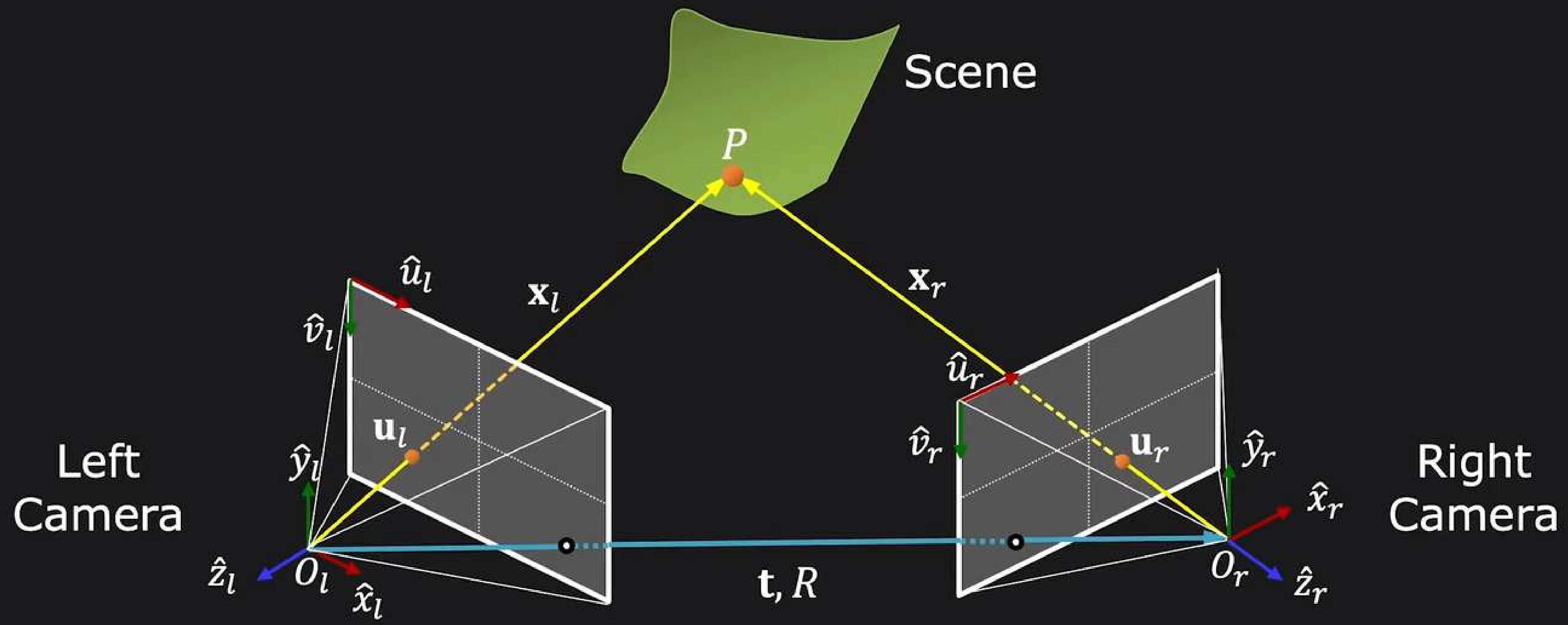
$$\boxed{E = K_l^T F K_r}$$

Step E: Extract R and $\mathbf{t}$ from E.

