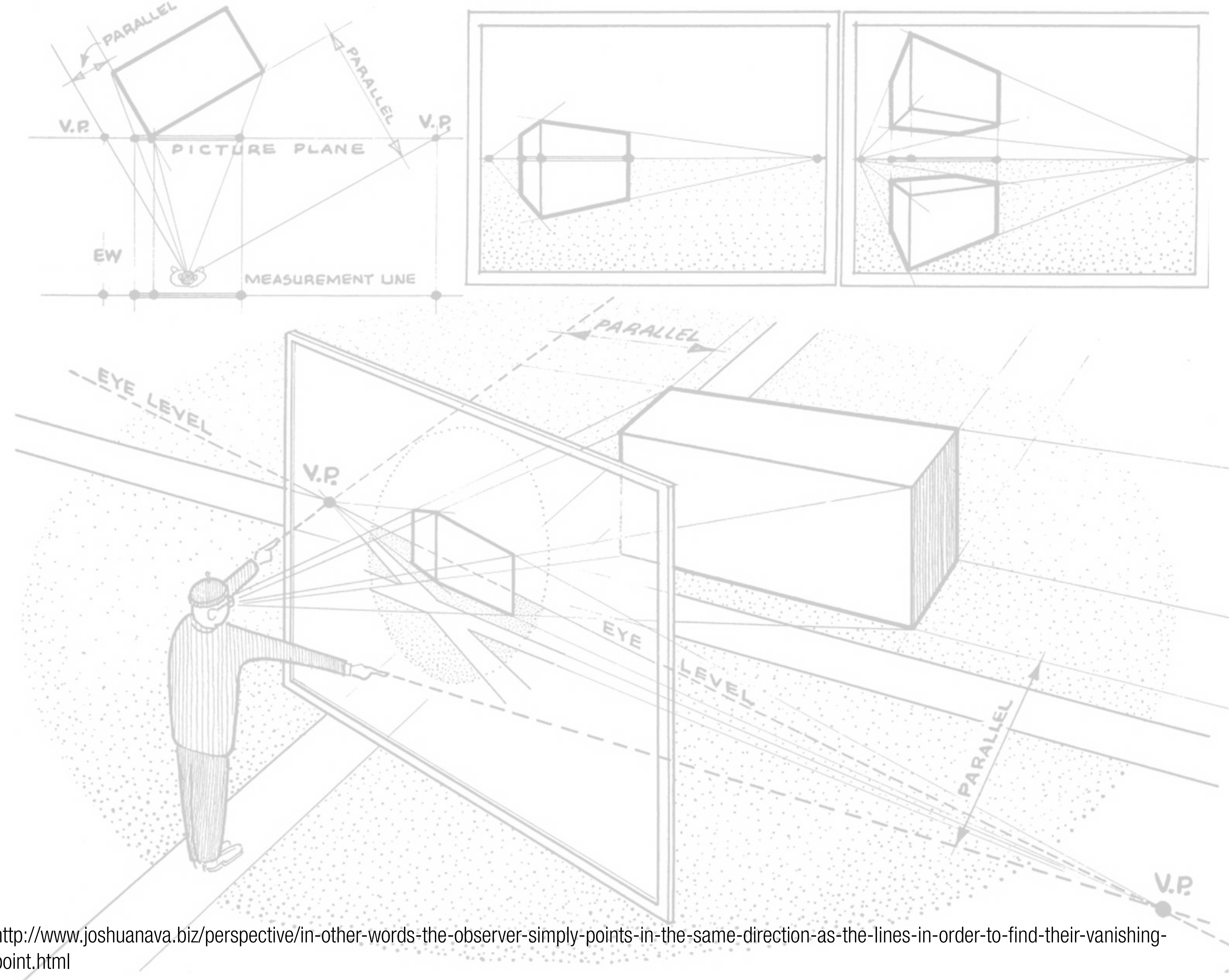
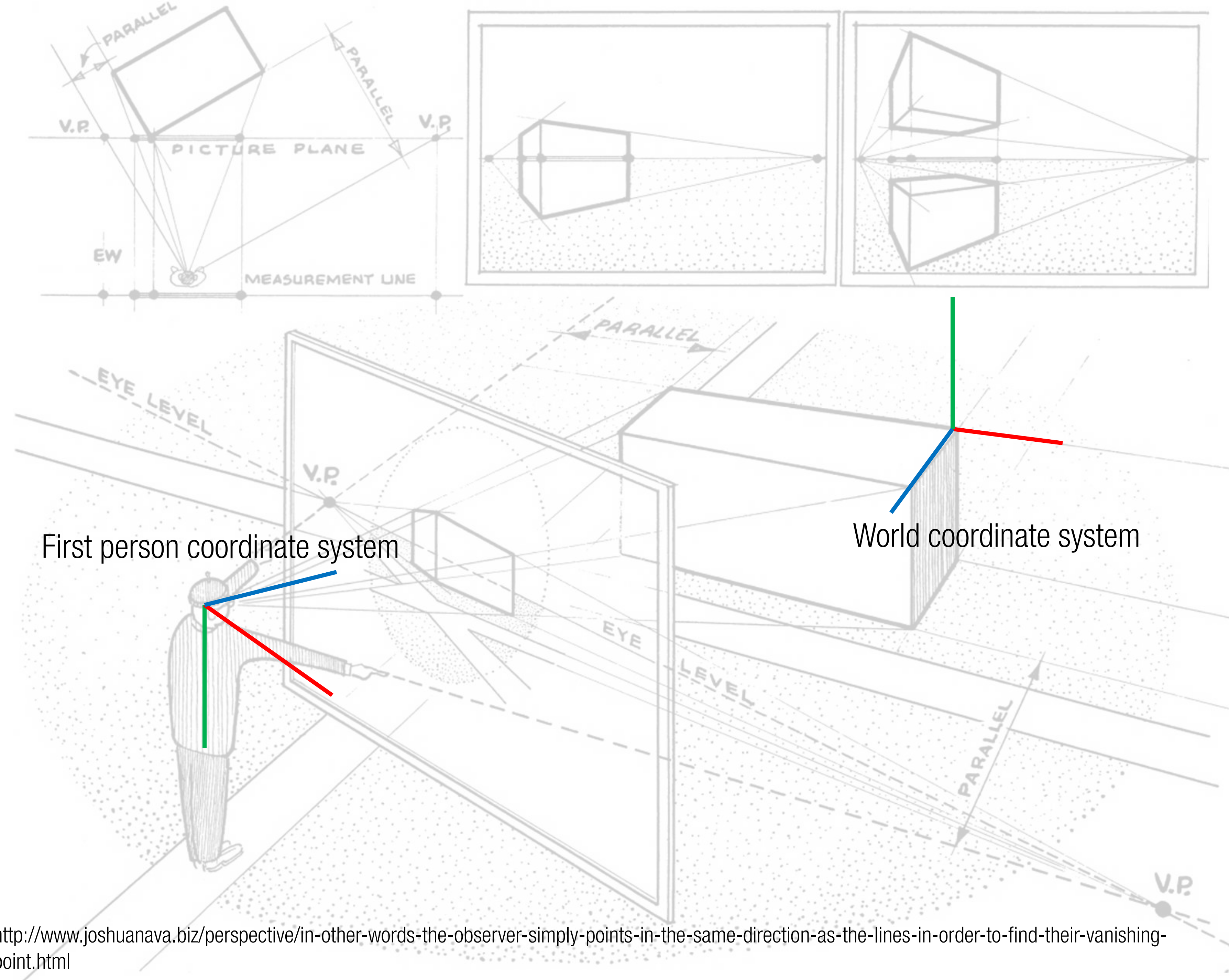
$$z\begin{bmatrix} u_{\text{img}} \\ v_{\text{img}} \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & s & p_x \\ & f_x & p_y \\ & & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

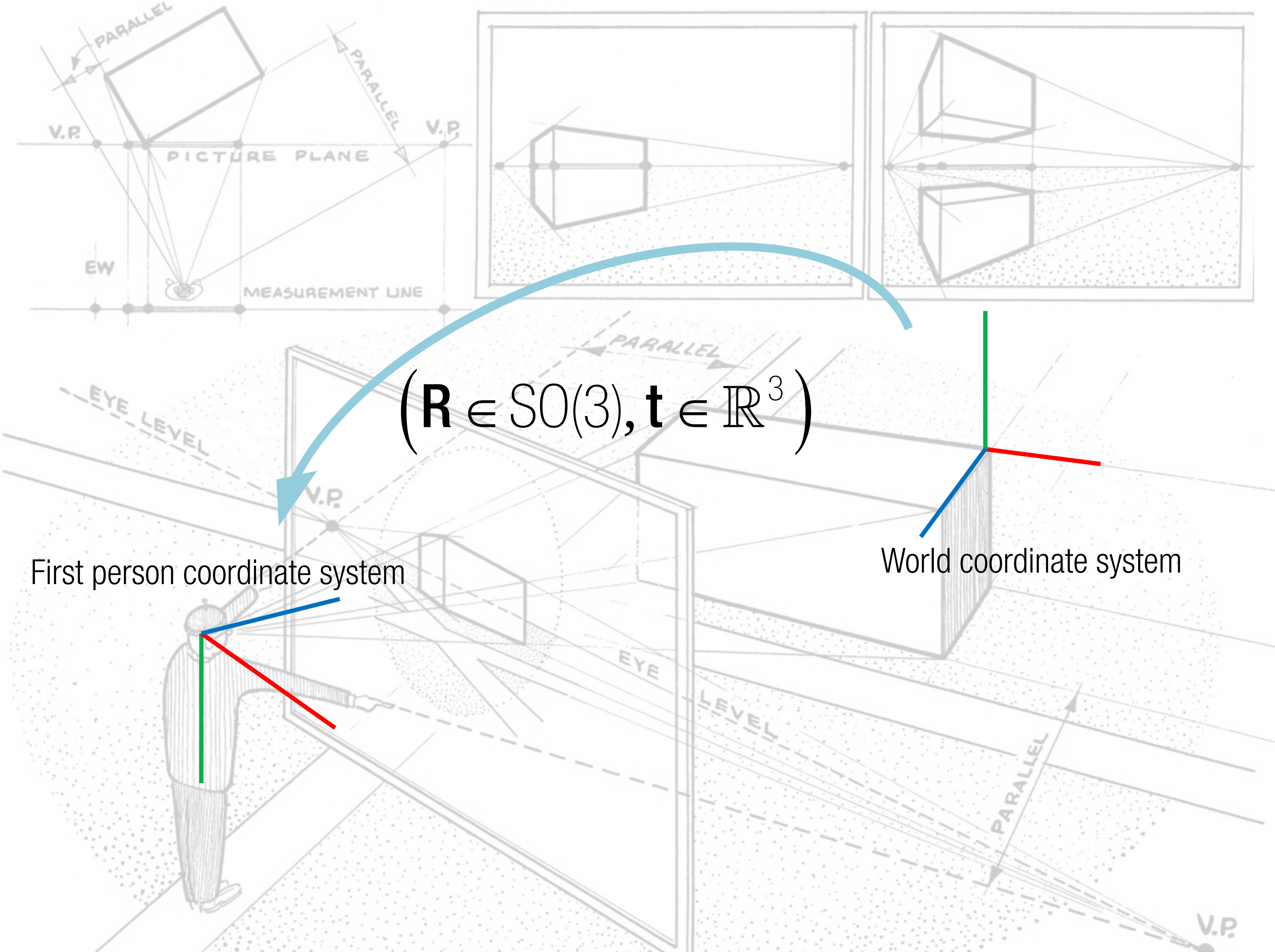
$\mathbf{x}$ $\mathbf{K}$ $\mathbf{R} \in \mathbb{R}^{3\times3}$ $\mathbf{t}$ $\mathbf{X}$



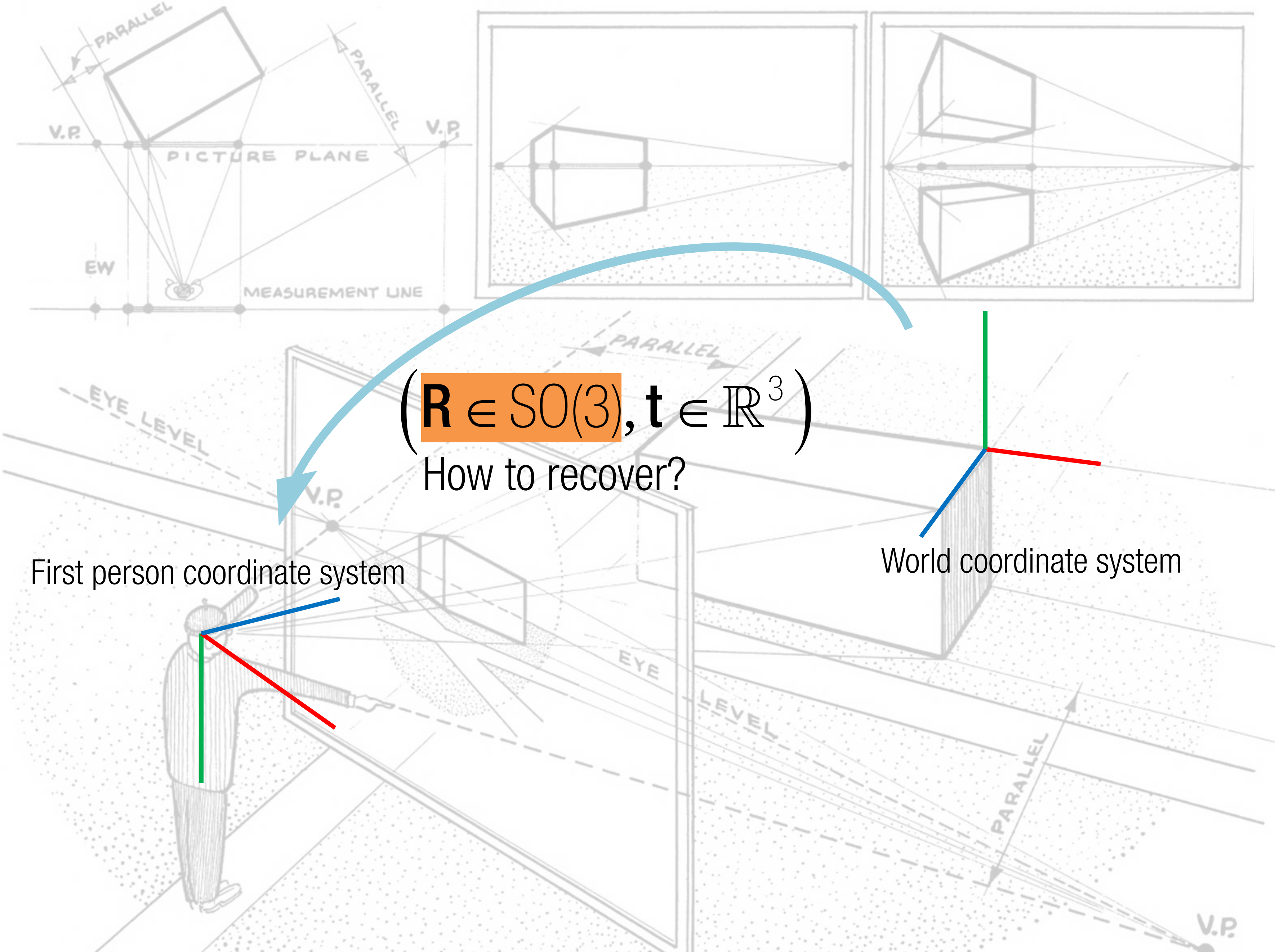
http://www.joshuanava.biz/perspective/in-other-words-the-observer-simply-points-in-the-same-direction-as-the-lines-in-order-to-find-their-vanishing-point.html



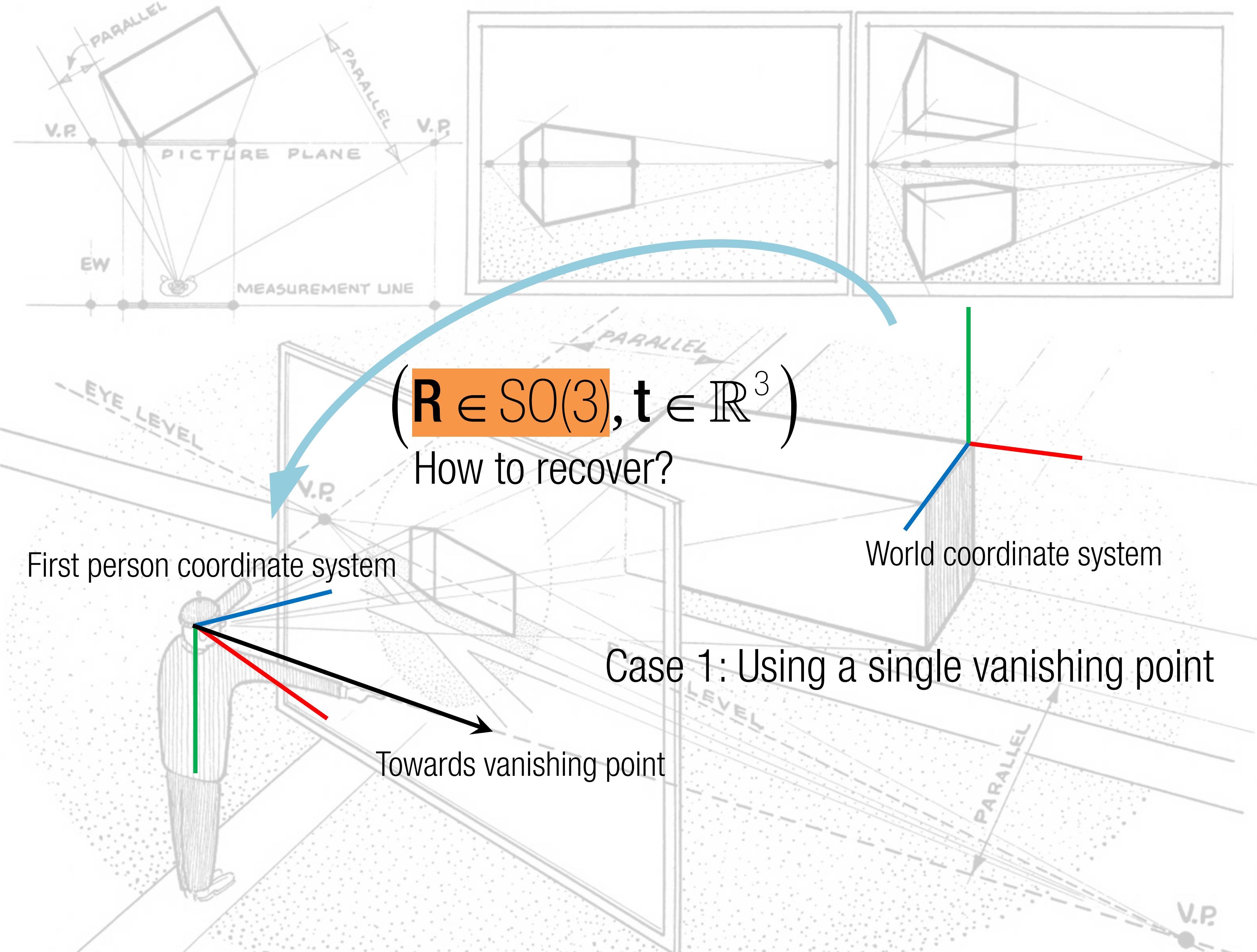
http://www.joshuanava.biz/perspective/in-other-words-the-observer-simply-points-in-the-same-direction-as-the-lines-in-order-to-find-their-vanishing-point.html



http://www.joshuanava.biz/perspective/in-other-words-the-observer-simply-points-in-the-same-direction-as-the-lines-in-order-to-find-their-vanishing-point.html



http://www.joshuanava.biz/perspective/in-other-words-the-observer-simply-points-in-the-same-direction-as-the-lines-in-order-to-find-their-vanishing-point.html

