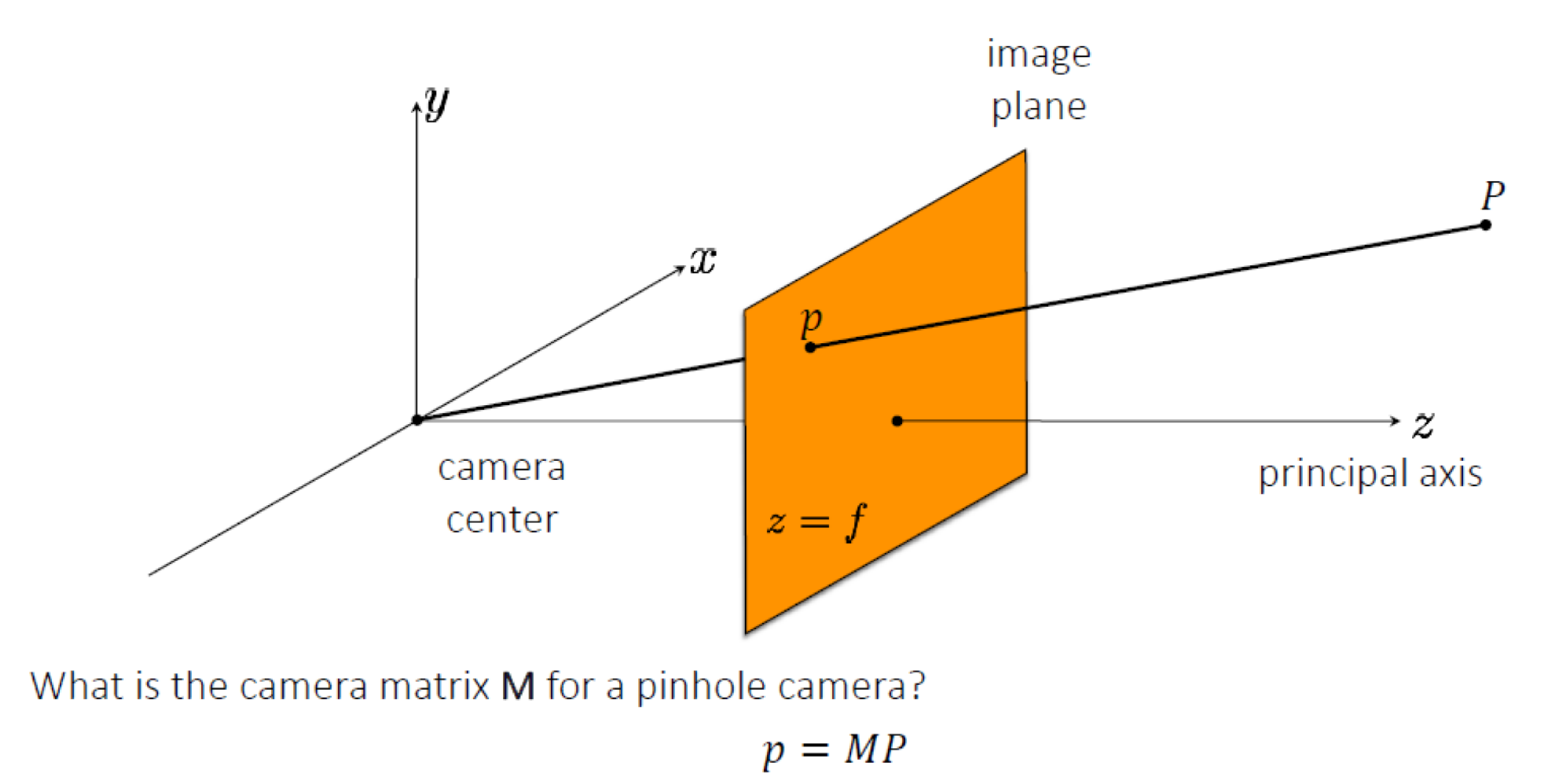
Tutorial 10 - Camera Calibration & Epipolar Geometry

Agenda

- Camera Model
- Camera Calibration
  - Estimating M
  - Chessboard Demo
  - Homography Quiz
- Epipolar Geometry
  - Essential Matrix
  - Fundamental Matrix
  - Fundamental Demo