$$\boxed{E = T_\times R}$$

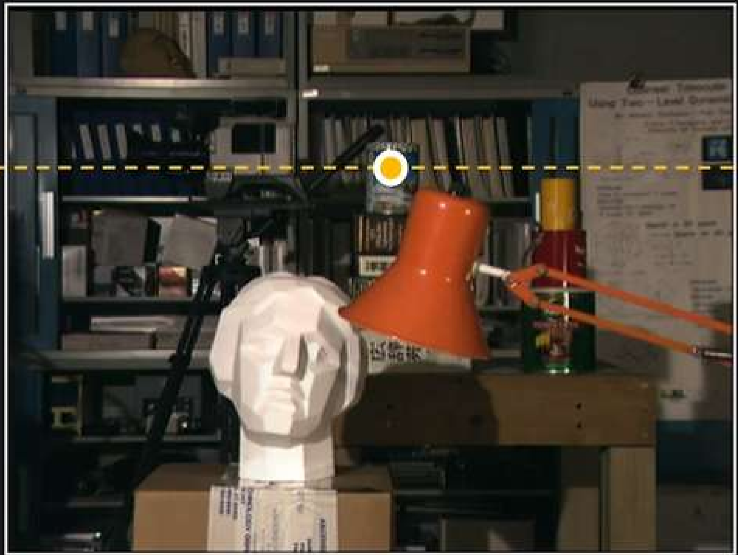
(Using Singular Value Decomposition)

Uncalibrated Stereo

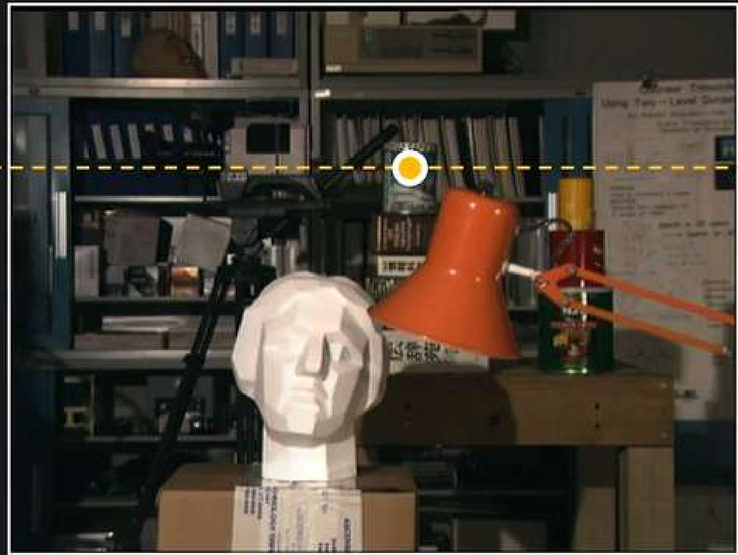


1. Assume Camera Matrix K is known for each camera
2. Find a few Reliable Corresponding Points
3. Find Relative Camera Position $\mathbf{t}$ and Orientation R
4. Find Dense Correspondence
5. Compute Depth using Triangulation