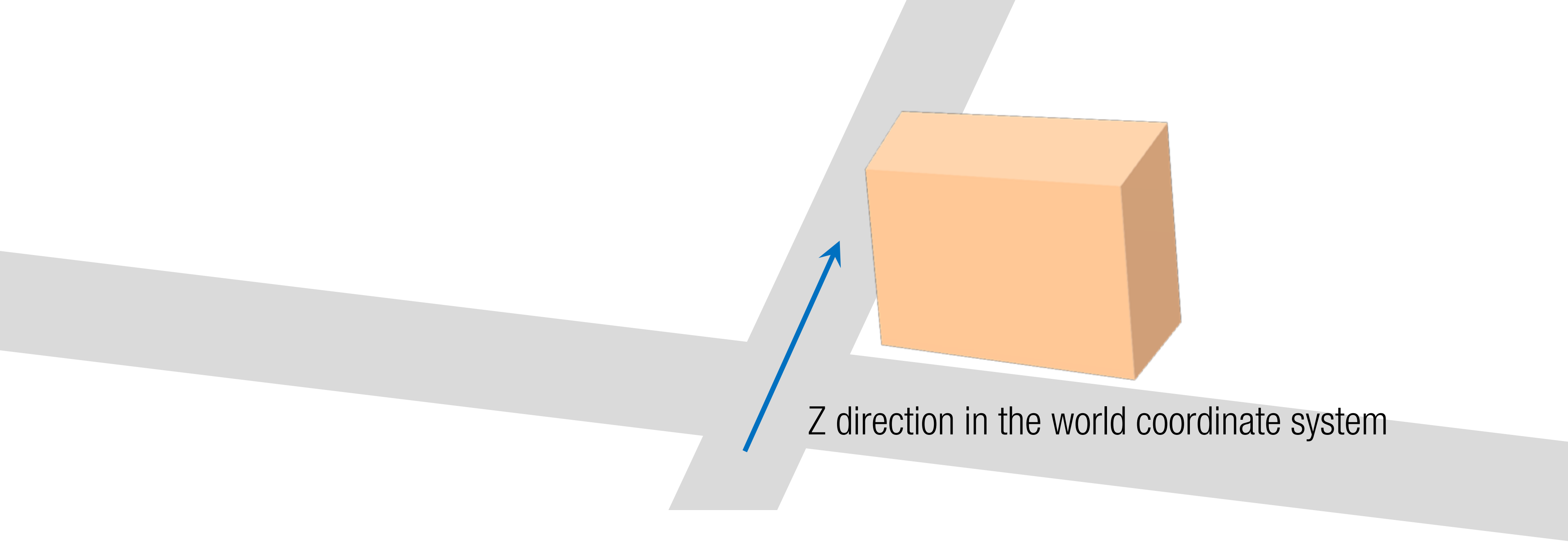


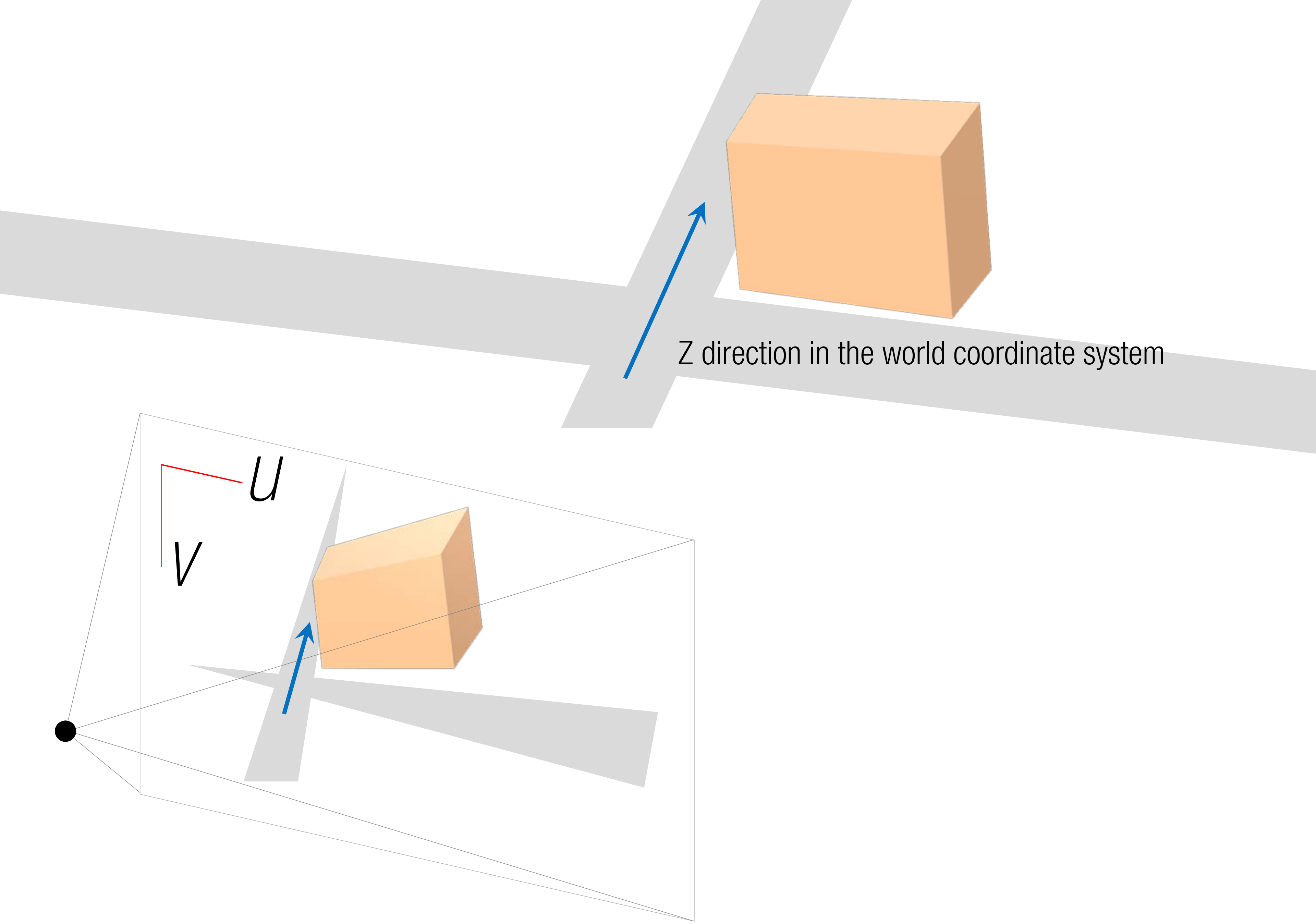
$\left(\boxed{\mathbf{R} \in SO(3)}, \mathbf{t} \in \mathbb{R}^3\right)$

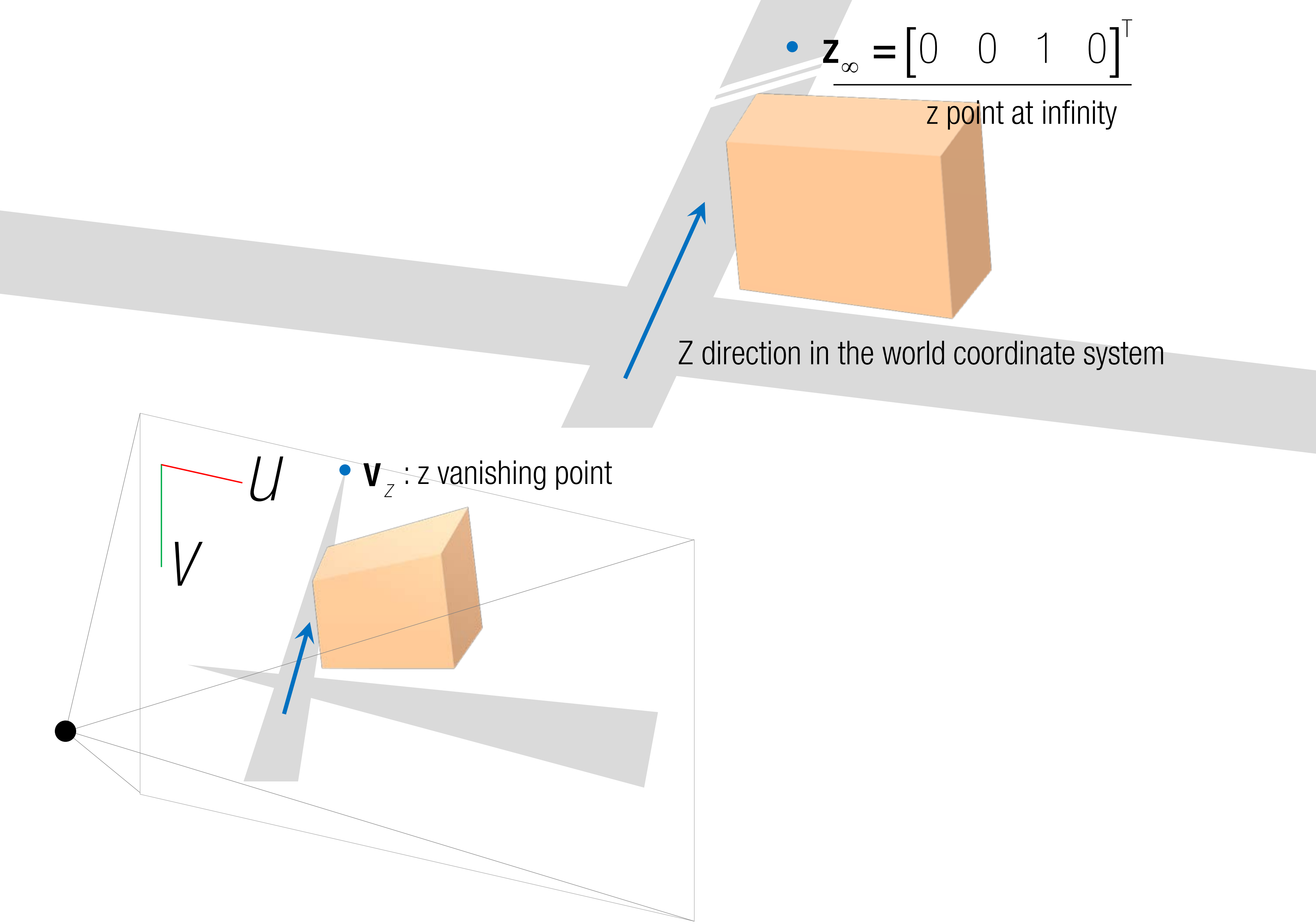
How to recover?

Case 1: Using a single vanishing point

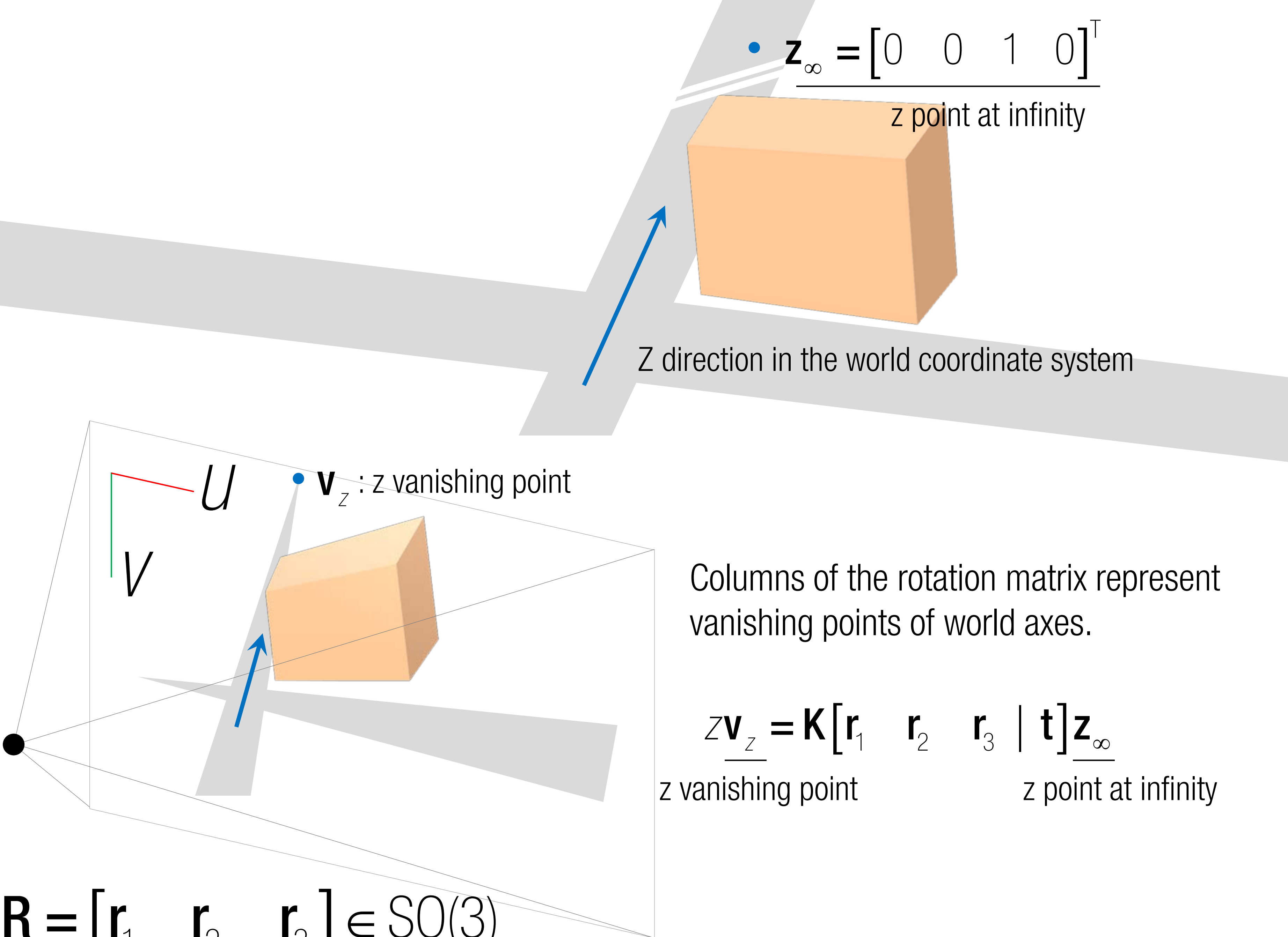
http://www.joshuanava.biz/perspective/in-other-words-the-observer-simply-points-in-the-same-direction-as-the-lines-in-order-to-find-their-vanishing-point.html


Z direction in the world coordinate system


Z direction in the world coordinate system
U
V



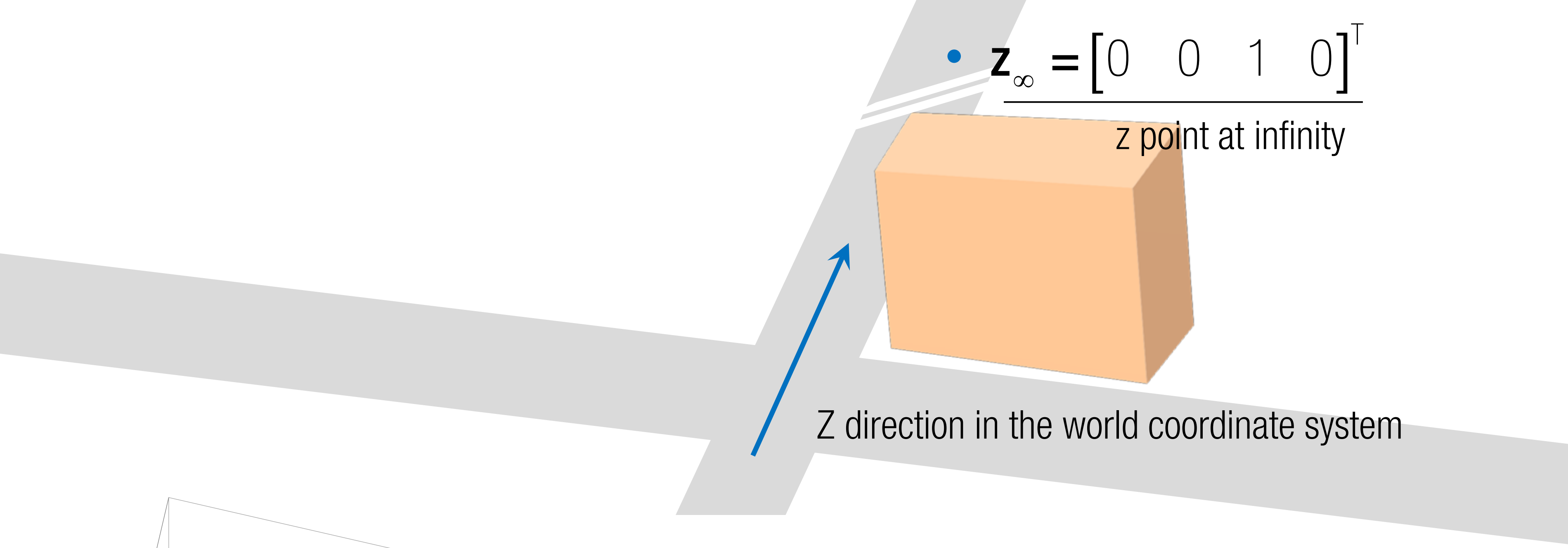
$\mathbf{z}_\infty = [0 \quad 0 \quad 1 \quad 0]^T$
z point at infinity
Z direction in the world coordinate system
U
V
$\mathbf{v}_Z$: z vanishing point