Camera Model

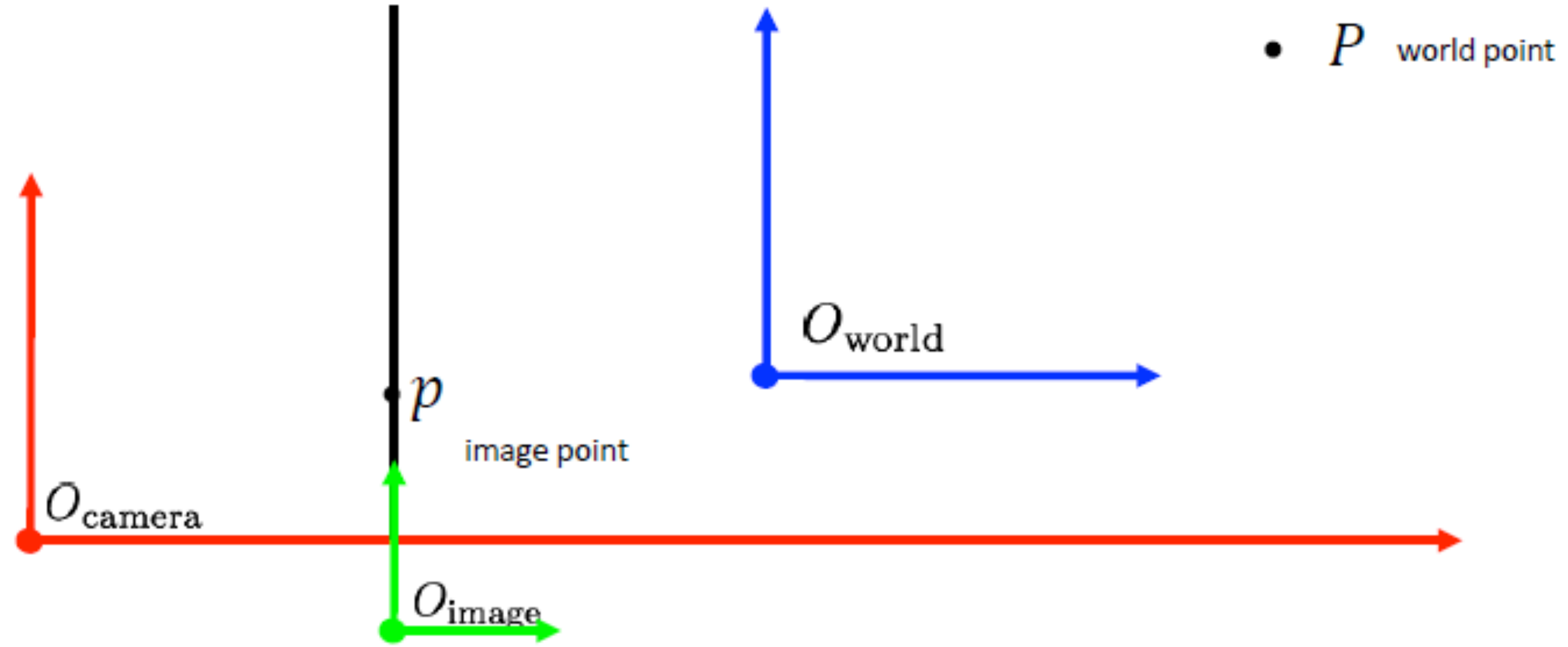
The (rearranged) pinhole camera



- A 3D world point P is projected by the camera matrix M to the 2D image point p

Camera Model

In general, there are *three*, generally different, coordinate systems.



We need to know the transformations between them.

- M is a 3×4 matrix comprised of two sets of parameters: Intrinsic and Extrinsic. What is the decomposed structure of M?

Camera Model

$$\mathrm{M} = K\,[R|t]$$

$$\mathrm{M} = \underbrace{\begin{bmatrix} f & 0 & m_x \\ 0 & f & m_y \\ 0 & 0 & 1 \end{bmatrix}}_{\text{intrinsic parameters}} \underbrace{\left[\begin{array}{ccc:c} r_1 & r_2 & r_3 & t_1 \\ r_4 & r_5 & r_6 & t_2 \\ r_7 & r_8 & r_9 & t_3 \end{array}\right]}_{\text{extrinsic parameters}}$$

$$\underbrace{\mathbf{R} = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix}}_{\text{3D rotation}} \qquad \underbrace{\mathbf{t} = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}}_{\text{3D translation}}$$

- M is a 3×4 matrix comprised of two sets of parameters: Intrinsic and Extrinsic. What is the decomposed structure of M?

Camera Model

- How many degress of freedom so far?
- And after switching f with f_x and f_y and adding skew s?

Camera Calibration

- Estimation of M
- Separating Extrinsic and Intrinsic Parameters

Geometric Camera Calibration: Estimating M

- Given a set of matched points $\{P_i, p_i\}$, we want to estimate M
  - Use the camera model: $P_i = Mp_i$
  - Where did we get such matched points?
- Same trick as in the Homography tutorial → switch to row-wise representation of the unknowns:

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} - & m_1^T & - \\ - & m_2^T & - \\ - & m_3^T & - \end{bmatrix} P$$

Estimating M

- Resulting equation for x and y in heterogeneous coordinates:

$$\tilde{x} = \frac{m_1^T P}{m_3^T P}, \tilde{y} = \frac{m_2^T P}{m_3^T P}$$

- Rearranging to solve for m_i:

$$m_1^T P - \tilde{x} m_3^T P = 0$$
$$m_2^T P - \tilde{y} m_3^T P = 0$$

- What is the dimension of $m_i^T P$

Estimating M

- Rearrange into a matrix for N_p points:

$$\begin{bmatrix} P_i^T & 0^T & -\tilde{x}_i P_i^T \\ 0^T & P_i^T & -\tilde{y}_i P_i^T \\ \vdots & \vdots & \vdots \\ P_{N_p}^T & 0^T & -\tilde{x}_{N_p} P_{N_p}^T \\ 0^T & P_{N_p}^T & -\tilde{y}_{N_p} P_{N_p}^T \end{bmatrix} \begin{bmatrix} | \\ m_1 \\ | \\ m_2 \\ | \\ m_3 \\ | \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \leftrightarrow Am = 0$$

- What are the dimensions? How much points N_p do we need?