Simple Stereo: Finding Correspondences



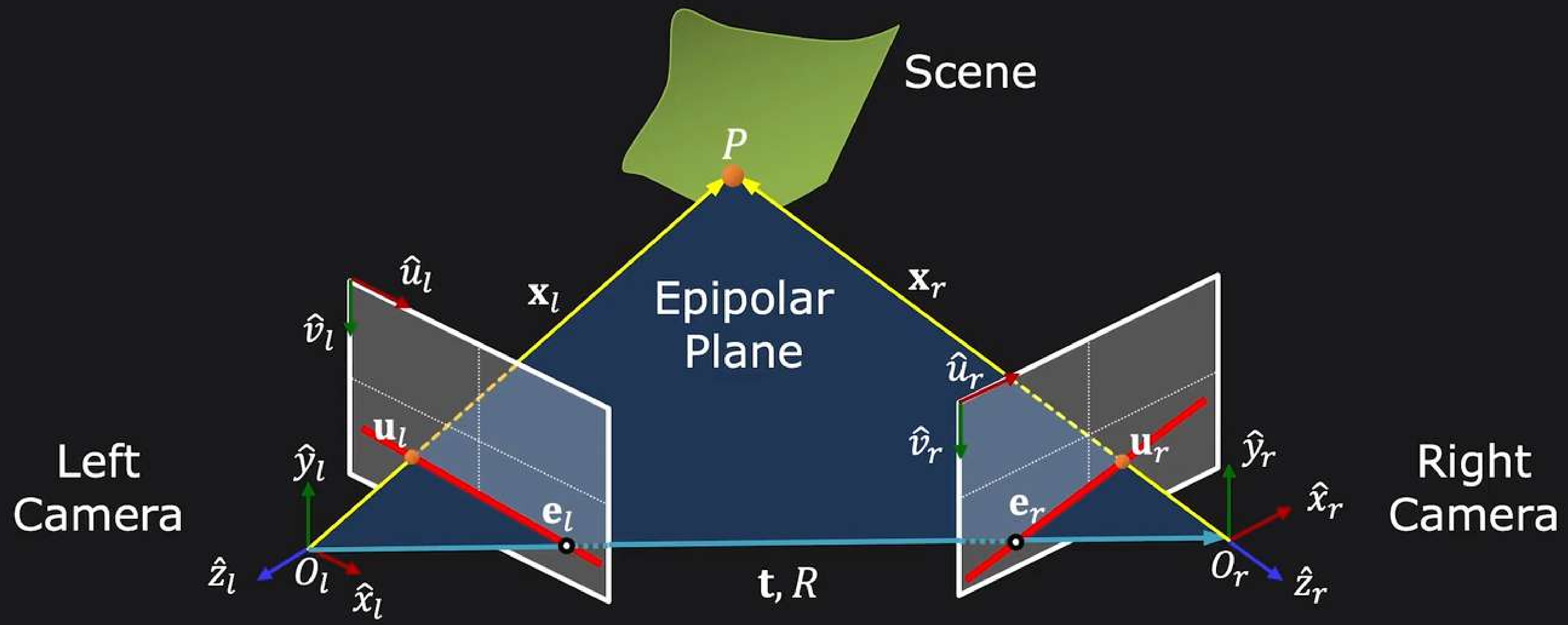
Left Camera Image



Right Camera Image

Corresponding scene points lie on the same horizontal scan-line.
Finding correspondence is a 1D search.