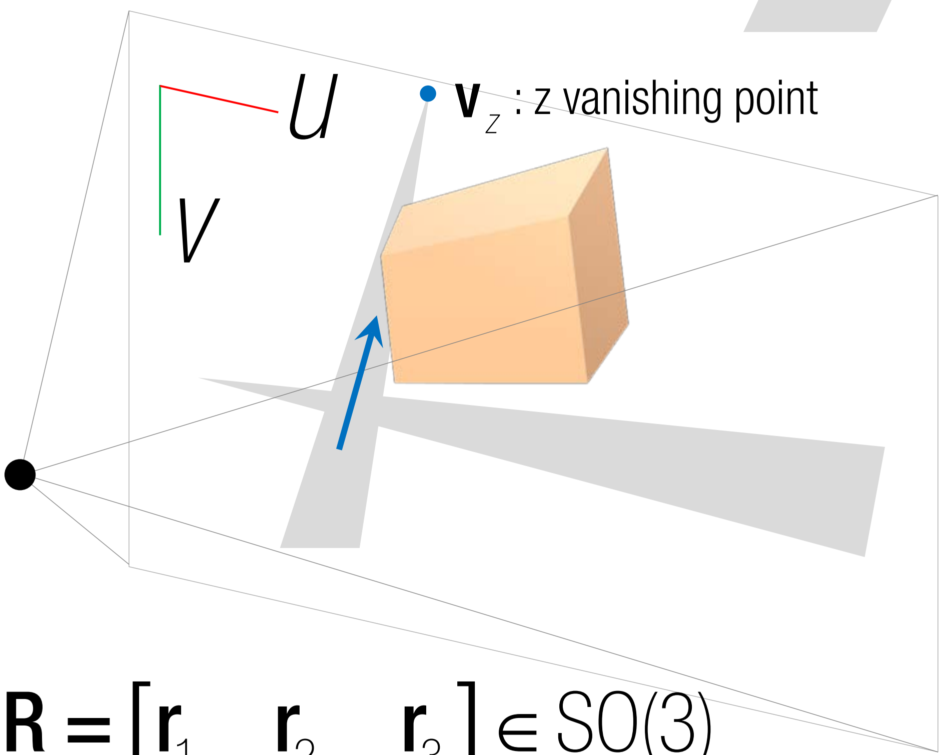


Columns of the rotation matrix represent vanishing points of world axes.

$$Z\underbrace{\mathbf{v}_Z}_{\text{z vanishing point}} = \mathbf{K}\begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 & | & \mathbf{t} \end{bmatrix}\underbrace{\mathbf{z}_\infty}_{\text{z point at infinity}}$$

$$\mathbf{R} = \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 \end{bmatrix} \in SO(3)$$

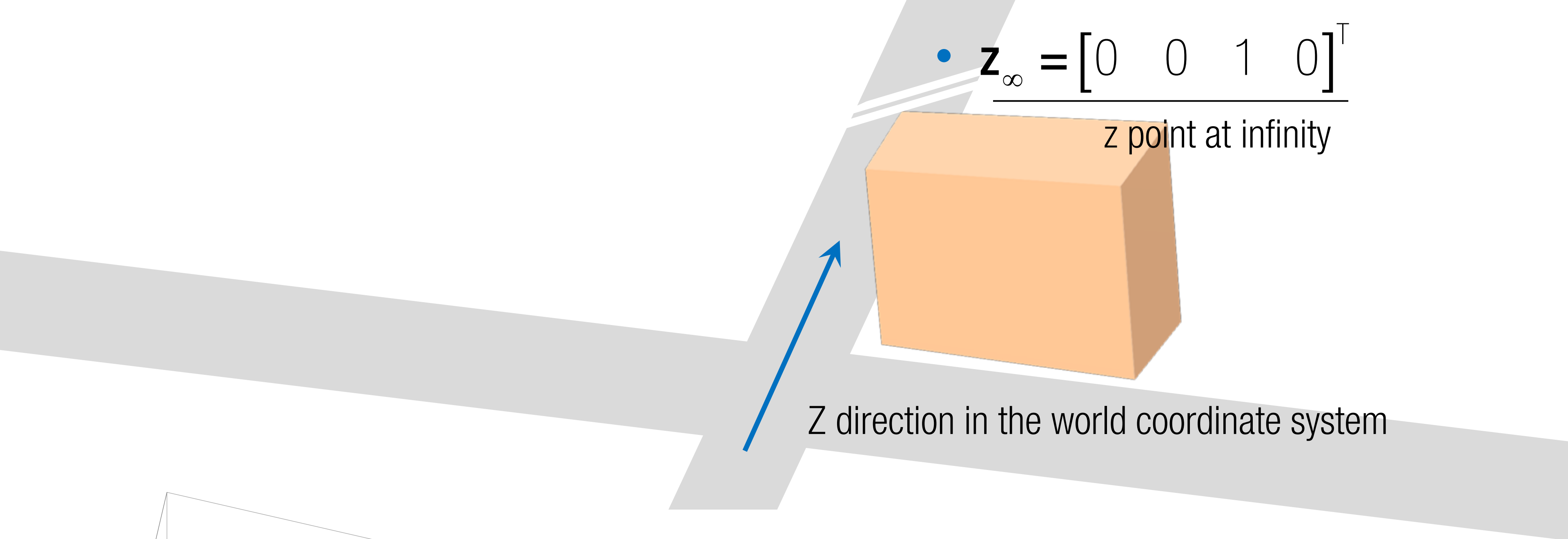


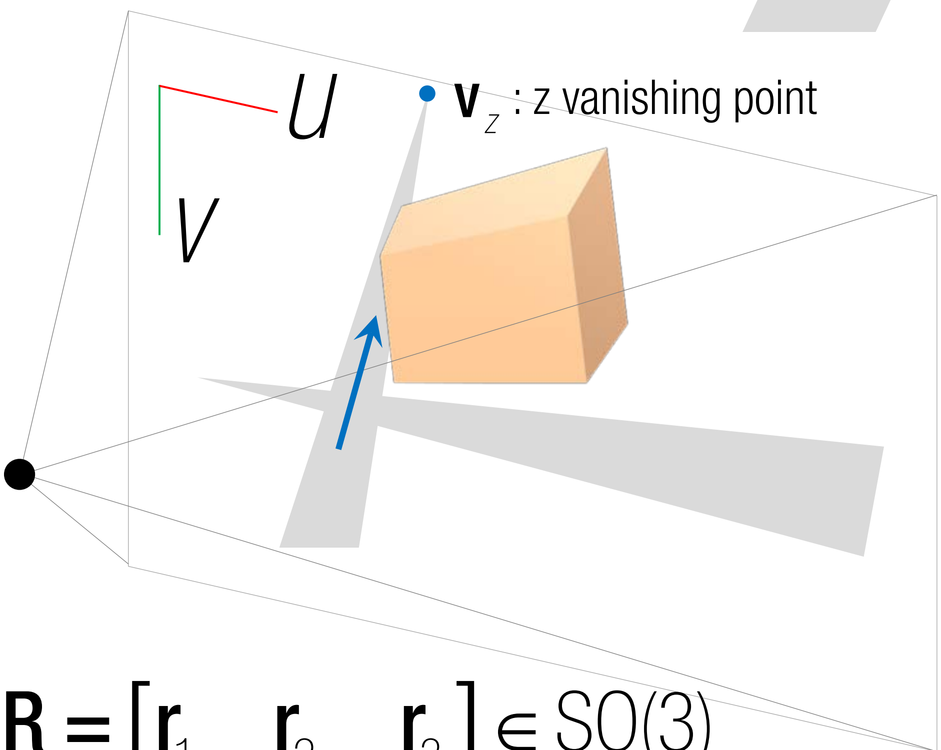


Columns of the rotation matrix represent vanishing points of world axes.

$$z\mathbf{v}_z = \mathbf{K}\begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 & | & \mathbf{t} \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 \end{bmatrix} \in SO(3)$$



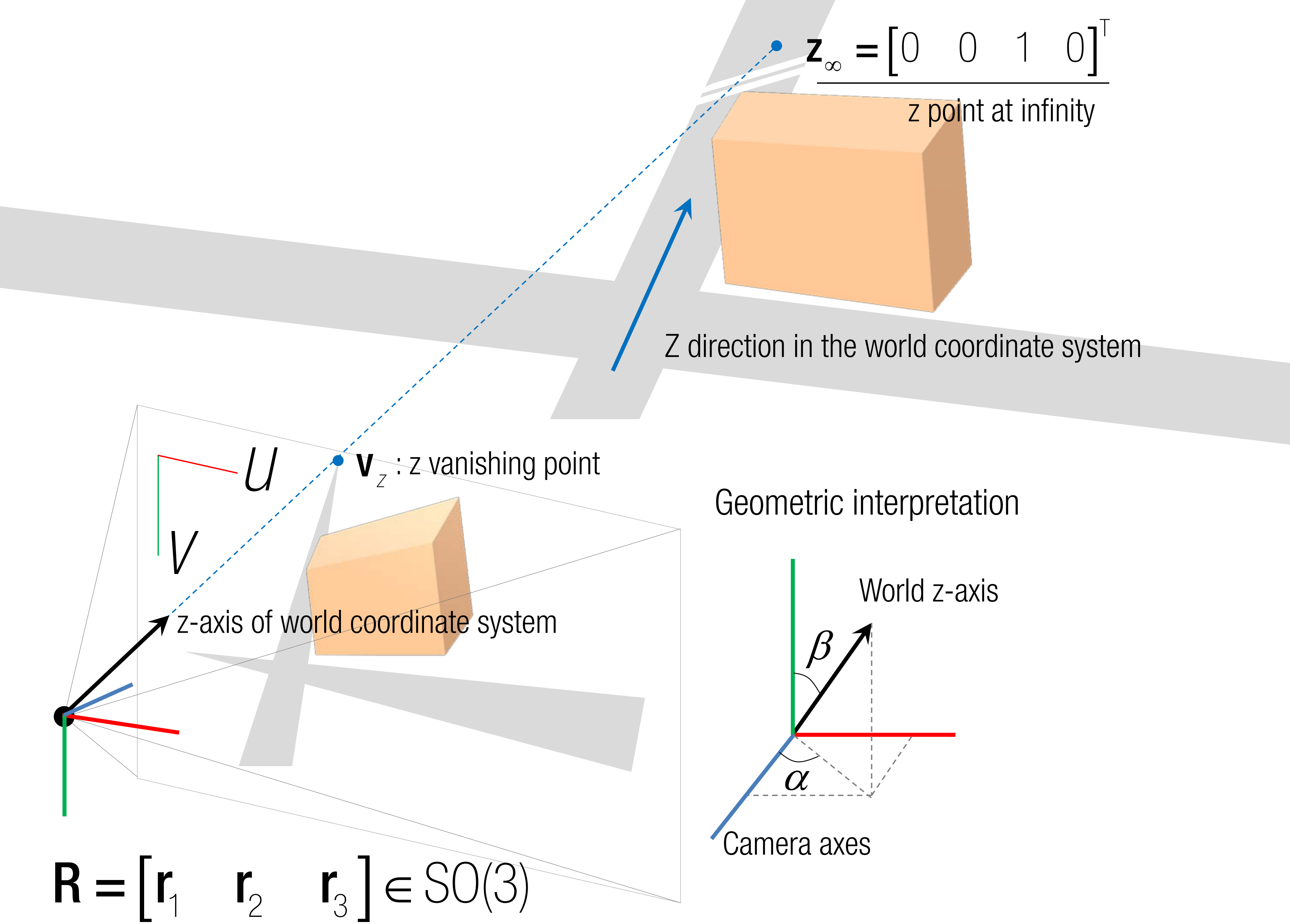


Columns of the rotation matrix represent vanishing points of world axes.

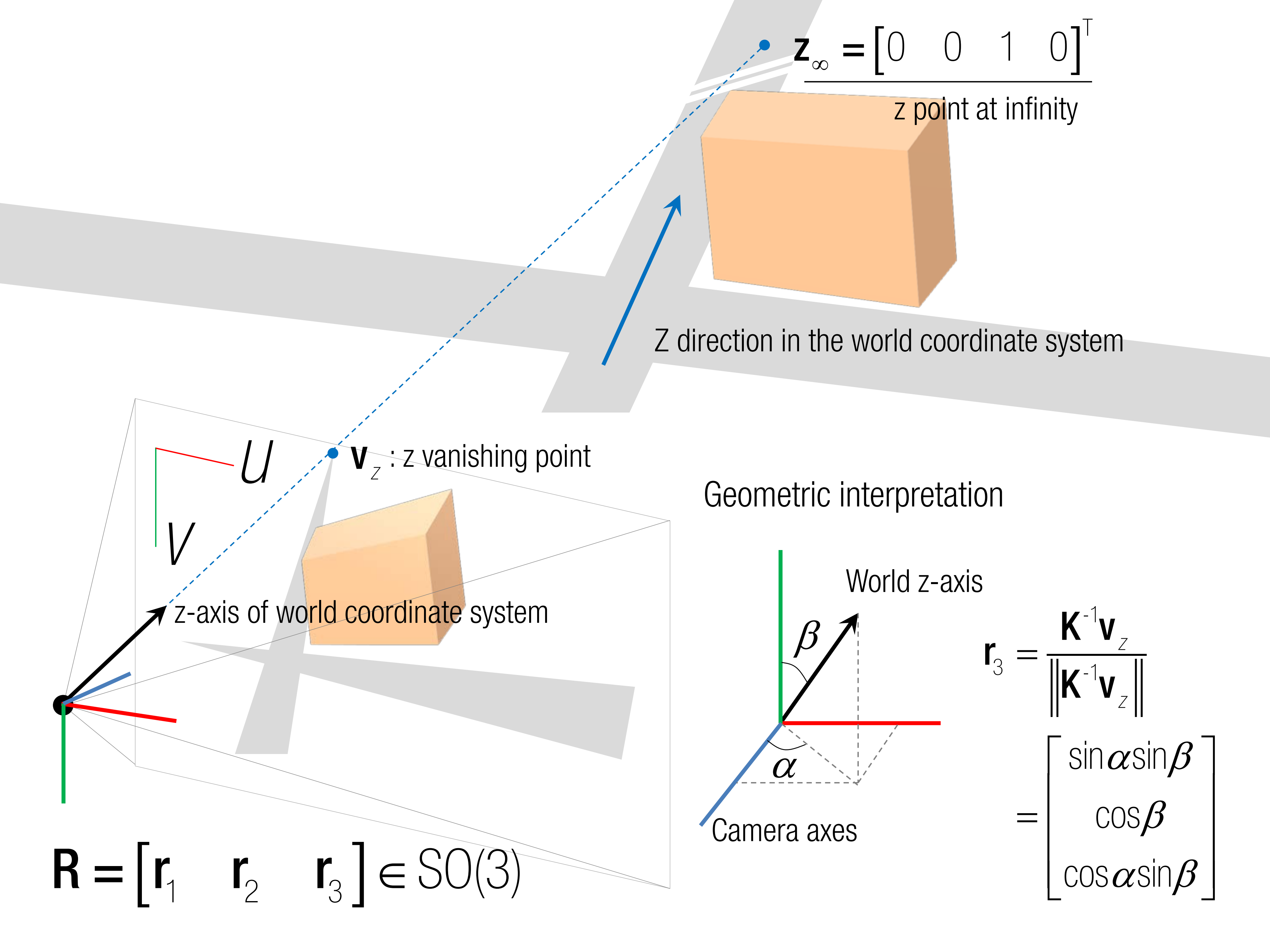
$$z\mathbf{v}_z = \mathbf{K}\mathbf{r}_3$$

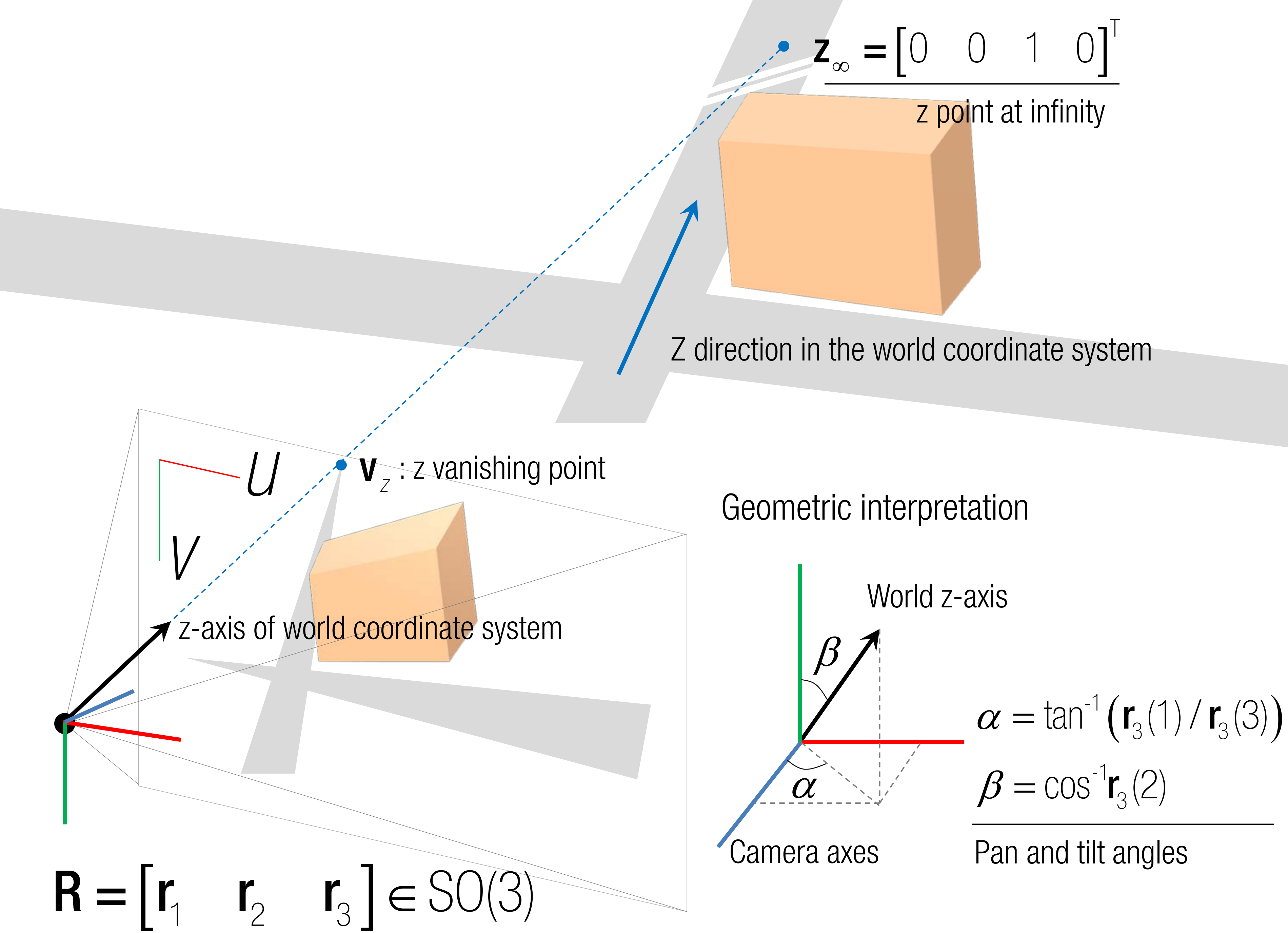
$$\mathbf{r}_3 = \mathbf{K}^{-1}\mathbf{v}_z / \left\|\mathbf{K}^{-1}\mathbf{v}_z\right\|$$

$$\mathbf{R} = \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 \end{bmatrix} \in SO(3)$$