Estimating M

- boils down to the problem:

$$\hat{m} = \arg\min_{m} ||Am||^2 \; s.t. ||m||^2 = 1$$

- Solution via SVD of $A = U\Sigma V^T$:
- $\hat{m}$ is the column of V corresponding to the smallest eigen-value.
- How about separating M to $K\left[R|t\right]$?

Decomposition of M to K, R & t

- rewrite M:

$$M = K\left[R|t\right] = K\left[R|-Rc\right] = \left[N|-Nc\right]$$

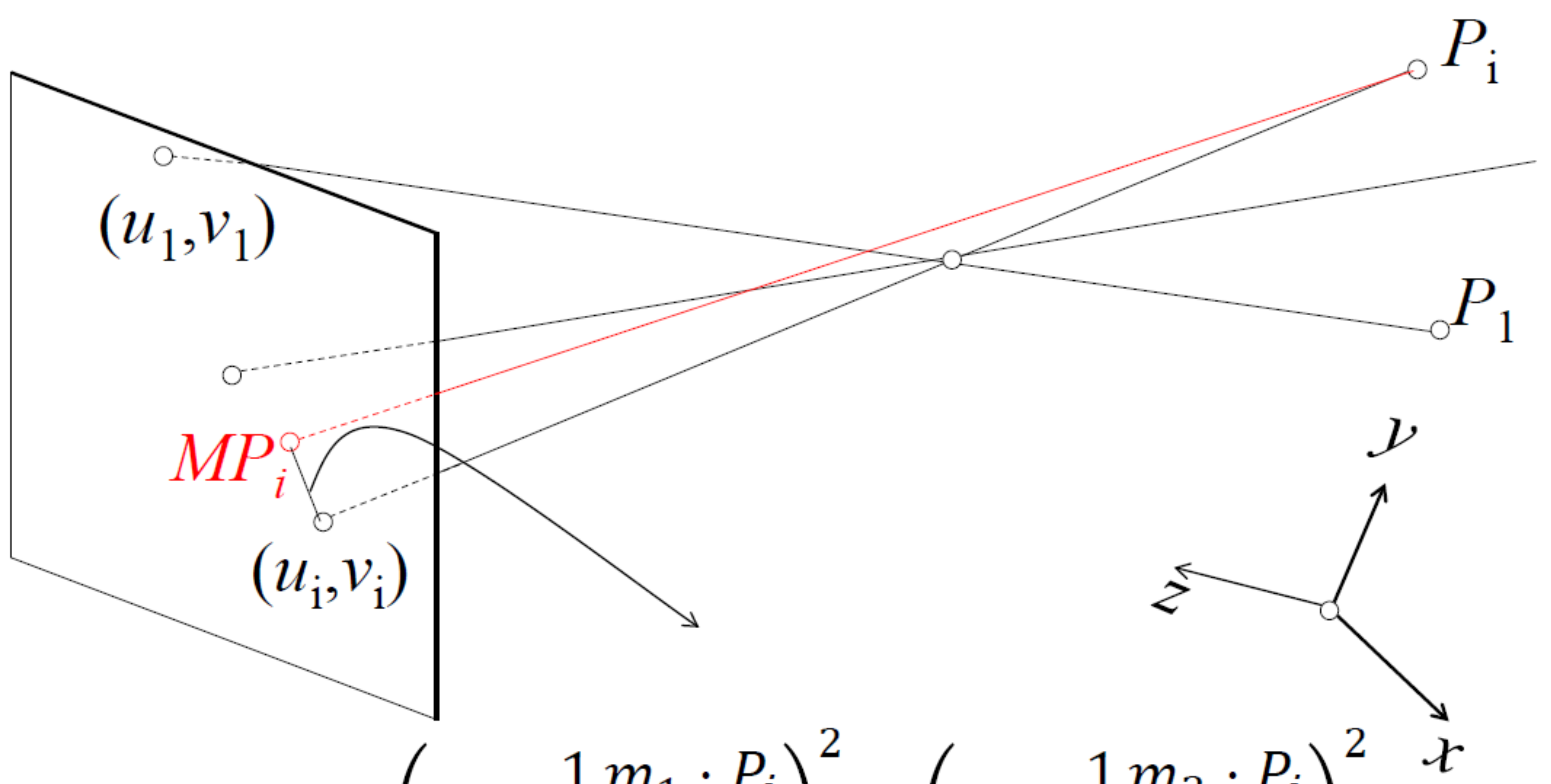
- c can be found via SVD of M due to the relation:

$$Mc = 0$$

- Then N can be further decomposed into $N = KR$:
  - How? using <u>QR</u> decomposition because K is upper triangular and R is orthogonal
- However..
  - Does not take into account noise, radial distortions, hard to impose prior knowledge (e.g. f), etc.
  - Solution?