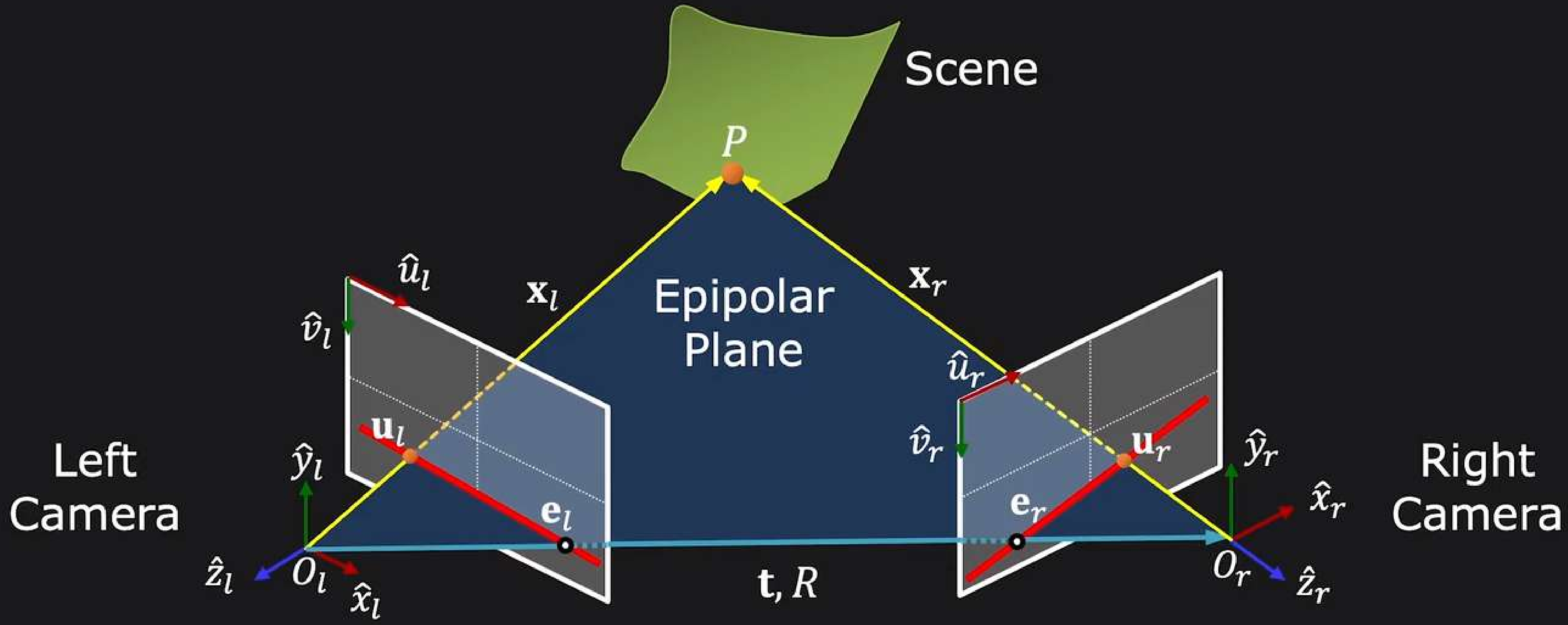
Epipolar Geometry: Epipolar Line



Epipolar Line: Intersection of image plane and epipolar plane.

Every scene point has two corresponding epipolar lines, one each on the two image planes.

Epipolar Geometry: Epipolar Line



Given a point in one image, the corresponding point in the other image must lie on the epipolar line.

Finding correspondence reduces to a 1D search.

Finding Epipolar Lines

Given: Fundamental matrix F and point on left image (u_l, v_l)

Find: Equation of Epipolar line in the right image

Epipolar Constraint Equation:

$$[u_l \quad v_l \quad 1] \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{bmatrix} u_r \\ v_r \\ 1 \end{bmatrix} = 0$$

Expanding the matrix equation gives:

$$(f_{11}u_l + f_{21}v_l + f_{31})u_r + (f_{12}u_l + f_{22}v_l + f_{32})v_r + (f_{13}u_l + f_{23}v_l + f_{33}) = 0$$

Equation for right epipolar line: $\boxed{a_l u_r + b_l v_r + c_l = 0}$

Similarly we can calculate epipolar line in left image for a point in right image.

Finding Epipolar Lines: Example

Given the Fundamental matrix,

$$F = \begin{bmatrix} -.003 & -.028 & 13.19 \\ -.003 & -.008 & -29.2 \\ 2.97 & 56.38 & -9999 \end{bmatrix}$$

and the left image point

$$\tilde{\boldsymbol{u}}_l = \begin{bmatrix} 343 \\ 221 \\ 1 \end{bmatrix}$$

Left Image

Right Image

The equation for the epipolar line in the right image is

$$[u_r \quad v_r \quad 1] \begin{bmatrix} -.003 & -.003 & 2.97 \\ -.028 & -.008 & 56.38 \\ 13.19 & -29.2 & -9999 \end{bmatrix} \begin{bmatrix} 343 \\ 221 \\ 1 \end{bmatrix} = 0$$

Finding Epipolar Lines: Example

Given the Fundamental matrix,

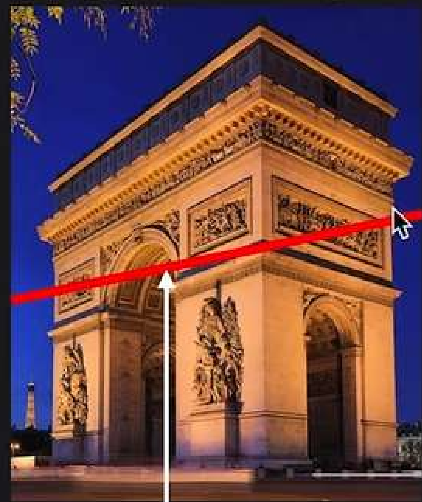
$$F = \begin{bmatrix} -.003 & -.028 & 13.19 \\ -.003 & -.008 & -29.2 \\ 2.97 & 56.38 & -9999 \end{bmatrix}$$