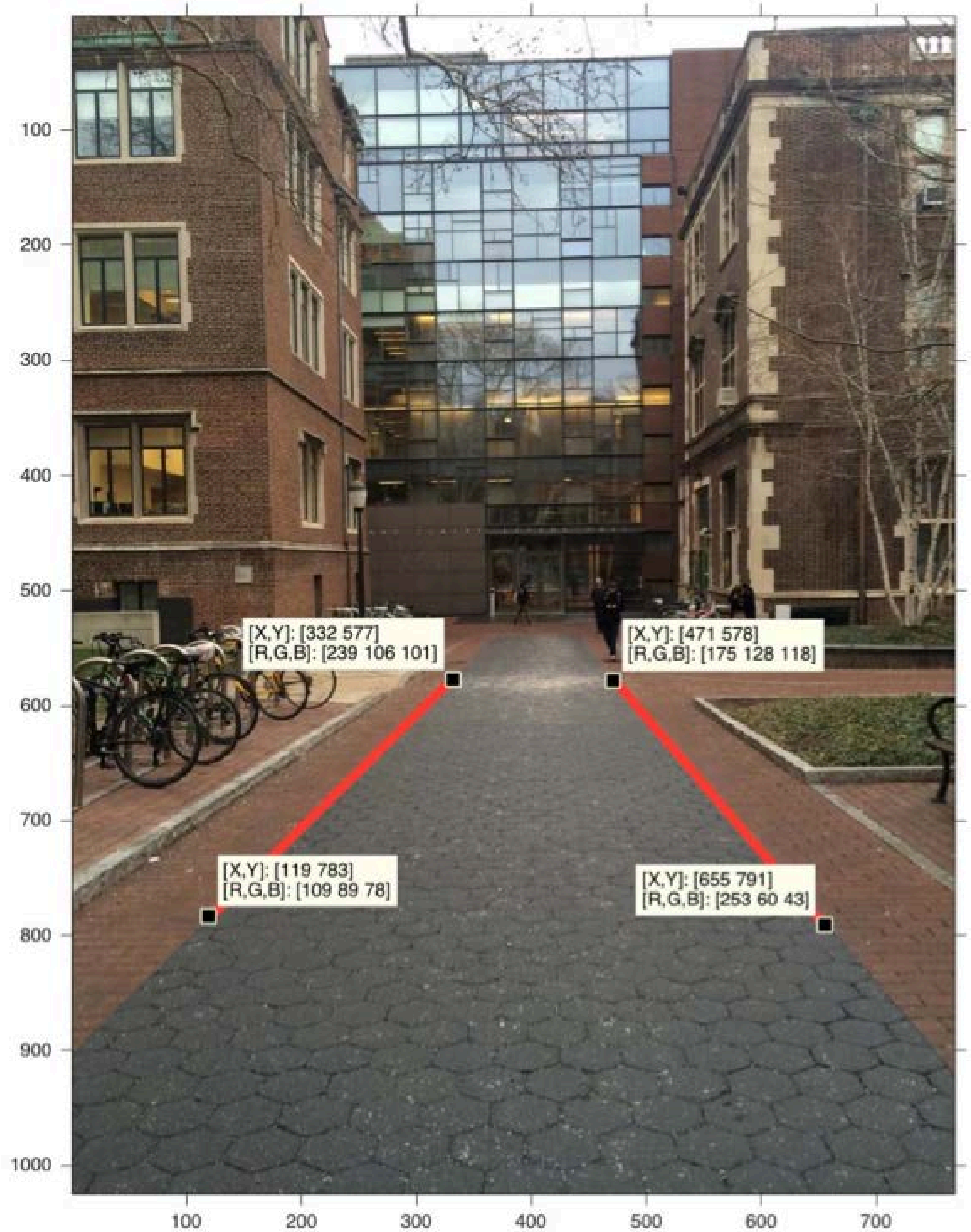


$\mathbf{z}_\infty = [0 \quad 0 \quad 1 \quad 0]^T$
z point at infinity
Z direction in the world coordinate system
U
V
$\mathbf{v}_z$: z vanishing point
Geometric interpretation
World z-axis
β
α
Camera axes
z-axis of world coordinate system
$\mathbf{R} = [\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{r}_3] \in SO(3)$

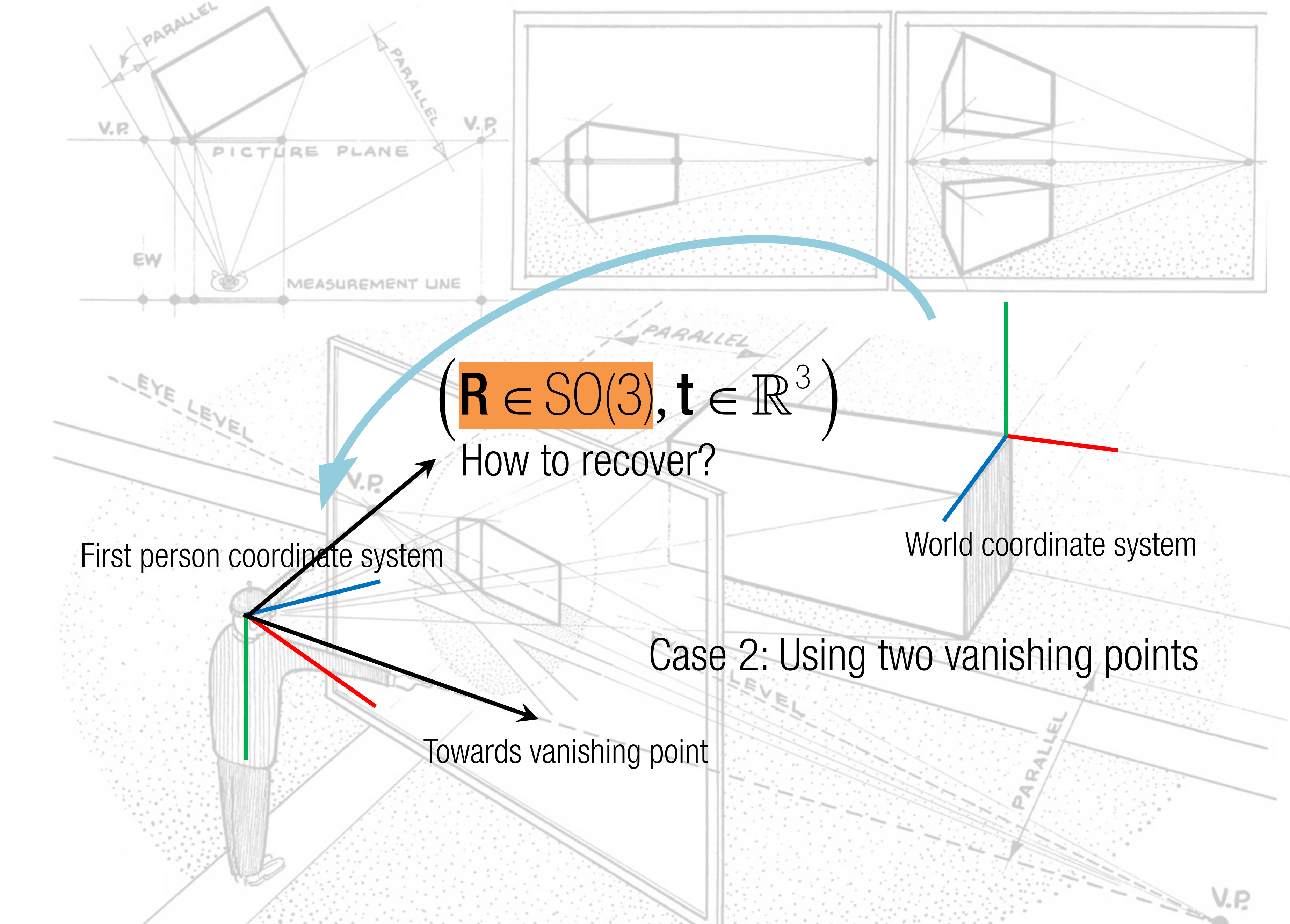

$\mathbf{z}_\infty = [0 \quad 0 \quad 1 \quad 0]^T$
z point at infinity
Z direction in the world coordinate system
U
V
$\mathbf{v}_z$: z vanishing point
z-axis of world coordinate system
Geometric interpretation
World z-axis
β
α
Camera axes
$\mathbf{r}_3 = \frac{\mathbf{K}^{-1}\mathbf{v}_z}{\|\mathbf{K}^{-1}\mathbf{v}_z\|} = \begin{bmatrix} \sin\alpha\sin\beta \\ \cos\beta \\ \cos\alpha\sin\beta \end{bmatrix}$
$\mathbf{R} = [\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{r}_3] \in SO(3)$



$\mathbf{z}_\infty = [0 \quad 0 \quad 1 \quad 0]^T$
z point at infinity
Z direction in the world coordinate system
U
V
$\mathbf{v}_z$: z vanishing point
z-axis of world coordinate system
Geometric interpretation
World z-axis
β
α
Camera axes
$\alpha = \tan^{-1}(\mathbf{r}_3(1) / \mathbf{r}_3(3))$
$\beta = \cos^{-1}\mathbf{r}_3(2)$
Pan and tilt angles
$\mathbf{R} = [\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{r}_3] \in SO(3)$



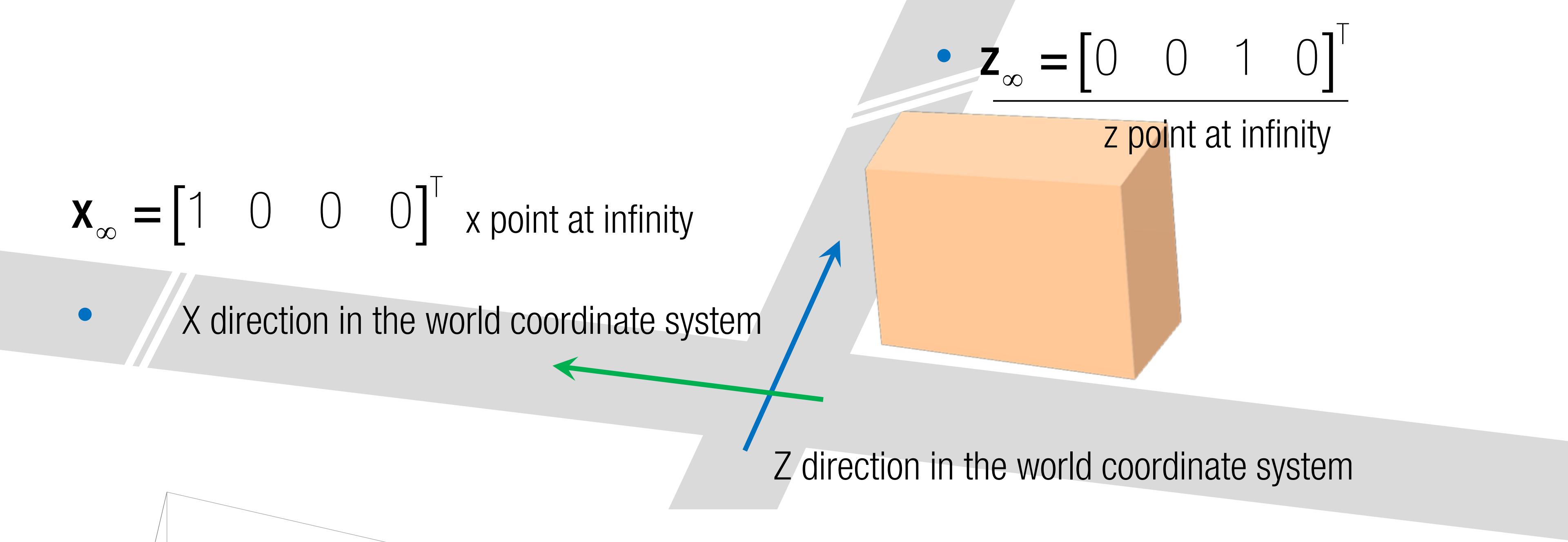
[X,Y]: [332 577]
[R,G,B]: [239 106 101]
[X,Y]: [471 578]
[R,G,B]: [175 128 118]
[X,Y]: [119 783]
[R,G,B]: [109 89 78]
[X,Y]: [655 791]
[R,G,B]: [253 60 43]
100
200
300
400
500
600
700
800
900
1000
100
200
300
400
500
600
700

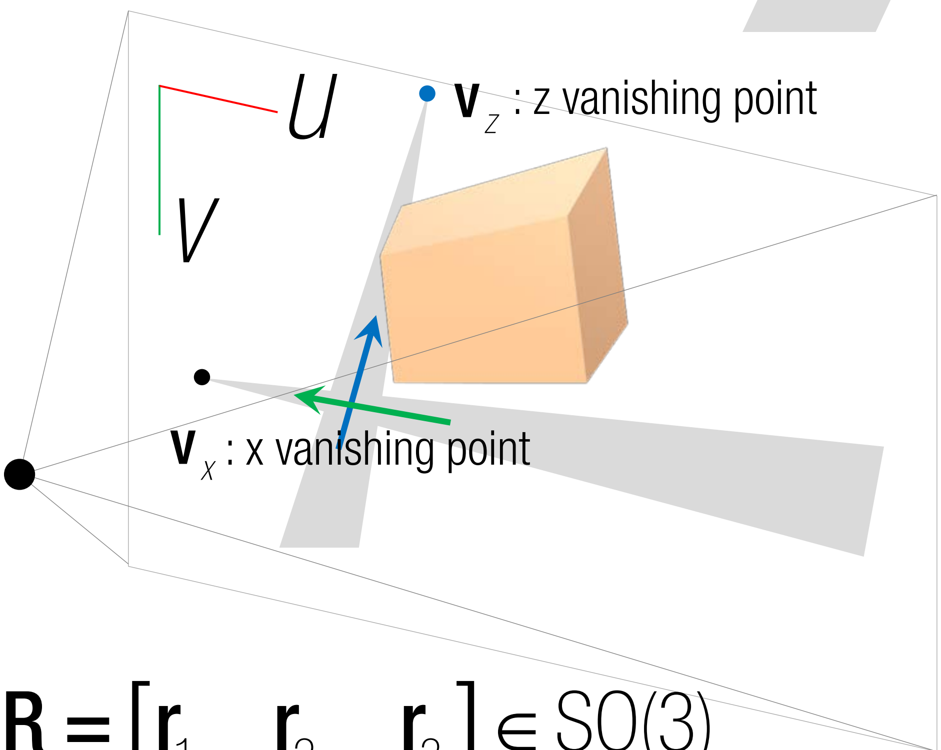


http://www.joshuanava.biz/perspective/in-other-words-the-observer-simply-points-in-the-same-direction-as-the-lines-in-order-to-find-their-vanishing-point.html

- $\mathbf{z}_\infty = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}^T$ z point at infinity

$\mathbf{x}_\infty = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}^T$ x point at infinity





Columns of the rotation matrix represent vanishing points of world axes.

$$\mathbf{r}_3 = \mathbf{K}^{-1} Z \mathbf{v}_Z$$

$$\mathbf{r}_1 = \mathbf{K}^{-1} Z \mathbf{v}_X$$

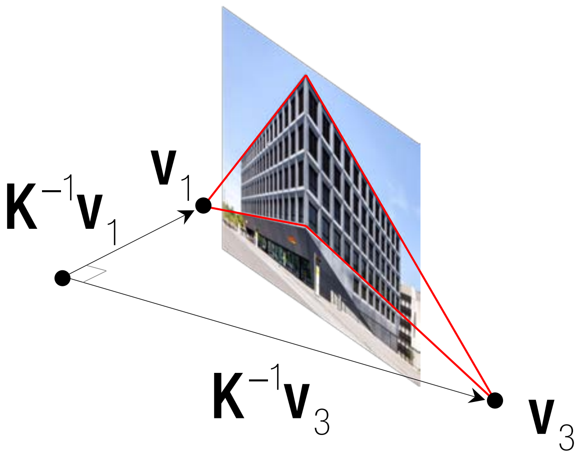
$$\mathbf{r}_2 = \mathbf{r}_3 \times \mathbf{r}_1$$

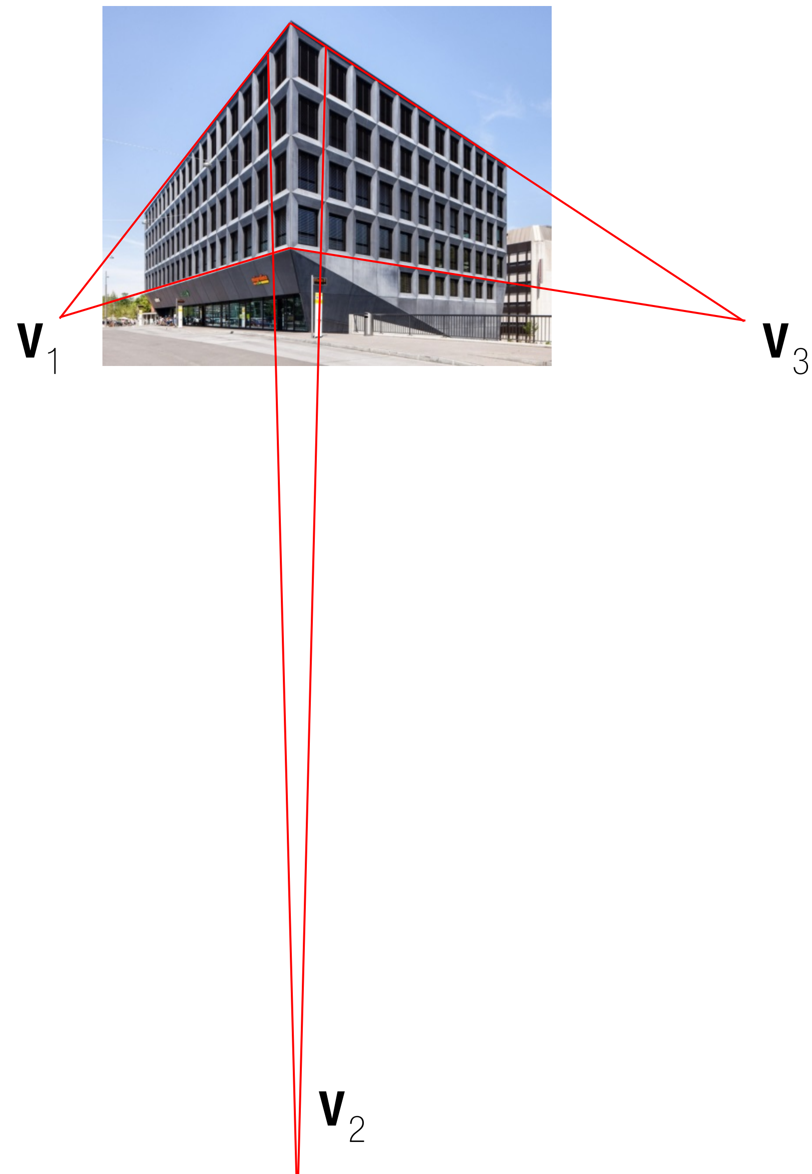
Orthogonal rotation matrix

$$\mathbf{R} = \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 \end{bmatrix} \in SO(3)$$


migrolino


migrolino





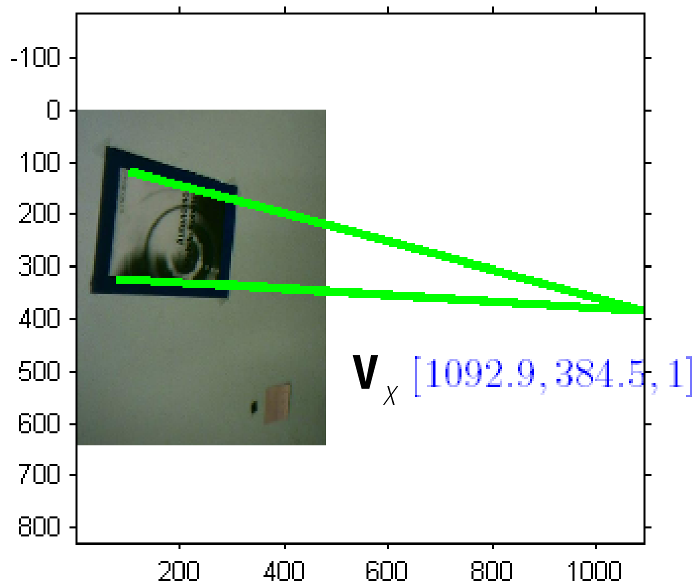
$$\mathbf{r}_3 = \mathbf{K}^{-1} Z \mathbf{v}_Z$$