Minimize reprojection error

Minimizing reprojection error with radial distortion



Add distortions to reprojection error:

$$\left(u_i - \frac{1}{\lambda}\frac{m_1 \cdot P_i}{m_3 \cdot P_i}\right)^2 + \left(v_i - \frac{1}{\lambda}\frac{m_2 \cdot P_i}{m_3 \cdot P_i}\right)^2$$

Where the radial distortion model is: $\lambda = 1 + k_1 r^2 + k_2 r^4 + k_3 r^6$

Minimize reprojection error

- Radial distortion is multiplicative:

$$x_{rad} = x\left[1 + k_1 r^2 + k_2 r^4 + k_3 r^6\right]$$

$$y_{rad} = y\left[1 + k_1 r^2 + k_2 r^4 + k_3 r^6\right]$$

- Usually we also include tangential distortion:

$$x_{tan} = x + \left[2p_1 xy + p_2\left(r^2 + 2x^2\right)\right]$$

$$y_{tan} = y + \left[p_1\left(r^2 + 2y^2\right) + 2p_2 xy\right]$$

- We end up with 5 parameters to estimate:

$$\text{distortion coefficients} = \left[k_1, k_2, k_3, p_1, p_2\right]^T$$

♜ Chessboard Calibration in OpenCV

- Take a notebook and paste a chesspattern
- Capture this pattern from several angles and positions
- Calibrate using OpenCV

♜ Chessboard Calibration in OpenCV

- Getting the 3D to 2D points correspondences from a known planar object
- Chessboard has fixed distances between squares known apriori
- Camera static and chessboard moves ↔ chessboard static and camera moves
- Camera moves ↔ Extrinsic parameters in each frame change
- Therefore we got the matches of real world points and camera points $\left\{P_i, p_i\right\}_{I=1}^{N}$!
- $P_i = \left[X_i, Y_i, Z_i = 0\right]$, where, X_i, Y_i set by periodicity of the chessbaord
  - $p_i = \left[x_i, y_i\right]$, detected corners in the image

Homography Quiz

Quiz 1

Prove that a 3×3 homography transform H is sufficient to describe the mapping between a planar 3D object and a camera, i.e. point matches of the form $\{p_i, P_i\}$, where $p_i = [x_i, y_i, w_i]^T$ and $P_i = [X_i, Y_i, Z_i, 1]^T$ satisfying $aX_i + bY_i + cZ_i + d = 0.$

Homography Quiz

Quiz 1

- In this special case of a **planar scene**, we do not need the full 3×4 camera matrix M, and we can make due with a 3×3 homography matrix H. The proof is relatively straight forward, and rely on the following observation:
- Since the points in 3D lie on a plane: $aX + bY + cZ + d = 0$
- we can switch sides and write down the plane equation for Z, such that

$$P_i = \left[X_i, Y_i, -\frac{a}{c}X_i - \frac{b}{c}Y_i - \frac{d}{c}, 1\right]^T.$$

Homography Quiz

Quiz 1

- This leads to the main conclusion that the 4D homogeneous coordinates are redundant and can be written down by a 3D homogeneous coordinates:

$$\begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{a}{c} & -\frac{b}{c} & -\frac{d}{c} \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ 1 \end{bmatrix}$$

Homography Quiz

Quiz 1

- Now coming back to the general camera matrix M how can we conclude it can be reduced to a homography?
  - The answer lies in simplifying the matrix-vector product.

Homography Quiz

Quiz 1

- A general M satisfy the relation:

$$\begin{bmatrix} x_i \\ y_i \\ w_i \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix}$$

- Assuming the scene is planar, we plug in the plane equation for $Z_i = -\frac{a}{c}X_i - \frac{b}{c}Y_i - \frac{d}{c}$:

$$\begin{bmatrix} x_i \\ y_i \\ w_i \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ -\frac{a}{c}X_i - \frac{b}{c}Y_i - \frac{d}{c} \\ 1 \end{bmatrix}$$

Homography Quiz
Quiz 1

- Now, let us look more closely at the formula for x_i:

$$x_i = m_{11}X_i + m_{12}Y_i + m_{13}\left(-\frac{a}{c}X_i - \frac{b}{c}Y_i - \frac{d}{c}\right) + m_{14}$$