and the left image point

$$\tilde{\boldsymbol{u}}_l = \begin{bmatrix} 343 \\ 221 \\ 1 \end{bmatrix}$$

Left Image

Right Image



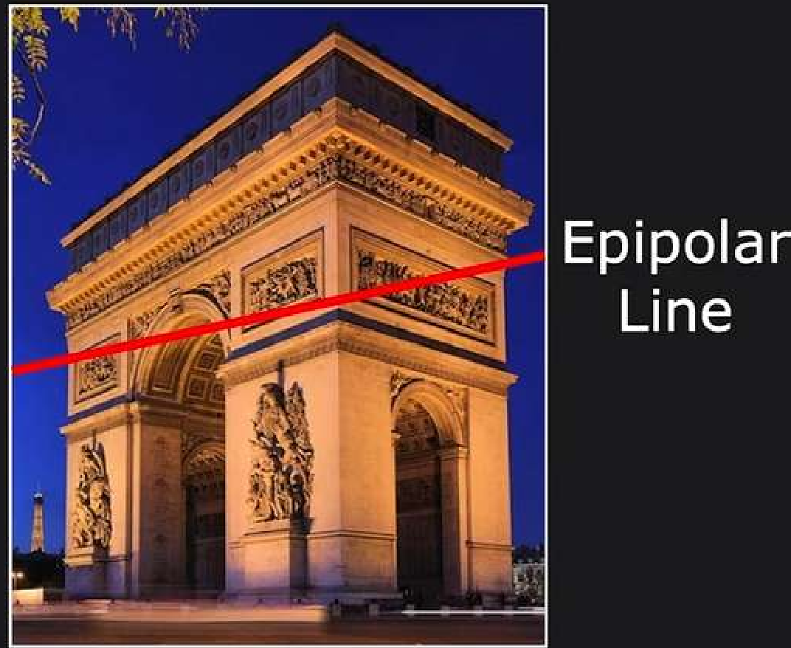
Epipolar Line

The equation for the epipolar line in the right image is

$$.03u_r + .99v_r - 265 = 0$$

Finding Correspondence

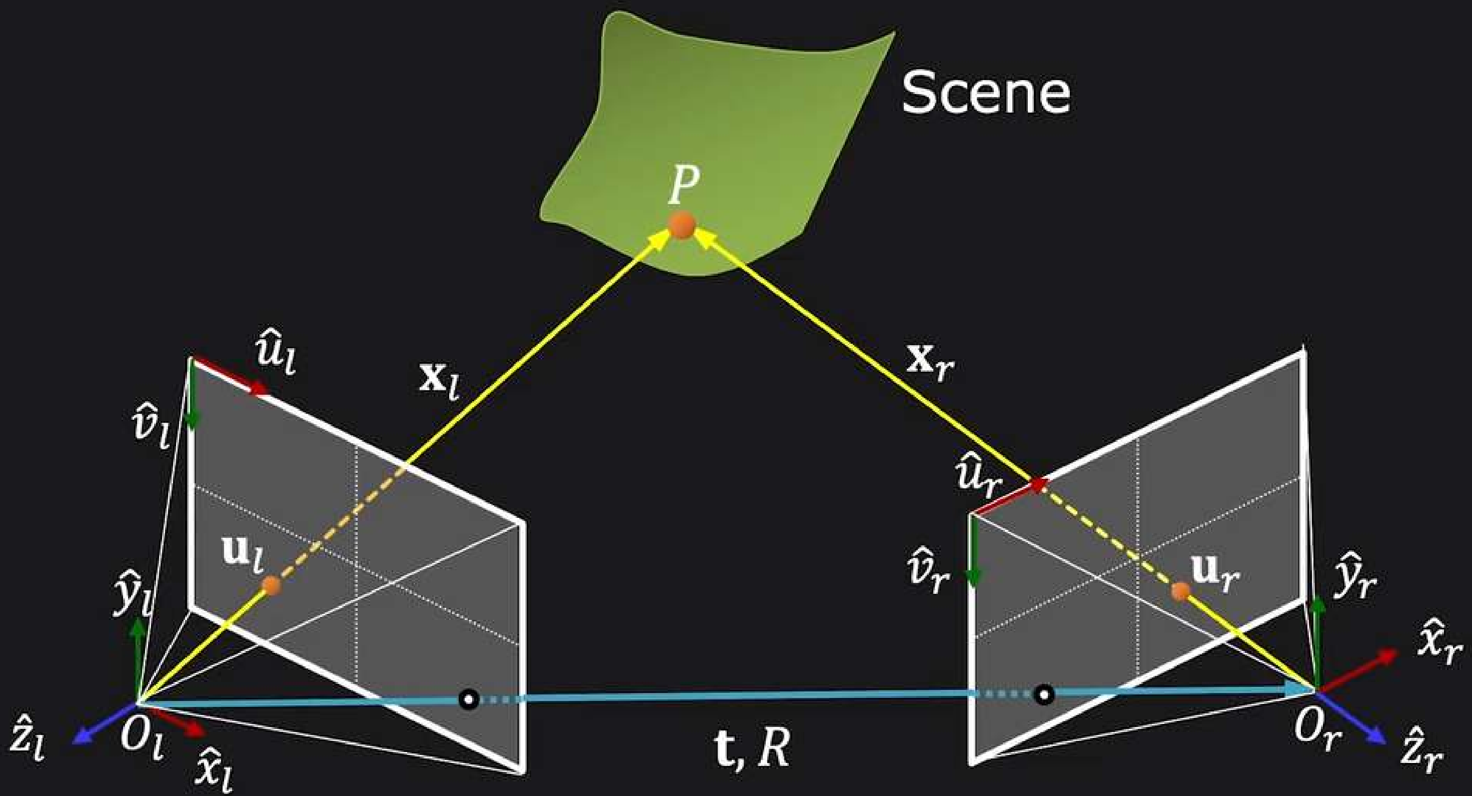
Left Image



Right Image

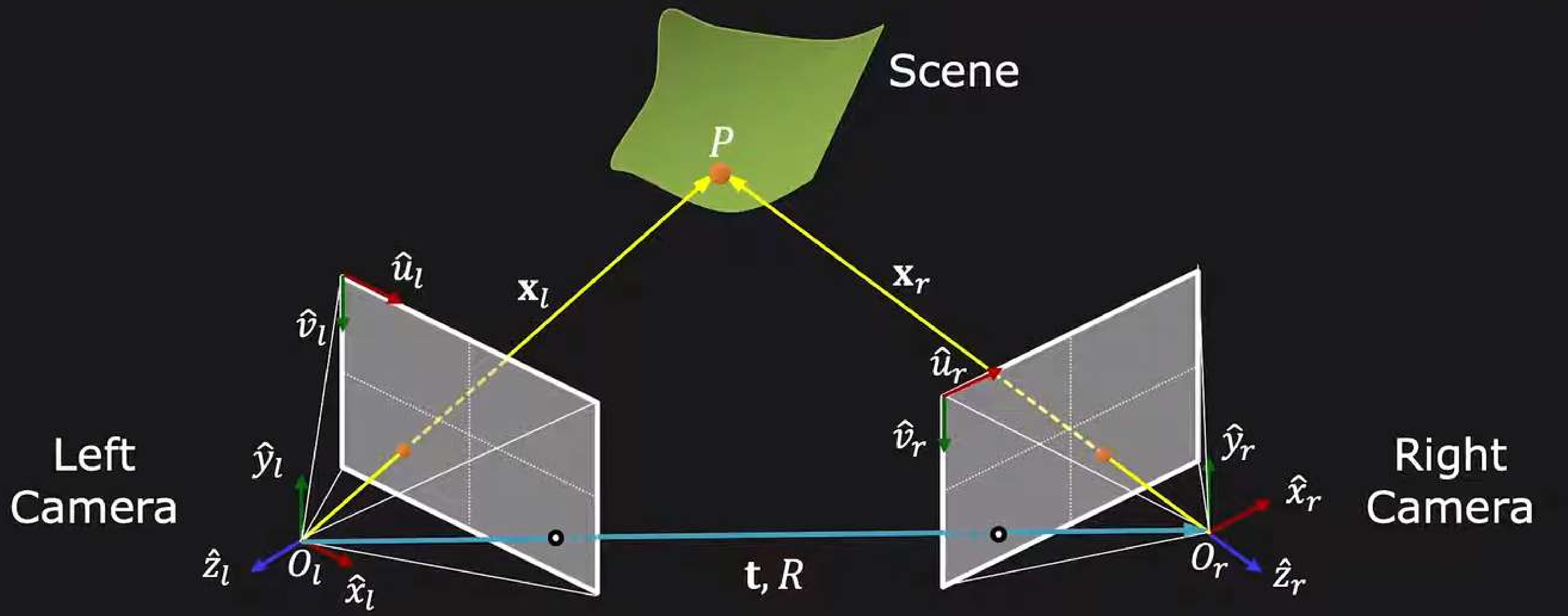
Corresponding scene points lie on the epipolar lines.
Finding correspondence is a 1D search.

Uncalibrated Stereo



1. Assume Camera Matrix K is known for each camera
2. Find a few Reliable Corresponding Points
3. Find Relative Camera Position $\mathbf{t}$ and Orientation R