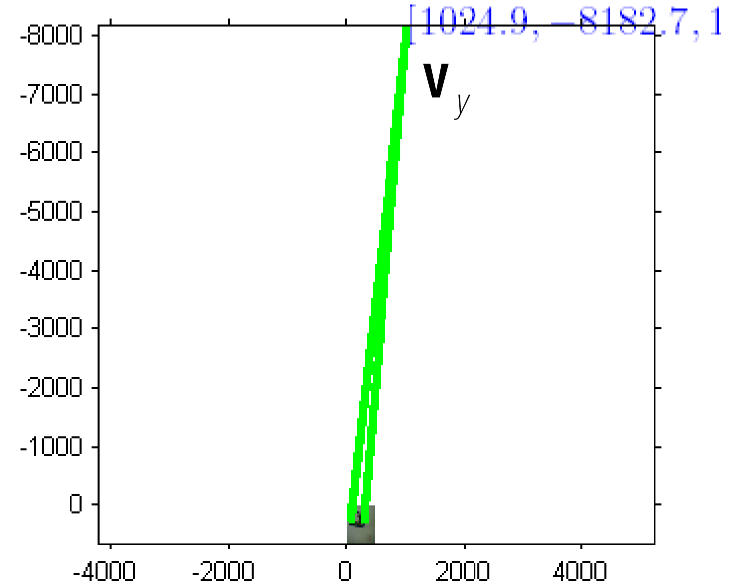
$$\mathbf{r}_1 = \mathbf{K}^{-1} Z \mathbf{v}_X$$

$$\mathbf{r}_2 = \mathbf{r}_3 \times \mathbf{r}_1$$

$$\mathbf{R} = \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 \end{bmatrix}$$

Exercise I

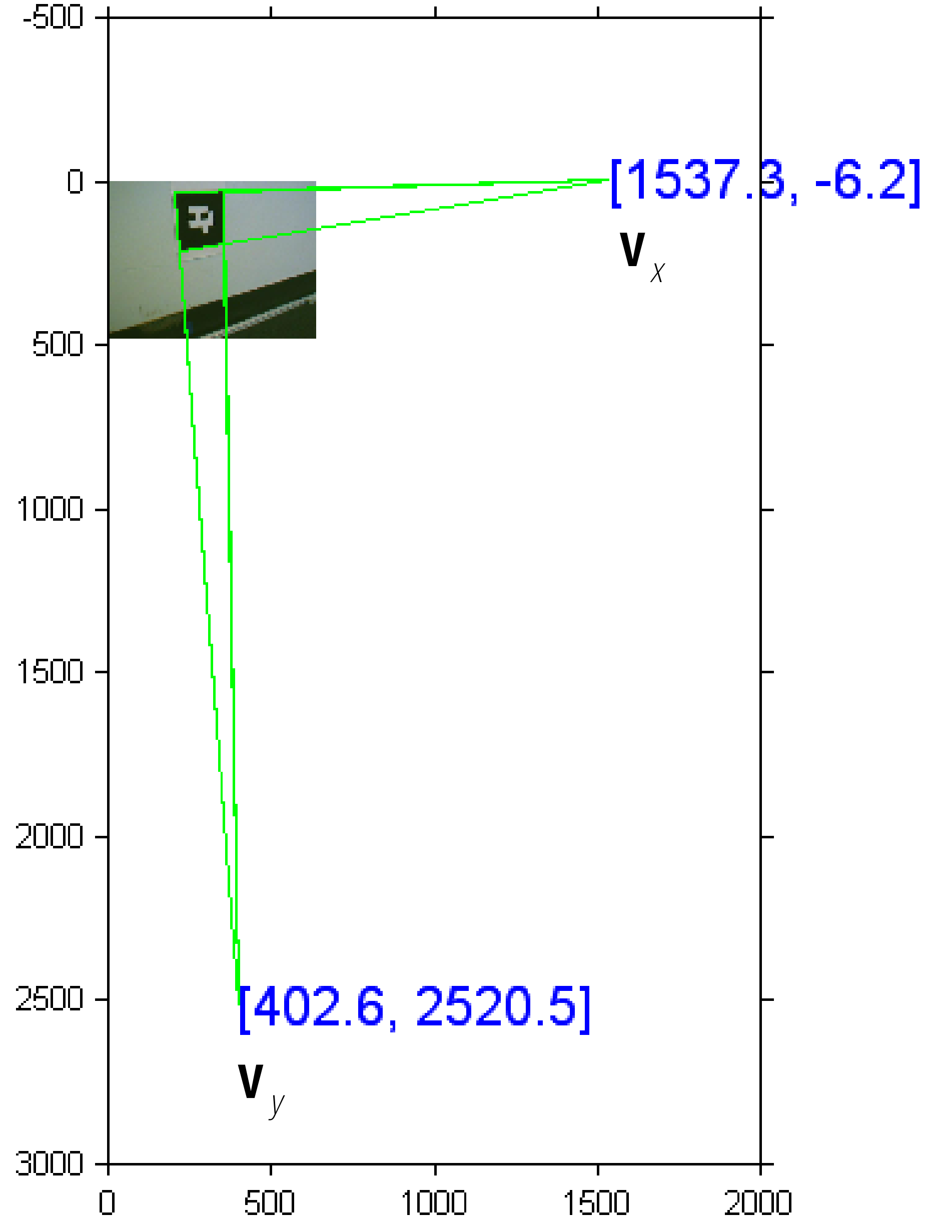




$$\mathbf{r}_1 = \mathbf{K}^{-1} Z \mathbf{v}_x \qquad \mathbf{r}_2 = \mathbf{K}^{-1} Z \mathbf{v}_y$$

$$\mathbf{R} = \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_1 \times \mathbf{r}_2 \end{bmatrix}$$

Exercise II



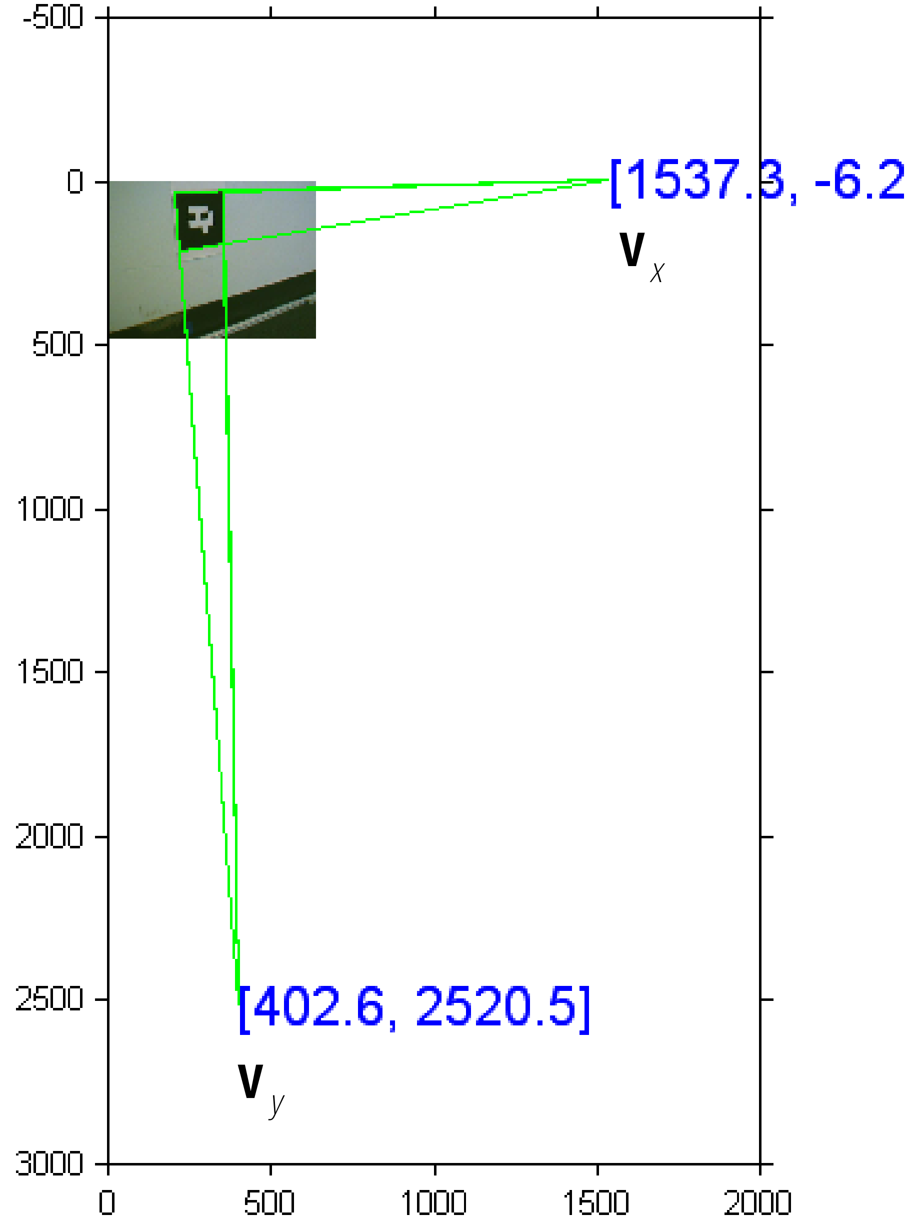
[1537.3, -6.2]
v_x
[402.6, 2520.5]
v_y
-500
0
500
1000
1500
2000
2500
3000
0
500
1000
1500
2000

Exercise II

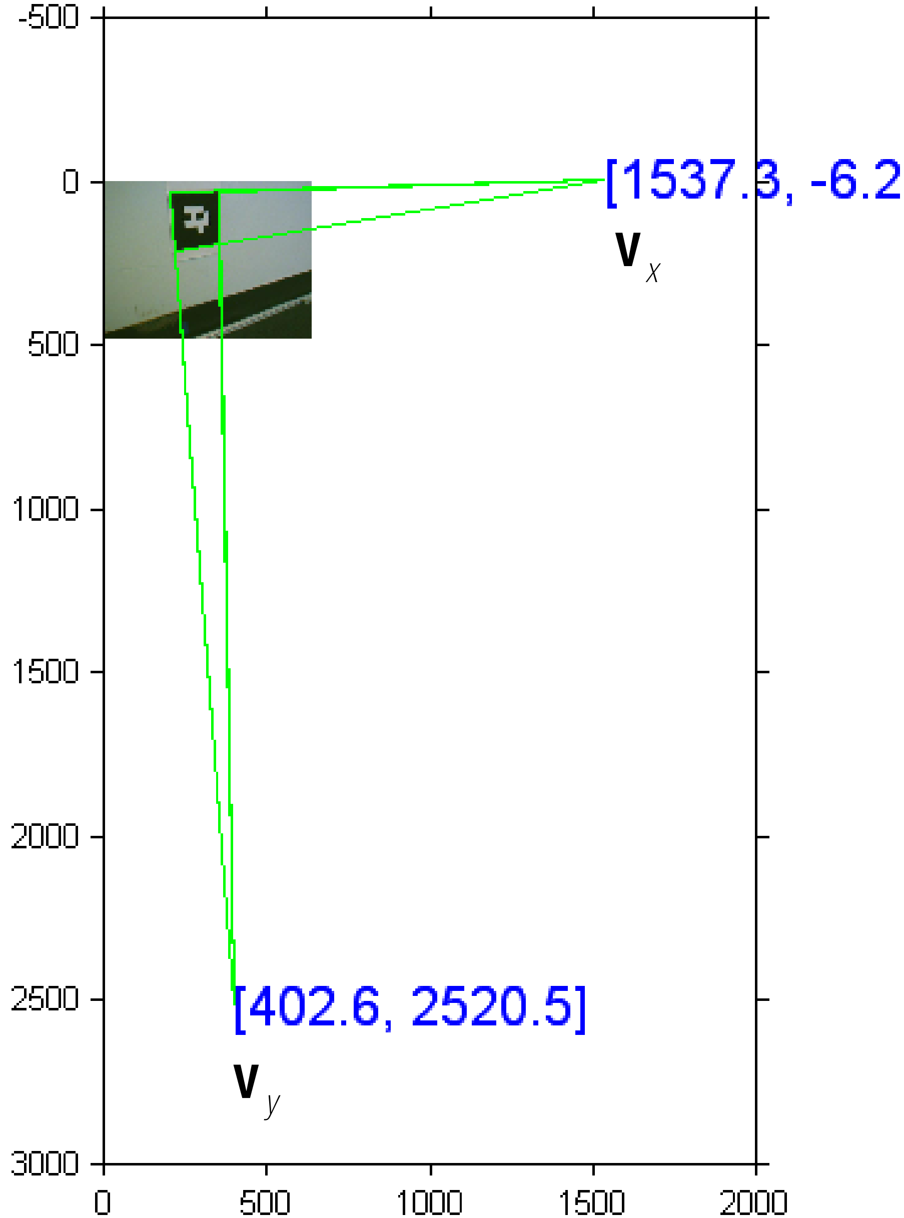
$$\mathbf{r}_1 = \mathbf{K}^{-1}\mathbf{v}_x \;/\; \left\|\mathbf{K}^{-1}\mathbf{v}_x\right\|$$

$$\mathbf{r}_2 = \mathbf{K}^{-1}\mathbf{v}_y \;/\; \left\|\mathbf{K}^{-1}\mathbf{v}_y\right\|$$

Scale normalization



Exercise II



$$
\begin{aligned}
r_1 &= (0.8017, -0.2086, 0.5602)^T \\
r_2 &= (0.0067, 0.9411, 0.3382)^T \\
r_3 &= r_1 \times r_2 = (-0.5988, -0.2673, 0.7558)^T
\end{aligned}
$$

Exercise II

Estimate pan/tilt from $\mathbf{r}_3$.

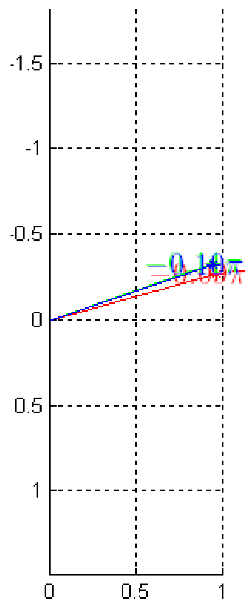
$$\alpha = \tan^{-1}\left(\mathbf{r}_3(1) / \mathbf{r}_3(3)\right)$$

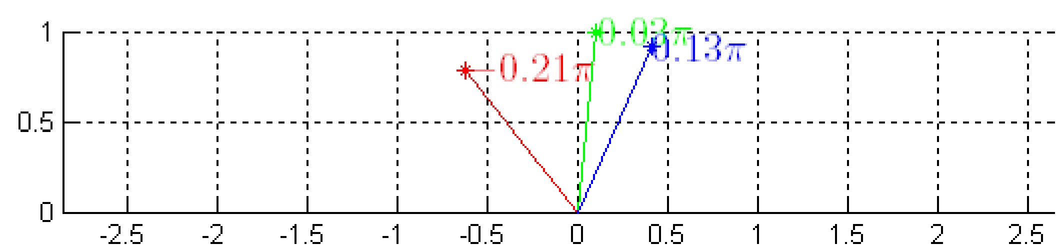
$$\beta = \sin^{-1}\mathbf{r}_3(2)$$

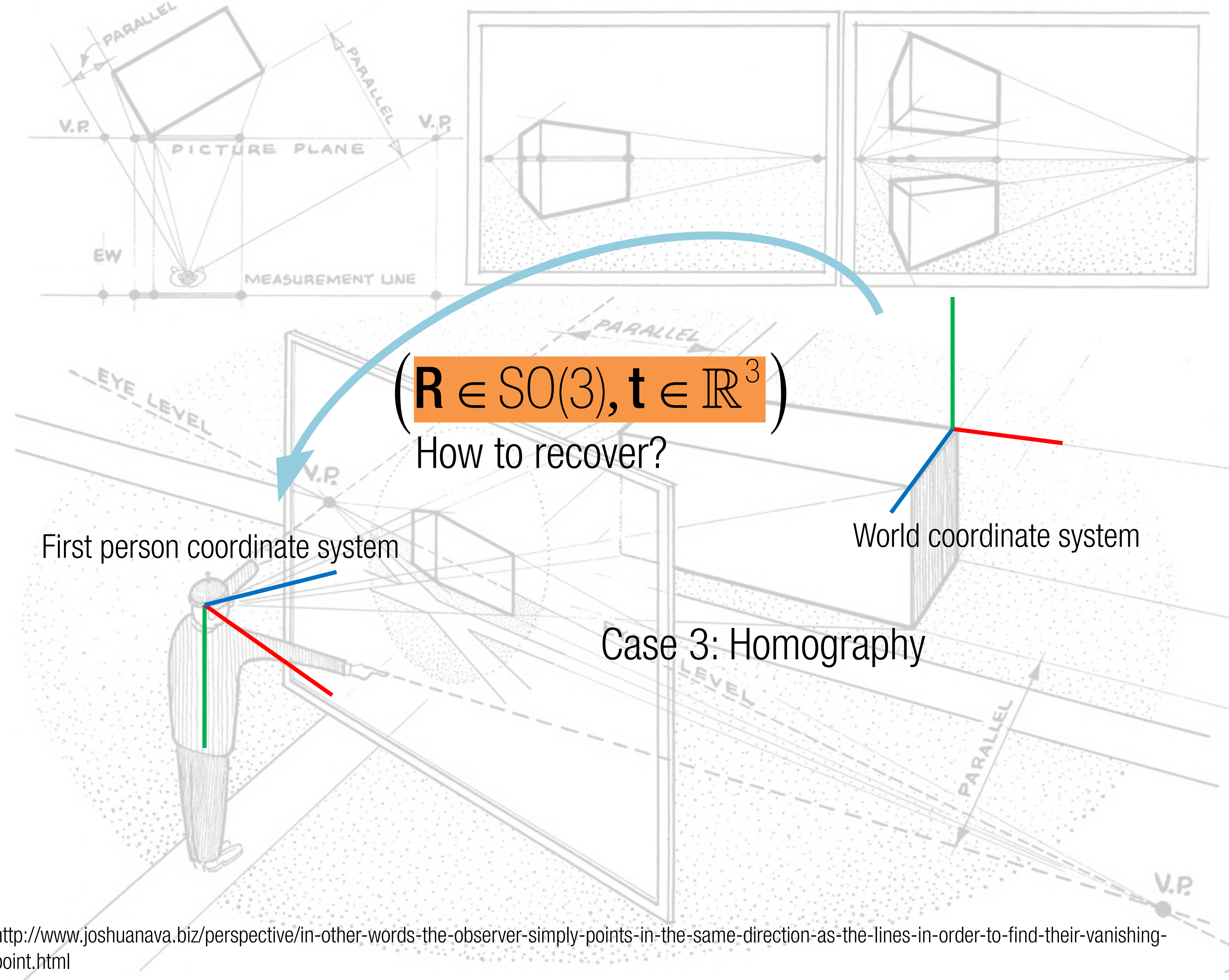
$$\begin{aligned} \alpha &= -0.6691 = -0.2130\pi \\ \beta &= -0.2706 = -0.0861\pi \end{aligned}$$

$$R = \begin{pmatrix} 0.8017 & 0.0067 & -0.5977 \\ -0.2086 & 0.9411 & -0.2673 \\ 0.5602 & 0.3382 & 0.7558 \end{pmatrix}$$

Exercise II


-1.5
-1
-0.5
0
0.5
1
0
0.5
1
Tilt angle β


1
0.5
0
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
-0.21π
0.03π
0.13π
Pan angle α



$\left(\mathbf{R} \in SO(3), \mathbf{t} \in \mathbb{R}^3\right)$

How to recover?