- The terms can be rearranged and written as:

$$x_i = \left(m_{11} - \frac{a}{c}m_{13}\right)X_i + \left(m_{12} - \frac{b}{c}m_{13}\right)Y_i + \left(m_{14} - m_{13}\frac{d}{c}\right)$$

Homography Quiz
Quiz 1

- Similarly, this can be done for y_i and w_i:

$$y_i = \left(m_{21} - \frac{a}{c} m_{23} \right) X_i + \left(m_{22} - \frac{b}{c} m_{23} \right) Y_i + \left(m_{24} - m_{23} \frac{d}{c} \right)$$

$$w_i = \left(m_{31} - \frac{a}{c} m_{33} \right) X_i + \left(m_{32} - \frac{b}{c} m_{33} \right) Y_i + \left(m_{34} - m_{33} \frac{d}{c} \right)$$

Homography Quiz

Quiz 1

- Therefore, rewriting this in matrix form we get the following relation:

$$\begin{bmatrix} x_i \\ y_i \\ w_i \end{bmatrix} = \begin{bmatrix} m_{11} - \frac{a}{c}m_{13} & m_{12} - \frac{b}{c}m_{13} & m_{14} - \frac{d}{c}m_{13} \\ m_{21} - \frac{a}{c}m_{23} & m_{22} - \frac{b}{c}m_{23} & m_{24} - \frac{d}{c}m_{23} \\ m_{31} - \frac{a}{c}m_{33} & m_{32} - \frac{b}{c}m_{33} & m_{34} - \frac{d}{c}m_{33} \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ 1 \end{bmatrix}$$

- Meaning, the 3D point P_i can be indeed reduced to a 3D homogeneous vector $\left[X_i, Y_i, 1\right]^T$, and is related to the 2D point in the image $p_i = \left[x_i, y_i, w_i\right]$ through a 3×3 homography H that is a function of the entries in the general 3×4 camera matrix M, and the normal to the plane $\left[a, b, c, d\right]^T$.

Homography Quiz

Quiz 2

- Prove that there exists a homography H that satisfies:
- $p_1 \equiv Hp_2$between the 2D points (in homogeneous coordinates) p_1 and p_2 in the images of a plane P_i captured by two 3×4 camera projection matrices M_1 and M_2, respectively.
- The symbol $\equiv$ stands for equality *up to scale*.

- (Note: A degenerate case happens when the plan Π contains both cameras' centers, in which case there are infinite choices of H satisfying the equation. You can ignore this special case in your answer.)

Homography Quiz
Quiz 2

- Plane in 3D using homogeneous coordinates is given by:
- $n^T P = 0$ Where n, P are homogeneous vectors (4 numbers each) and n is the normal to the plane.
- Therefore, we can find a basis of 3 vectors u_1, u_2, u_3 in $\mathbb{R}$, such that each point on the plane is given by:

$$P = \sum_{i=1}^{3} \alpha_i u_i$$

- The projection of 3D point P to the j^{th} image point p_j is given by:

$$p_j = M_j P = \sum_{i=1}^{3} \alpha_i M_j u_i$$

Homography Quiz

Quiz 2

- If we denote $v_j^i = M_j u_i$ we get:

$$p_1 = \sum_{i=1}^{3} \alpha_i v_1^i$$

$$p_2 = \sum_{i=1}^{3} \alpha_i v_2^i$$

- Hence, the relation between the two points is a 3×3 matrix satisfying:

$$\begin{bmatrix} | & | & | \\ v_1^1 & v_1^2 & v_1^3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix} \begin{bmatrix} | & | & | \\ v_2^1 & v_2^2 & v_2^3 \\ | & | & | \end{bmatrix}$$

Homography Quiz

Quiz 2

- Relation to camera
- Recall that we have 11 degrees of freedom in M
- If all the calibration points are on a plane, we get at most 8 independent equations out of 4
- Any 5^{th} point will result in constraints that are linearly dependent on the constraints from the previous 4 pts on the
- Therefore, in estimating M we can't rely on a single image of the chessboard.

Homography Quiz

Quiz 3

- Prove that there exists a homography H that satisfies the equation $p_1 = Hp_2$, given two cameras separated by a pure rotation. That is, for camera 1, $p_1 = K_1 \left[I|0\right] P$, and for camera 2, $p_2 = K_2 \left[R|0\right] P$. Note that K_1 and K_2 are the 3×3 intrinsic matrices of the cameras and are different. I is 3×3 identity matrix, 0 is a 3×1 zero vector and P is a point in 3D space. R is the 3×3 rotation matrix of the camera.