4. Find Dense Correspondence
5. Compute Depth using Triangulation

Computing Depth



Given the intrinsic parameters, the projections of scene point on the two image sensors are:

$$\begin{bmatrix} u_l \\ v_l \\ 1 \end{bmatrix} \equiv \begin{bmatrix} f_x^{(l)} & 0 & o_x^{(l)} & 0 \\ 0 & f_y^{(l)} & o_y^{(l)} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_l \\ y_l \\ z_l \\ 1 \end{bmatrix} \qquad \begin{bmatrix} u_r \\ v_r \\ 1 \end{bmatrix} \equiv \begin{bmatrix} f_x^{(r)} & 0 & o_x^{(r)} & 0 \\ 0 & f_y^{(r)} & o_y^{(r)} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_r \\ y_r \\ z_r \\ 1 \end{bmatrix}$$

Computing Depth

Left Camera Imaging Equation

$$\begin{bmatrix} u_l \\ v_l \\ 1 \end{bmatrix} \equiv \begin{bmatrix} f_x^{(l)} & 0 & o_x^{(l)} & 0 \\ 0 & f_y^{(l)} & o_y^{(l)} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \boxed{\begin{bmatrix} x_l \\ y_l \\ z_l \\ 1 \end{bmatrix}}$$

Right Camera Imaging Equation

$$\begin{bmatrix} u_r \\ v_r \\ 1 \end{bmatrix} \equiv \begin{bmatrix} f_x^{(r)} & 0 & o_x^{(r)} & 0 \\ 0 & f_y^{(r)} & o_y^{(r)} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_r \\ y_r \\ z_r \\ 1 \end{bmatrix}$$

We also know the relative position and orientation between the two cameras.

$$\boxed{\begin{bmatrix} x_l \\ y_l \\ z_l \\ 1 \end{bmatrix}} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_r \\ y_r \\ z_r \\ 1 \end{bmatrix}$$