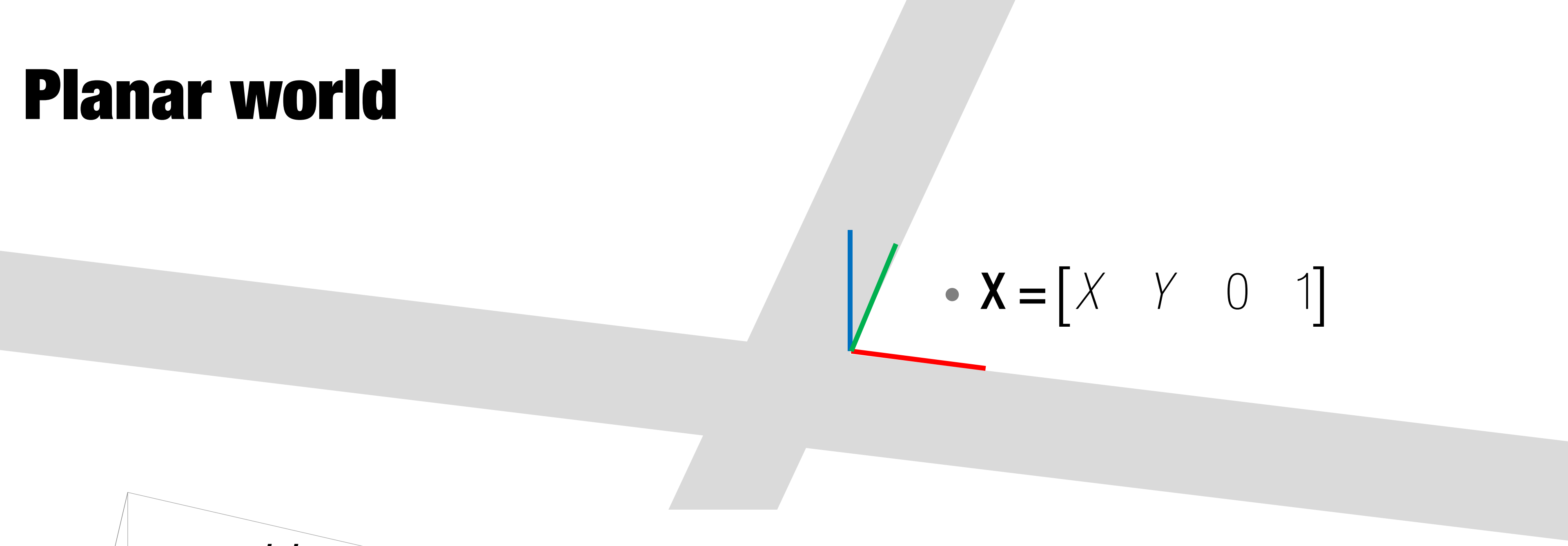
First person coordinate system

World coordinate system

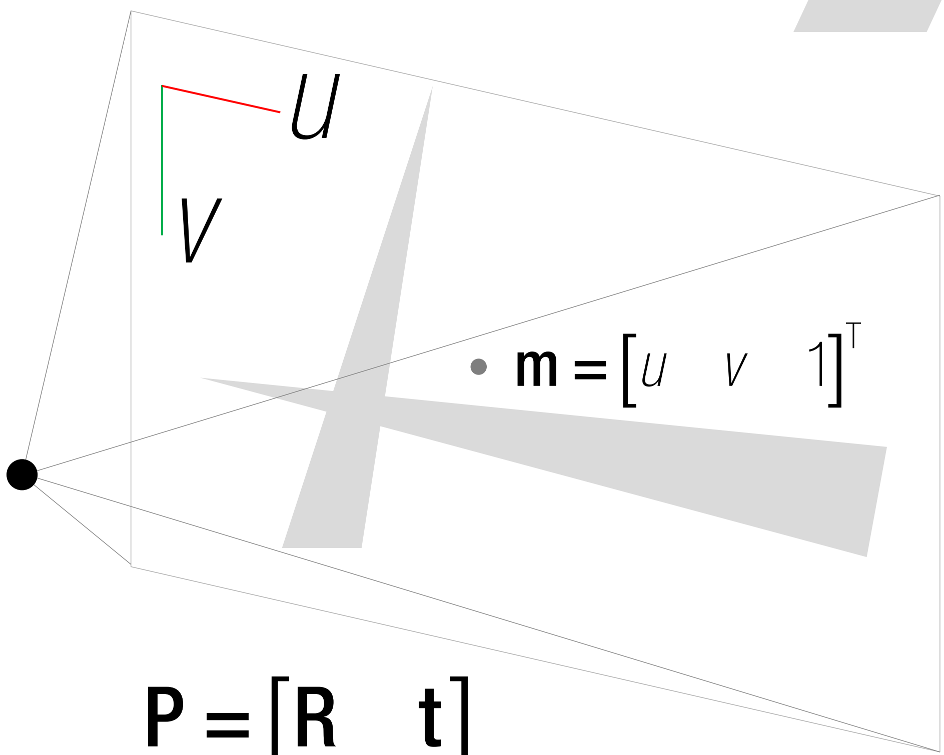
Case 3: Homography

http://www.joshuanava.biz/perspective/in-other-words-the-observer-simply-points-in-the-same-direction-as-the-lines-in-order-to-find-their-vanishing-point.html

Planar world



$\mathbf{X} = \begin{bmatrix} X & Y & 0 & 1 \end{bmatrix}$



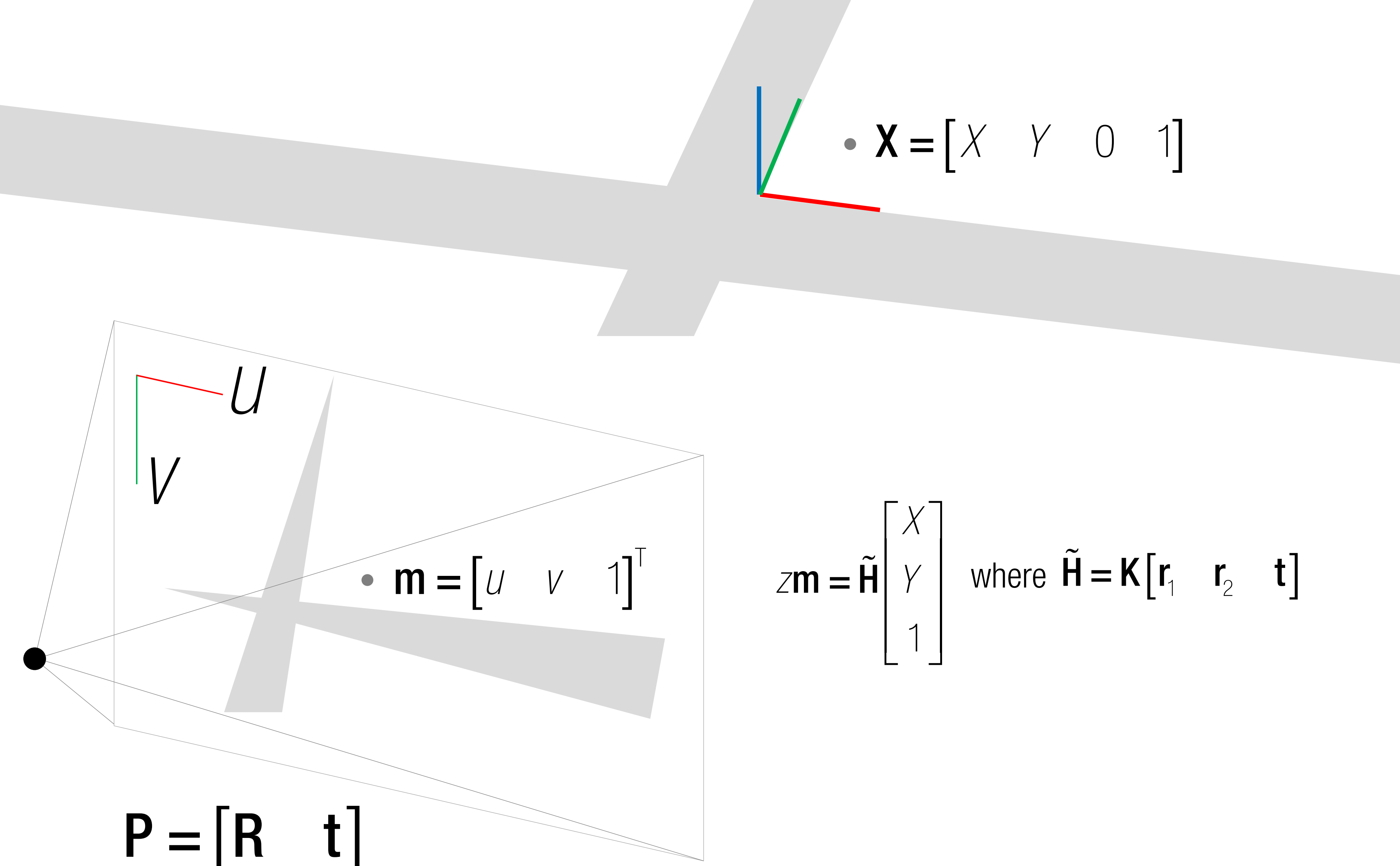
$\mathbf{m} = \begin{bmatrix} u & v & 1 \end{bmatrix}^{\mathrm{T}}$

$$z\mathbf{m} = \mathbf{K}\begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 & | & \mathbf{t} \end{bmatrix}\mathbf{X}$$

$$= \mathbf{K}\begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 & | & \mathbf{t} \end{bmatrix}\begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = \underbrace{\mathbf{K}\begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix}}_{\text{2D homography}}\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}$$

$\mathbf{P} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix}$

Planar world



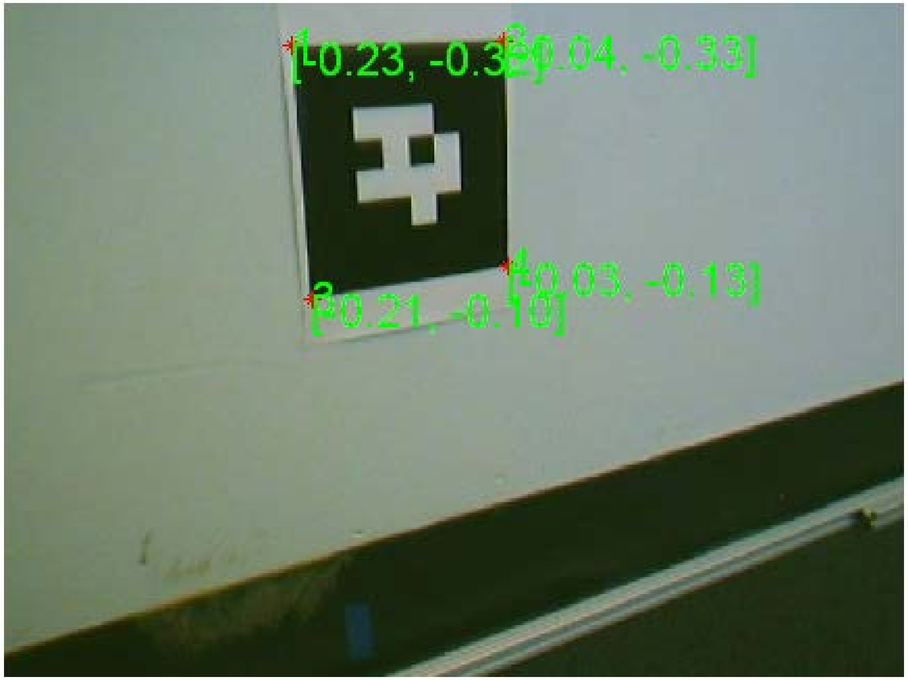
$\mathbf{X} = \begin{bmatrix} X & Y & 0 & 1 \end{bmatrix}$

$\mathbf{m} = \begin{bmatrix} u & v & 1 \end{bmatrix}^{\mathrm{T}}$

$$z\mathbf{m} = \tilde{\mathbf{H}} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \text{ where } \tilde{\mathbf{H}} = \mathbf{K}\begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix}$$

$$\mathbf{P} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix}$$

Exercise

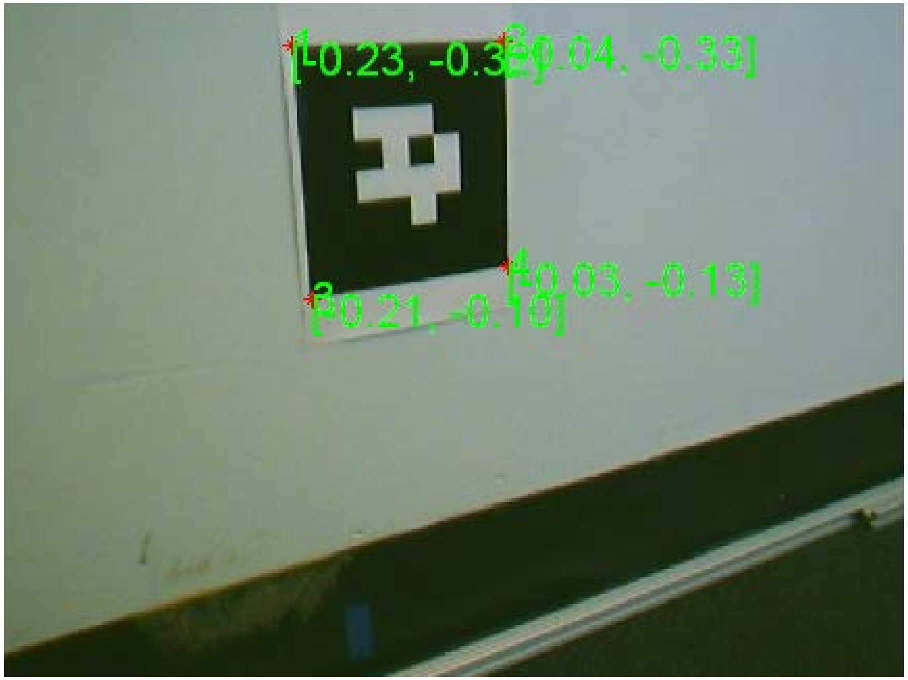


Homography from four points:

$$H = \begin{pmatrix} 0.4430 & 0.0037 & -0.1071 \\ -0.1153 & 0.5216 & 0.1506 \\ 0.3096 & 0.1875 & 0.5944 \end{pmatrix}$$

$$\mathbf{H} = \mathbf{K}^{-1}\tilde{\mathbf{H}} = \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix}$$ Note that $\|\mathbf{r}_1\| = \|\mathbf{r}_2\| = 1$

Exercise



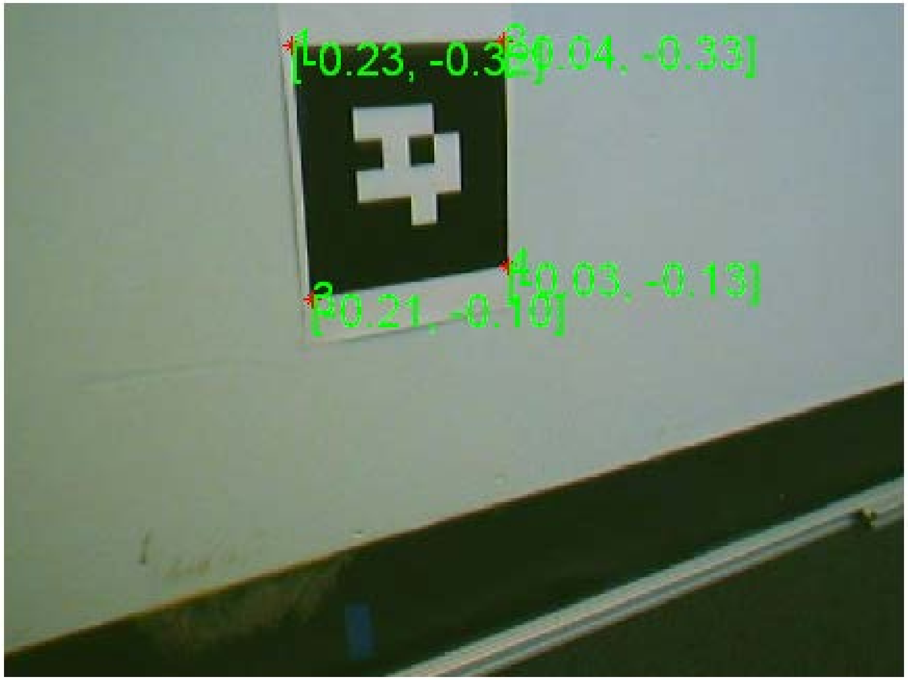
Homography from four points:

$$H = \begin{pmatrix} 0.4430 & 0.0037 & -0.1071 \\ -0.1153 & 0.5216 & 0.1506 \\ 0.3096 & 0.1875 & 0.5944 \end{pmatrix}$$

$$a = \|(H_{11}, H_{21}, H_{31})\|$$: Normalization factor

$$\begin{aligned} t &= H(:,3)/a = (-0.1937, 0.2726, 1.0756)^T \\ r_1 &= H(:,1)/a = (0.8017, -0.2086, 0.5602)^T \\ r_2 &= H(:,2)/a = (0.0067, 0.9439, 0.3392)^T \end{aligned}$$