Homography Quiz

Quiz 3

- Since the last column is zero, we can see that:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = R^{-1} K_2^{-1} p_2$$

- Substituting this in the second equation we get:

$$p_1 = K_1 R^{-1} K_2^{-1} p_2$$

- Therefore, the resulting homography is given by:

$$H = K_1 R^{-1} K_2^{-1}$$

Homography Quiz

Quiz 3

- Take away is 2 cameras differing only in rotation can't triangulate!
- Remember where this was useful?
  - Panorama stitching! (There we did not care about recovering depth)

Homography Quiz

Quiz 3

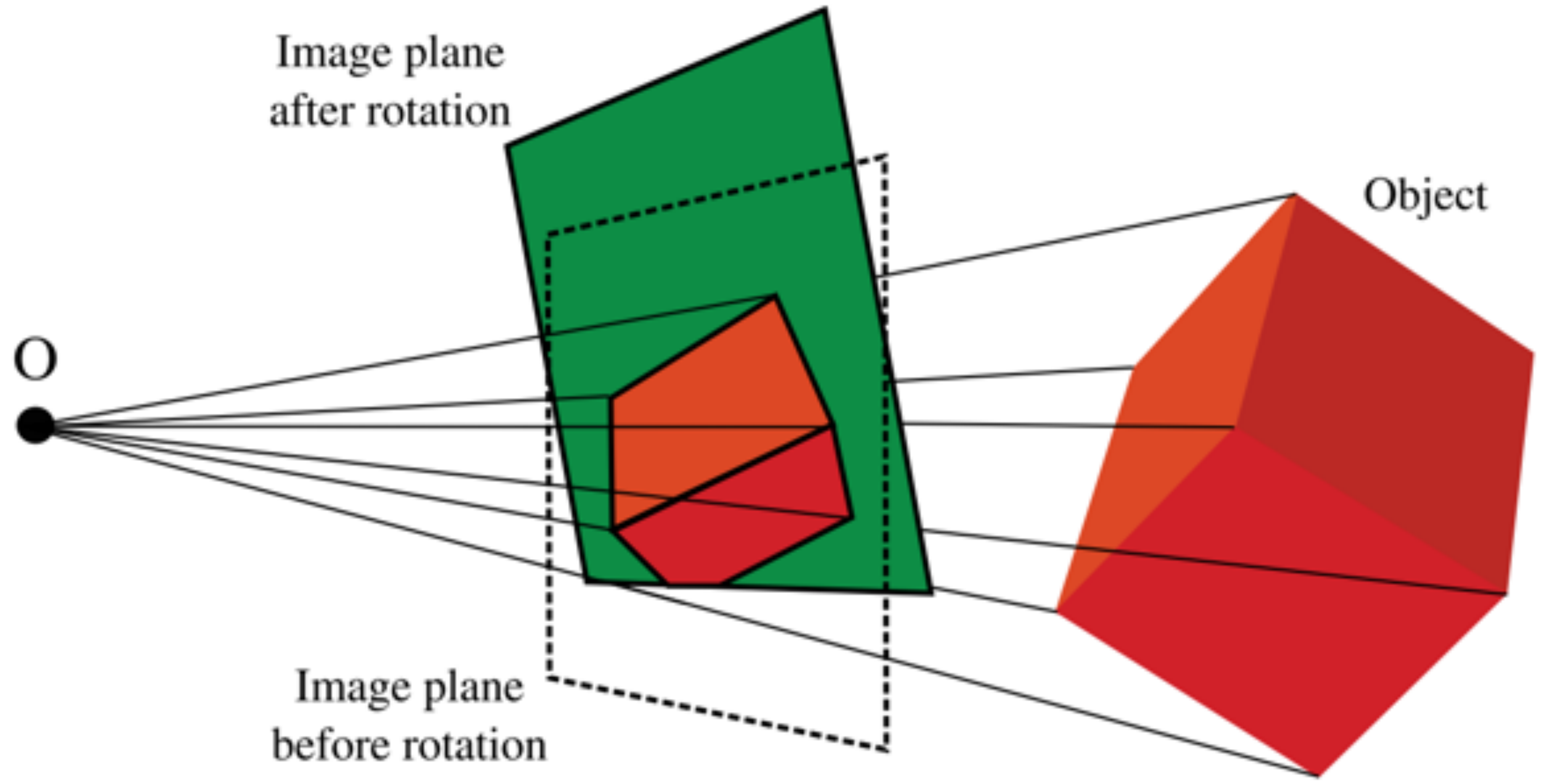


Figure 15.14 Images under pure camera rotation. When the camera rotates but does not translate, the bundle of rays remains the same, but is cut by a different plane. It follows that the two images are related by a homography.

Homography Quiz

Quiz 4

- Suppose that a camera is rotating about its center C, keeping the intrinsic parameters K constant. Let H be the homography that maps the view from one camera orientation to the view at a second orientation. Let θ be the angle of rotation between the two. Show that H^2 is the homography corresponding to a rotation of 2θ.

Homography Quiz

Quiz 4

- We have just shown that for such a scenario:

$$H_{2\to1} = K_1 R_\theta^{-1} K_2^{-1}$$

$$H_{1\to2} = K_2 R_\theta K_1^{-1}$$

- Applying the constraint $K_1 = K_2 \equiv K$ gets us:

$$H_{1\to2} = K R_\theta K^{-1}$$

- Applying $H_{1\to2}$ twice gets us:

$$H_{1\to2}^2 = K R_\theta K^{-1} K R_\theta K^{-1} = K R_\theta R_\theta K^{-1}$$

- Since $R_\theta R_\theta = R_{2\theta}$, we indeed get:

$$H_{1\to2}^2 = K R_{2\theta} K^{-1}$$

- Which is a homography that corresponds to a rotation of 2θ.