Computing Depth

Left Camera Imaging Equation:

$$\begin{bmatrix} u_l \\ v_l \\ 1 \end{bmatrix} \equiv \begin{bmatrix} f_x^{(l)} & 0 & o_x^{(l)} & 0 \\ 0 & f_y^{(l)} & o_y^{(l)} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_r \\ y_r \\ z_r \\ 1 \end{bmatrix}$$

$$\boxed{\tilde{\mathbf{u}}_l = P_l \, \tilde{\mathbf{x}}_r}$$

Right Camera Imaging Equation:

$$\begin{bmatrix} u_r \\ v_r \\ 1 \end{bmatrix} \equiv \begin{bmatrix} f_x^{(r)} & 0 & o_x^{(r)} & 0 \\ 0 & f_y^{(r)} & o_y^{(r)} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_r \\ y_r \\ z_r \\ 1 \end{bmatrix}$$

$$\boxed{\tilde{\mathbf{u}}_r = M_{int_r} \, \tilde{\mathbf{x}}_r}$$

Computing Depth

The imaging equations:

$$\tilde{\mathbf{u}}_r = M_r \, \tilde{\mathbf{x}}_r \qquad\qquad \tilde{\mathbf{u}}_l = P_l \, \tilde{\mathbf{x}}_r$$

$$\underbrace{\begin{bmatrix} u_r \\ v_r \\ 1 \end{bmatrix} \equiv \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix}}_{\text{Known}} \underbrace{\begin{bmatrix} x_r \\ y_r \\ z_r \\ 1 \end{bmatrix}}_{\text{Unknown}} \qquad \underbrace{\begin{bmatrix} u_l \\ v_l \\ 1 \end{bmatrix} \equiv \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix}}_{\text{Known}} \underbrace{\begin{bmatrix} x_r \\ y_r \\ z_r \\ 1 \end{bmatrix}}_{\text{Unknown}}$$

Rearranging the terms:

$$\begin{bmatrix} u_r m_{31} - m_{11} & u_r m_{32} - m_{12} & u_r m_{33} - m_{13} \\ v_r m_{31} - m_{21} & v_r m_{32} - m_{22} & v_r m_{33} - m_{23} \\ u_l p_{31} - p_{11} & u_l p_{32} - p_{12} & u_l p_{33} - p_{13} \\ v_l p_{31} - p_{21} & v_l p_{32} - p_{22} & v_l p_{33} - p_{23} \end{bmatrix} \begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} = \begin{bmatrix} m_{14} - m_{34} \\ m_{24} - m_{34} \\ p_{14} - p_{34} \\ p_{24} - p_{34} \end{bmatrix}$$

Computing Depth: Least Squares Solution

$$\begin{bmatrix} u_r m_{31} - m_{11} & u_r m_{32} - m_{12} & u_r m_{33} - m_{13} \\ v_r m_{31} - m_{21} & v_r m_{32} - m_{22} & v_r m_{33} - m_{23} \\ u_l p_{31} - p_{11} & u_l p_{32} - p_{12} & u_l p_{33} - p_{13} \\ v_l p_{31} - p_{21} & v_l p_{32} - p_{22} & v_l p_{33} - p_{23} \end{bmatrix} \begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} = \begin{bmatrix} m_{14} - m_{34} \\ m_{24} - m_{34} \\ p_{14} - p_{34} \\ p_{24} - p_{34} \end{bmatrix}$$

$A_{4\times3}$ (Known) $\quad$ $\mathbf{x}_r$ (Unknown) $\quad$ $\mathbf{b}_{4\times1}$ (Known)

Find least squares solution using pseudo-inverse:

$$A\mathbf{x}_r = \mathbf{b}$$

$$A^T A \mathbf{x}_r = A^T \mathbf{b}$$

$$\boxed{\mathbf{x}_r = (A^T A)^{-1} A^T \mathbf{b}}$$

3D Reconstruction with Internet Images

St. Peter's Basilica (1275 Images)

[Snavely 2006]