Exercise



Homography from four points:

$$H = \begin{pmatrix} 0.4430 & 0.0037 & -0.1071 \\ -0.1153 & 0.5216 & 0.1506 \\ 0.3096 & 0.1875 & 0.5944 \end{pmatrix}$$

$$a = \|(H_{11}, H_{21}, H_{31})\|$$: Normalization factor

$$\begin{aligned} t &= H(:,3)/a = (-0.1937, 0.2726, 1.0756)^T \\ r_1 &= H(:,1)/a = (0.8017, -0.2086, 0.5602)^T \\ r_2 &= H(:,2)/a = (0.0067, 0.9439, 0.3392)^T \\ r_3 &= r_1 \times r_2 = (-0.1937, 0.2726, 1.0756)^T \end{aligned}$$

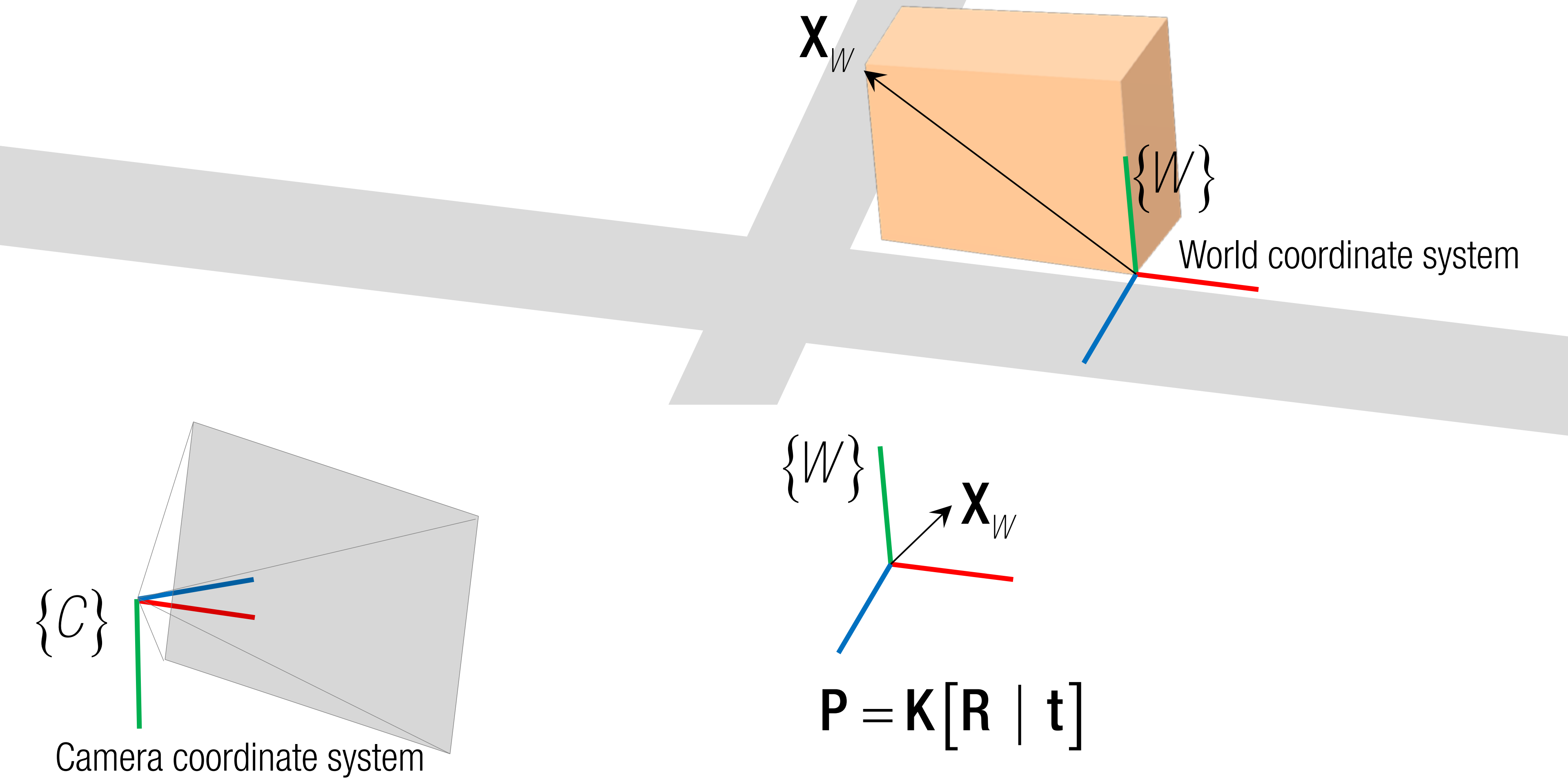
How to estimate the rotation and translation of the robot from the world point of view?

In the case of moving robot(rather than moving target), we need to know the orientation/position of the robot in the world ==>

we need to how to pan/tilt the world oriented to the robot.

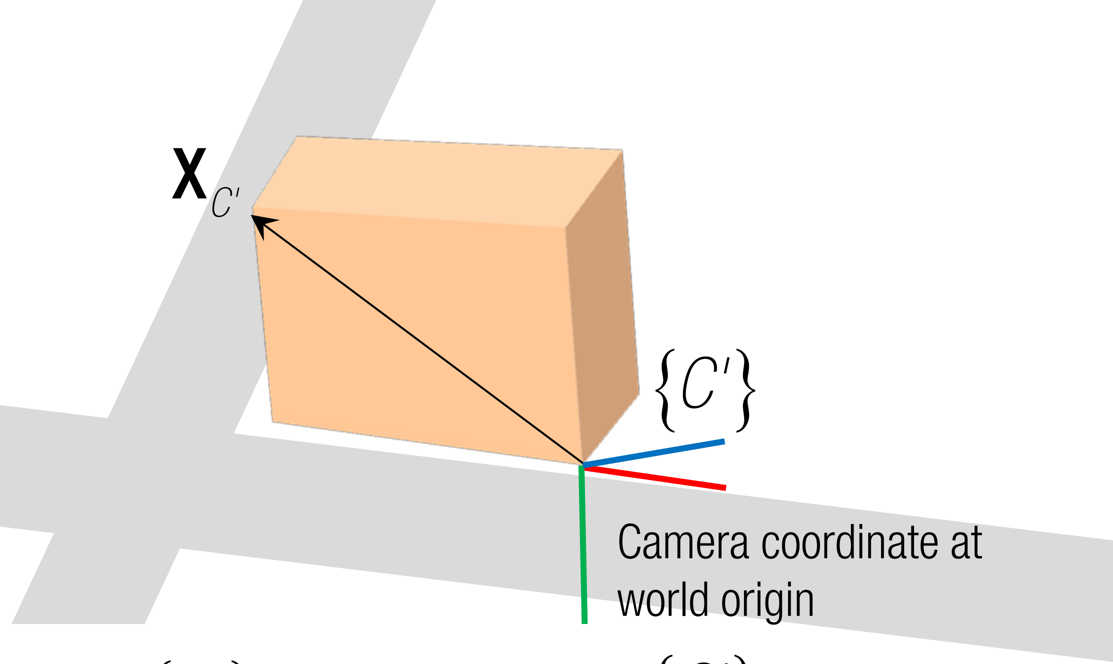
Note: pan/tilt of the camera is very different from the pan/tilt of the world!

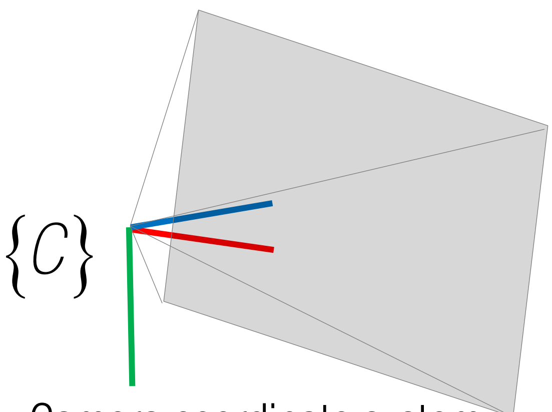
Third person (world) perspective



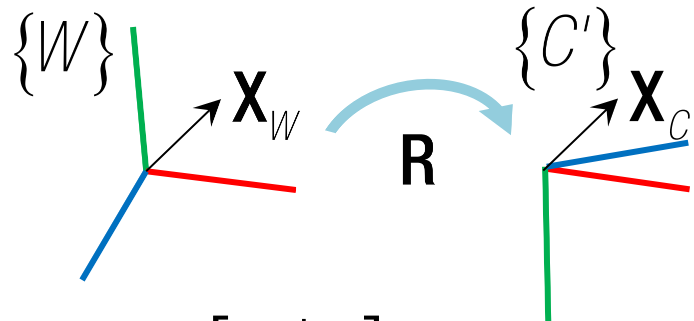
$$\mathbf{P} = \mathbf{K}\left[\mathbf{R} \mid \mathbf{t}\right]$$

First person perspective





Camera coordinate system

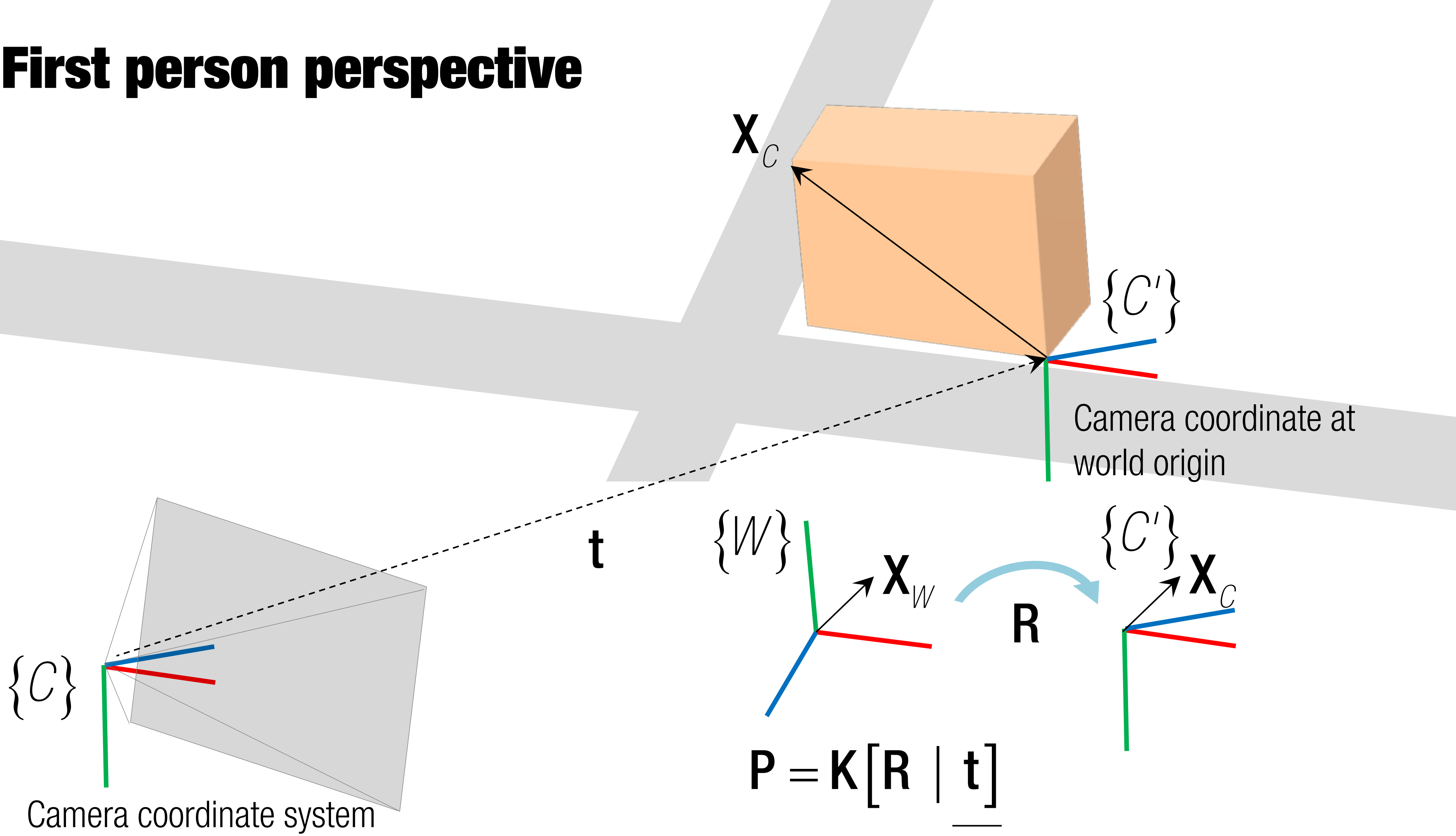


$$\mathbf{P} = \mathbf{K}[\underline{\mathbf{R} \mid} \ \mathbf{t}]$$

Coordinate transform from $\{W\}$ to $\{C'\}$

$$\mathbf{X}_{C'} = \mathbf{R}\mathbf{X}_W$$

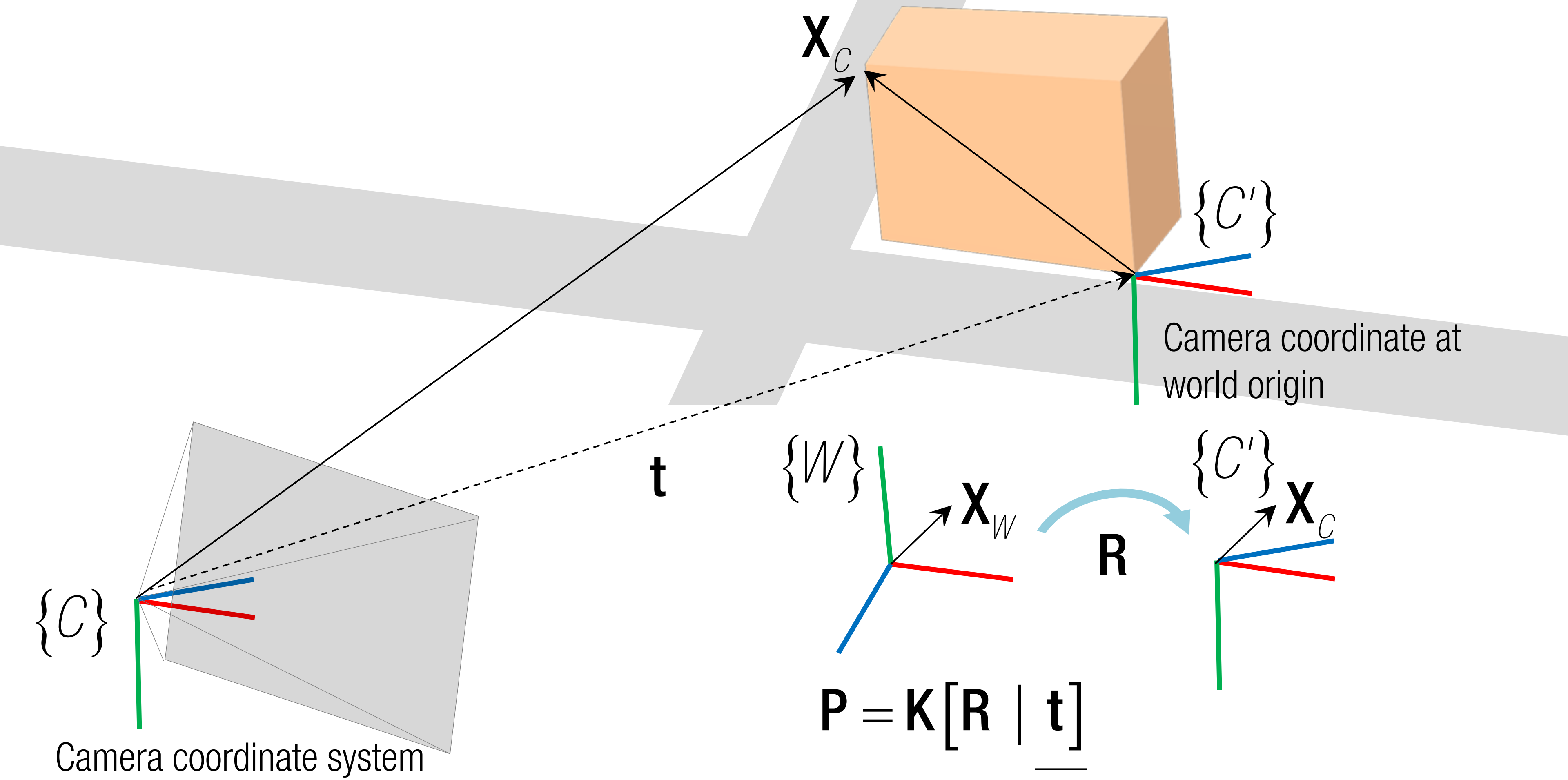
First person perspective



$$\mathbf{P} = \mathbf{K}\left[\mathbf{R} \mid \underline{\mathbf{t}}\right]$$

Camera translation seen from $\{C\}$

First person perspective



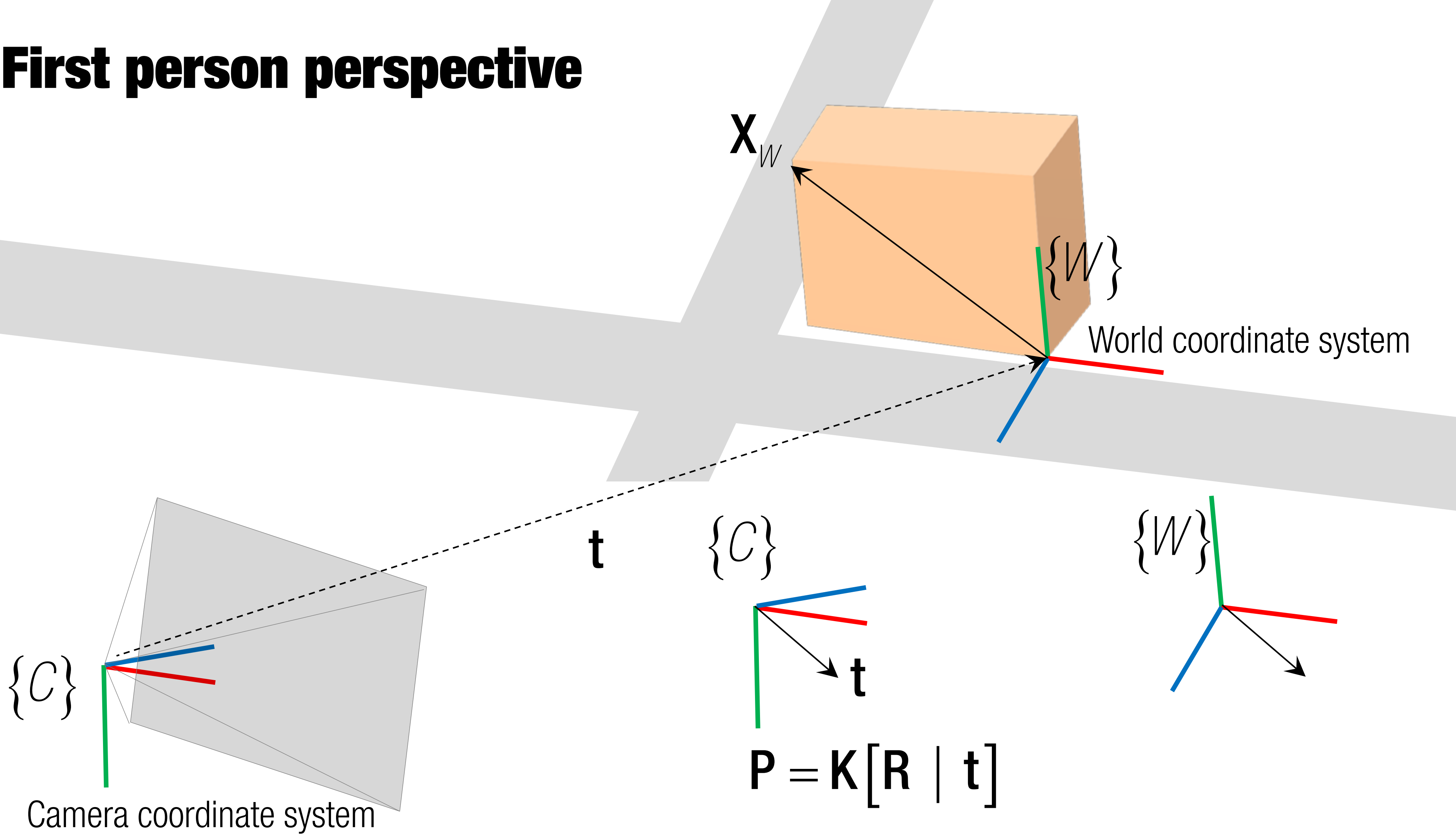
$$\mathbf{P} = \mathbf{K}\left[\mathbf{R} \mid \underline{\mathbf{t}}\right]$$