Homography Quiz

Quiz 5

- Prove that points on a single line do not uniquely constrain the homography H. In other words, prove that we need points on at least 2 different directions within the plane to estimate a homography reliably.

Homography Quiz

Quiz 5

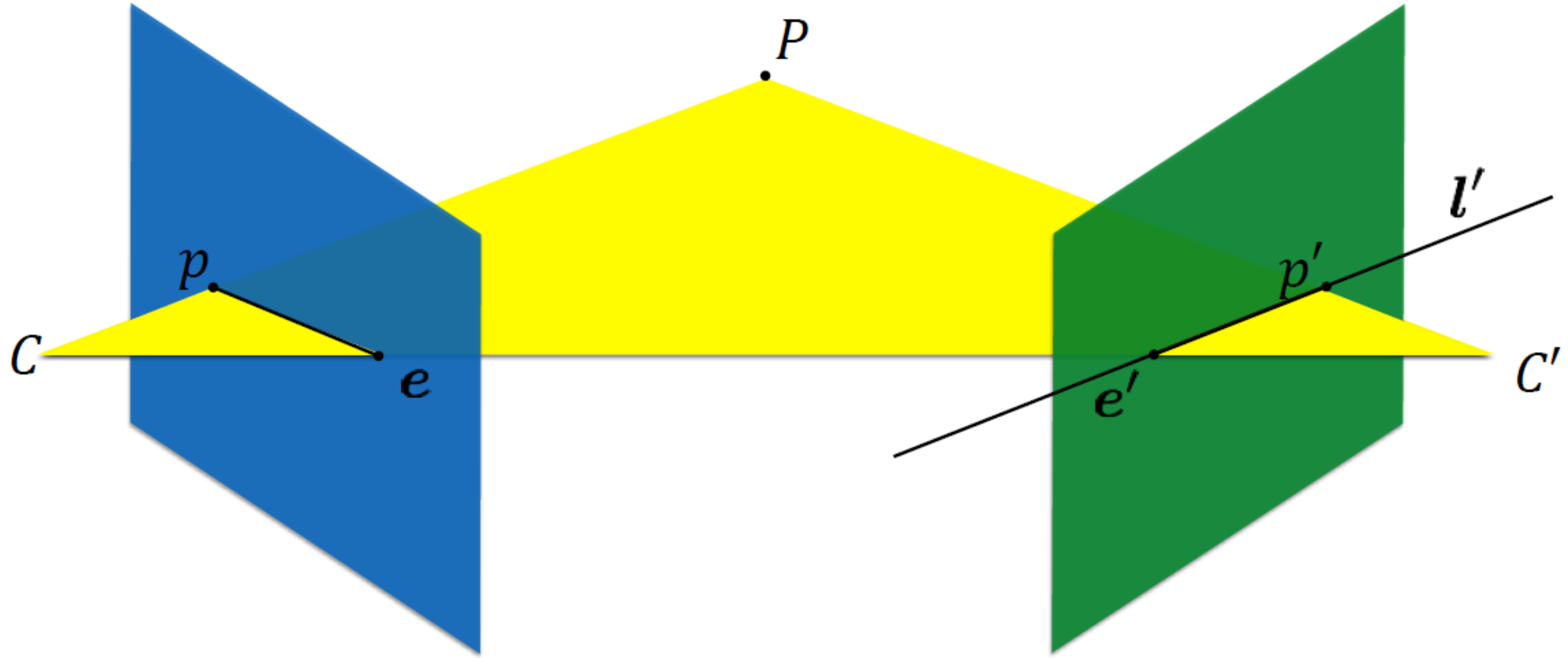
- If all points lie on a line there is a 3×1 vector l such the $l^T p = 0$ for all points p.
- Now suppose you found a homography matrix H such that $p'_j = Hp_j$, and yet all your points satisfy $l^T p_j = 0$.
- Then it is easy to see that for every 3×1 vector v the matrix $H' = H + vl^T$ will also satisfy $p'_j = H'p_j$
- * This implies that there is no unique H which explains points on the same line.

Epipolar Geometry

Epipolar Lingo

Given a point in one image,
multiplying by the **essential matrix** will tell us
the **epipolar line** in the second view.

$$\boxed{Ep = l'}$$



Epipolar Geometry

Epipolar Lingo

Essential Matrix vs Homography

What's the difference between the essential matrix and a homography?

They are both 3 x 3 matrices but …

$$Ep = l'$$

Essential matrix maps a **point** to a **line**

$$Hp = p'$$

Homography maps a **point** to a **point**

When can we use a homography? And when only an essential matrix?

Homography applies only for planer scenes

Epipolar Geometry

Epipolar Lingo

- Given by the equation $p'^T E p = 0$
- and also equals $E = R[dC]_\times, \quad E = [t]_\times R$, where $t = -RdC$

Properties:

$p'^T E p = 0$

- $p^T l = 0, \quad p'^T l' = 0$
- $l' = Ep, \quad l = E^T p'$
- $e'^T E = 0, \quad Ee = 0$

Fundamental Matrix

$$\widehat{p'}^T E \hat{p} = 0$$

The Essential matrix operates on image points expressed in
normalized coordinates
(points have been aligned (normalized) to camera coordinates)

$$\widehat{p'} = K'^{-1} p' \qquad \underset{\text{camera point}}{\hat{p}} = \underset{\text{image point}}{K^{-1} p}$$

Writing out the epipolar constraint in terms of image coordinates

$$p'^T K'^{-T} E K^{-1} p = 0$$

$$p'^T (K'^{-T} E K^{-1}) p = 0$$

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

The Fundamental matrix

SolutionFrom Spring 2019 Q4

- Explain how you can calculate F from the known K, R, and t of the camera.

The Fundamental matrix

SolutionFrom Spring 2019 Q4 Solution

- $F \triangleq K^{-T} E K'^{-1}$
- $E = [t]_\times R$
- $F = K^{-T}[t]_\times R K'^{-1}$

where, $[t]_\times = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}$

Fundamental Matrix Properties

Properties

- $p'^T F p = 0$
- $p^T l = 0, \quad p'^T l' = 0$
- $l' = Fp, \quad l = F^T p'$
- $e'^T F = 0, \quad Fe = 0$

The Fundamental matrix
(From Spring 2019 Q4)

- Explain shortly the steps of the algorithm for estimating F (when K, R, t are unknown). How many DOF does F have? How many sets of points are needed to find it?

The Fundamental matrix

(From Spring 2019 Q4) Solution

- Assume we are given 2D to 2D M matched image points:

$$\left\{p_i, p_i'\right\}_{i=1}^{M}$$

- Each correspondence should satisfy:

$$p_i^T F p_i' = 0 \leftrightarrow \begin{bmatrix} x_i & y_i & 1 \end{bmatrix}^T \begin{bmatrix} f_1 & f_2 & f_3 \\ f_4 & f_5 & f_6 \\ f_7 & f_8 & f_9 \end{bmatrix} \begin{bmatrix} x_i' \\ y_i' \\ 1 \end{bmatrix} = 0$$

- How to solve?
- The 8-point algorithm ↔ arrange into homogeneous linear equations and SVD..

The Fundamental matrix

(From Spring 2019 Q4) Solution

$$\begin{bmatrix} x_1x'_1 & x_1y'_1 & x_1 & y_1x'_1 & y_1y'_1 & y_1 & x'_1 & y'_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_Mx'_M & x_My'_M & x_M & y_Mx'_M & y_My'_M & y_M & x'_M & y'_M & 1 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \\ f_7 \\ f_8 \\ f_9 \end{bmatrix} = 0$$

8 DOF, solve **SVD** under the constraint $||f|| = 1$

The Fundamental matrix
From Spring 2019 Q4 Solution

- Under what conditions does $F = E$?
- Since $F \triangleq K'^{-T} E K^{-1}$, when $K' = K = I$ we'll get $F = I^{-1} E I^{-1} = E$

- note that $F = K' E K$, that's a hint that we want to constraint K and K'

The Fundamental matrix

From Spring 2019 Q4

Suppose we know that the camera planes are parallel (after rectification - i.e. they are only translated on the x axis). Does that information allow us to use less sets of matching points to find F?

The Fundamental matrix

From Spring 2019 Q4

The camera matrices are given by $M = K[I \mid 0]$ and $M' = K'[R \mid t]$, where

$$K = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}, K' = \begin{bmatrix} f' & 0 & 0 \\ 0 & f' & 0 \\ 0 & 0 & 1 \end{bmatrix}, R = I, t = \begin{bmatrix} t_x \\ 0 \\ 0 \end{bmatrix}$$

The Fundamental matrix

From Spring 2019 Q4

- We know that $F = K'^{-T}[t]_{\times}RK^{-1}$, so by plugging in the relevant parameters:

$$F = \begin{bmatrix} 1/f' & 0 & 0 \\ 0 & 1/f' & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_x \\ 0 & t_x & 0 \end{bmatrix} \begin{bmatrix} 1/f & 0 & 0 \\ 0 & 1/f & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$F = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_x/f' \\ 0 & -t_x/f & 0 \end{bmatrix}$$

The Fundamental matrix

From Spring 2019 Q4

- F now has only 2 DoFs, so we only need 2 points.
- Note that Even if we look at more complex intrinsic matrices, and 3 degrees of translation, we can still spare 3 DoFs (and 3 points) by not having a rotation matrix!

Questions?