Camera translation seen from $\{C\}$

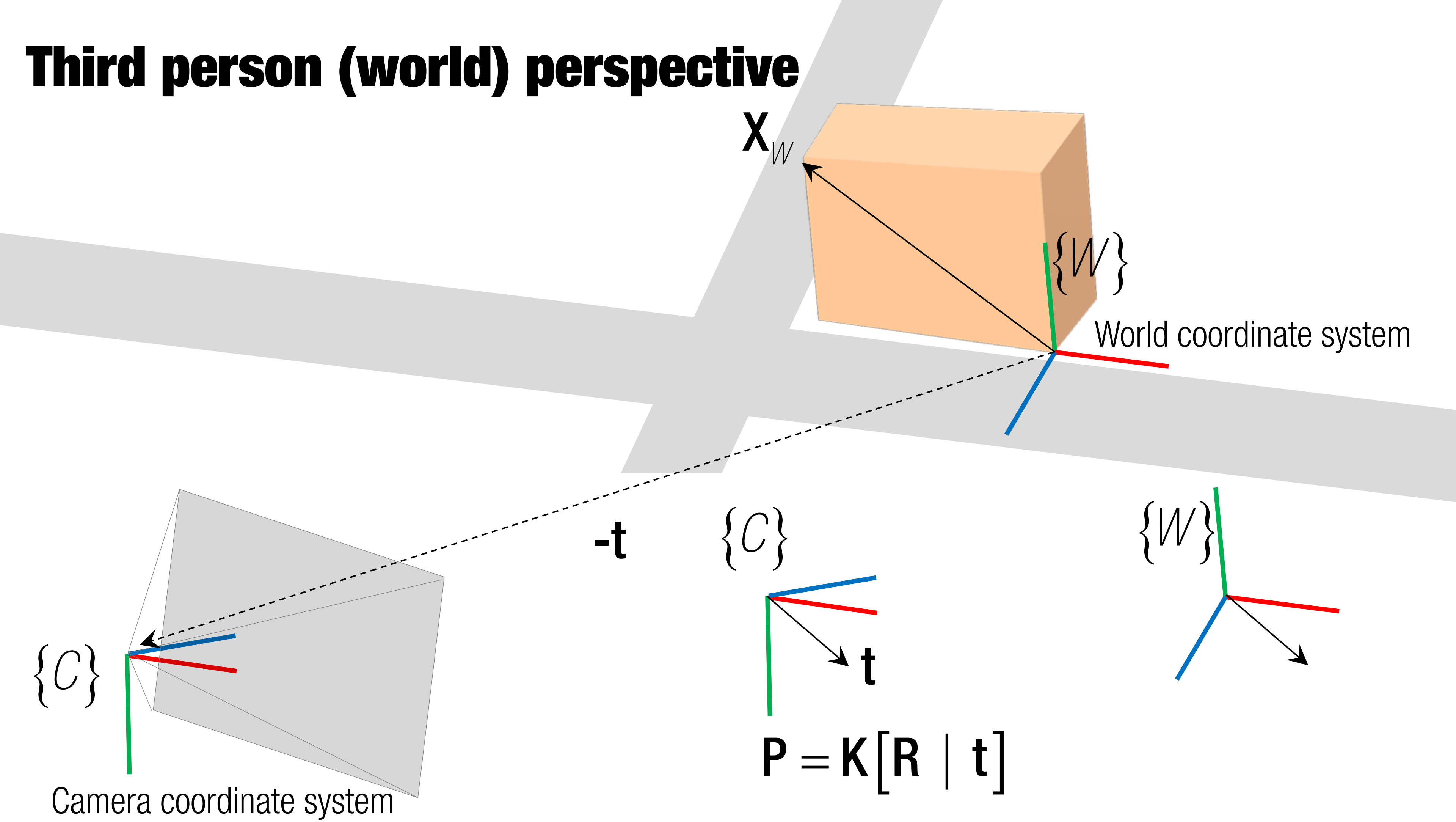
$$\mathbf{X}_C = \mathbf{R}\mathbf{X}_W + \mathbf{t}$$

Looking a point in world through the camera view point

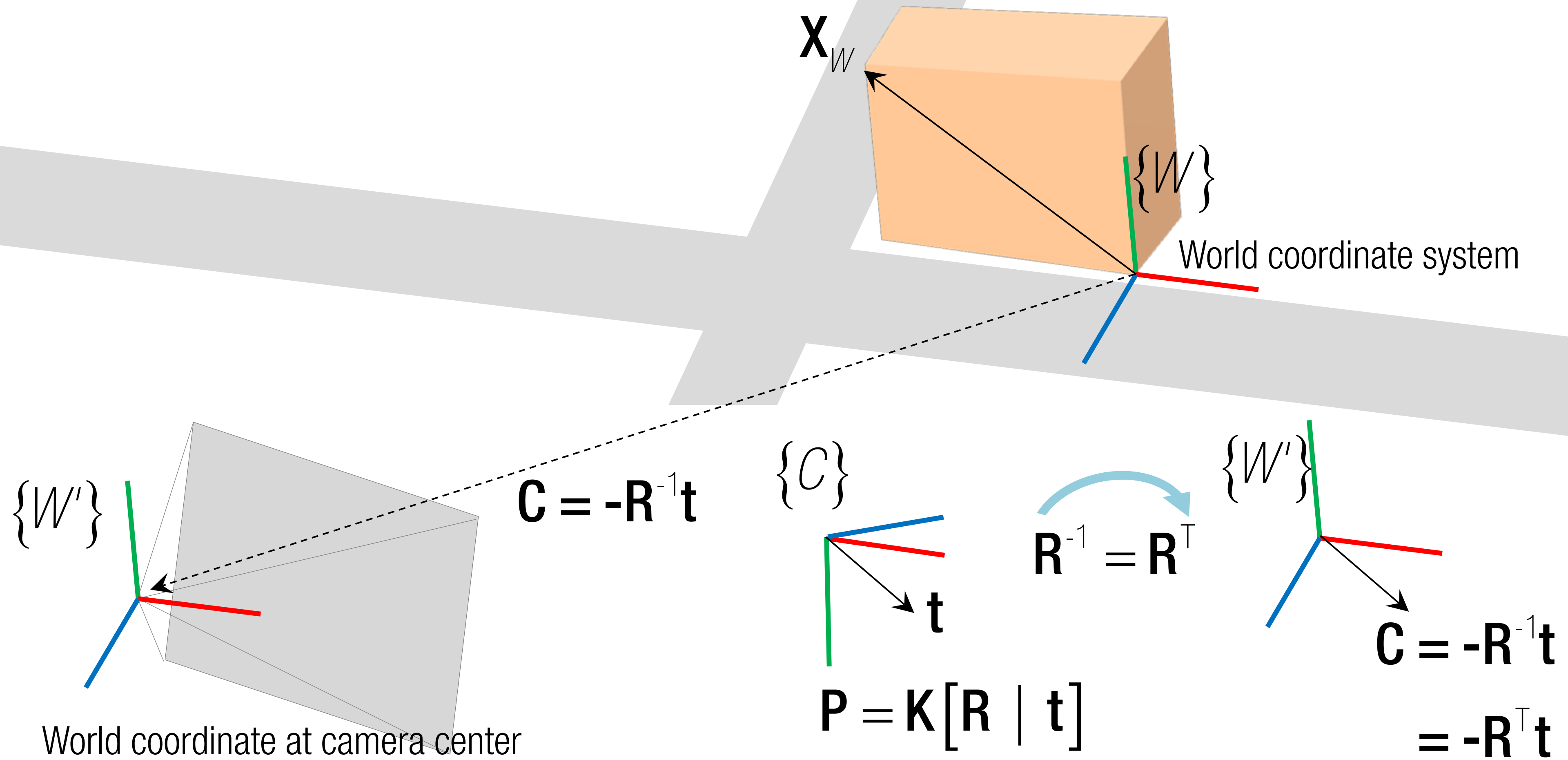
First person perspective



Third person (world) perspective



Third person (world) perspective



$$\mathbf{C} = -\mathbf{R}^{-1}\mathbf{t}$$
$$= -\mathbf{R}^{\mathrm{T}}\mathbf{t}$$

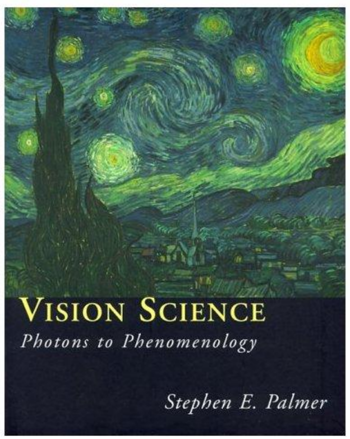
$$\mathbf{P} = \mathbf{K}\left[\mathbf{R} \mid \mathbf{t}\right]$$
$$= \mathbf{K}\left[\mathbf{R} \mid -\mathbf{RC}\right]$$
$$= \mathbf{KR}\left[\mathbf{I}_{3\times3} \mid \underline{-\mathbf{C}}\right]$$

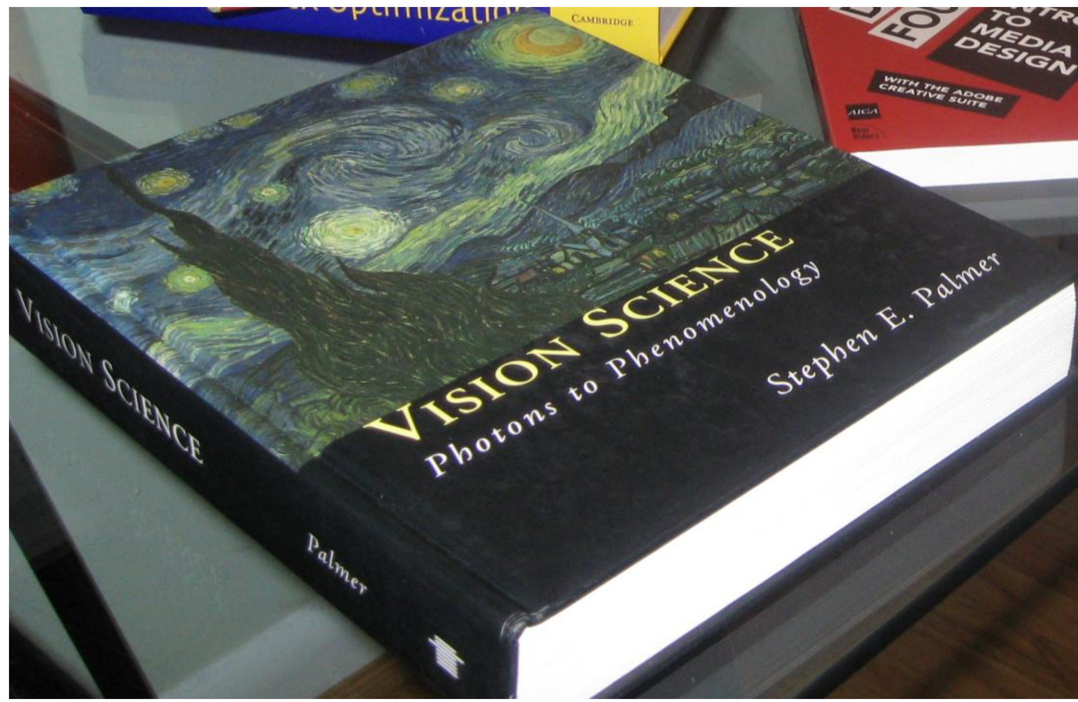
Camera center seen from world coordinate system

Robot Perception: Compute Projective Transformations

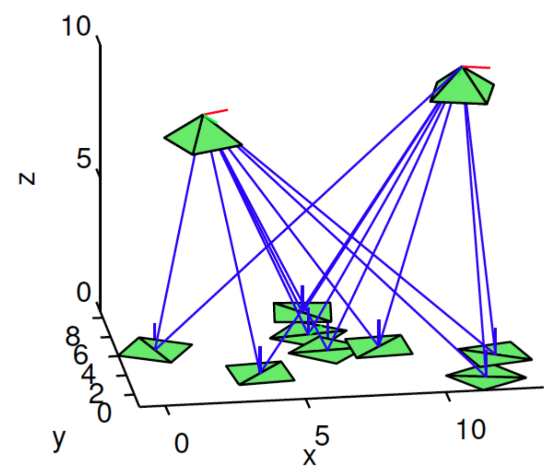
Kostas Daniilidis

A perspective projection of a plane (like a camera image) is always a projective transformation





Using the projective transformation the pose of a robot with respect to a planar pattern:



Projective Transformation

Definition

A **projective transformation** is any invertible matrix transformation $\mathbb{P}^2 \to \mathbb{P}^2$.

A projective transformation A maps p to $p' \sim Ap$.

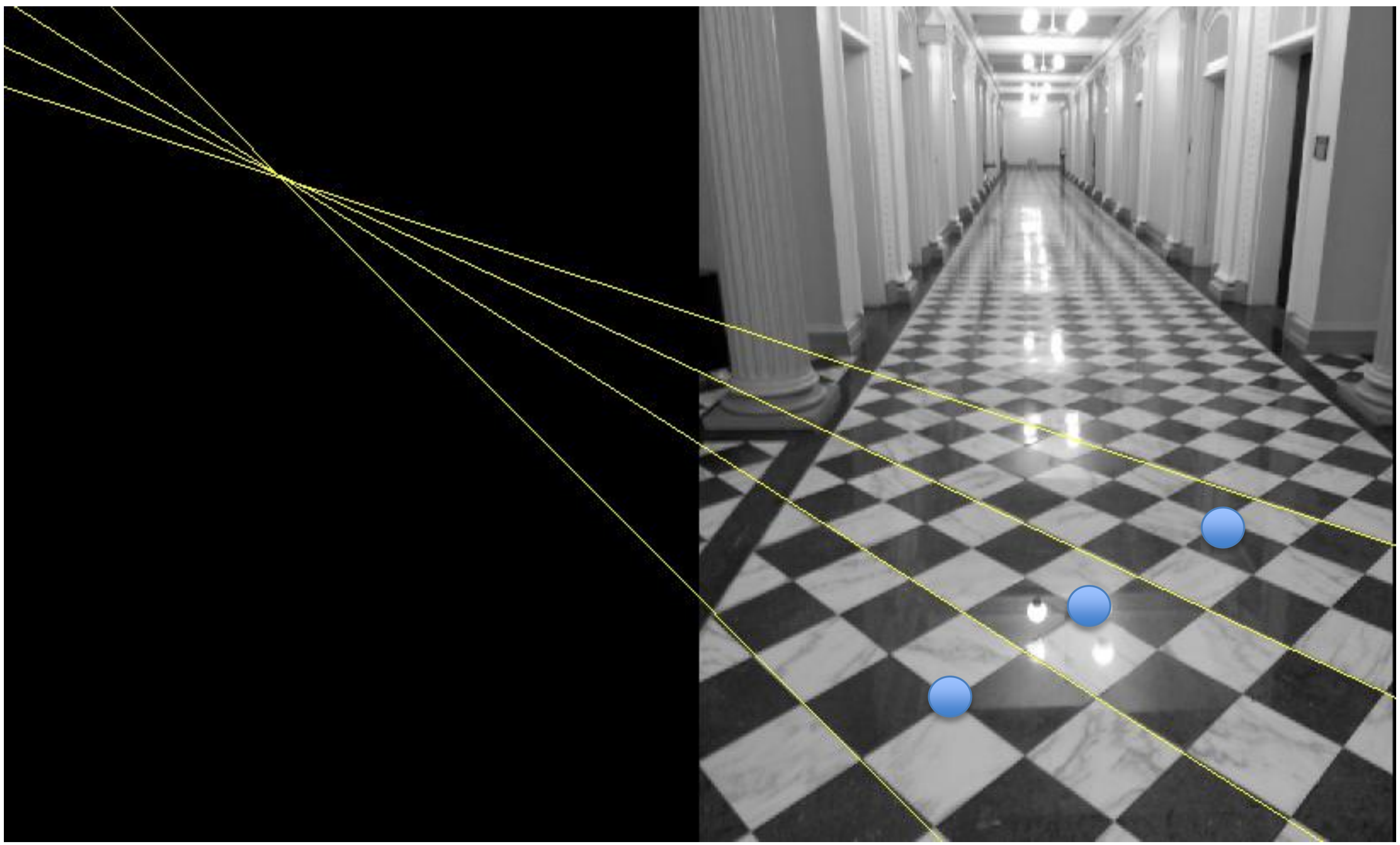
Invertibility means that $\det(A) \neq 0$ and that there exists $\lambda \neq 0$ such that $\lambda p' = Ap$.

Observe that we will write either $p' \sim Ap$ or $\lambda p' = Ap$.

A projective transformation is also known as **collineation or homography.**

A projective transformation preserves incidence:

- Three collinear points are mapped to three collinear points.
- and three concurrent lines are mapped to three concurrent lines.



Projective transformation of lines

If A maps a point to Ap, then where does a line l map to?

Line equation in original plane

$$l^T p = 0$$

Line equation in image plane $p' \sim Ap$

$$l^T A^{-1} p' = 0$$

implies that $l' = A^{-T} l$.

Projective transformation of lines

If A maps a point to Ap, then where does a line l map to?

Line equation in original plane

$$l^T p = 0$$

Line equation in image plane $p' \sim Ap$

$$l^T A^{-1} p' = 0$$

implies that $l' = A^{-T} l$.

Projective transformation of lines

If A maps a point to Ap, then where does a line l map to?

Line equation in original plane

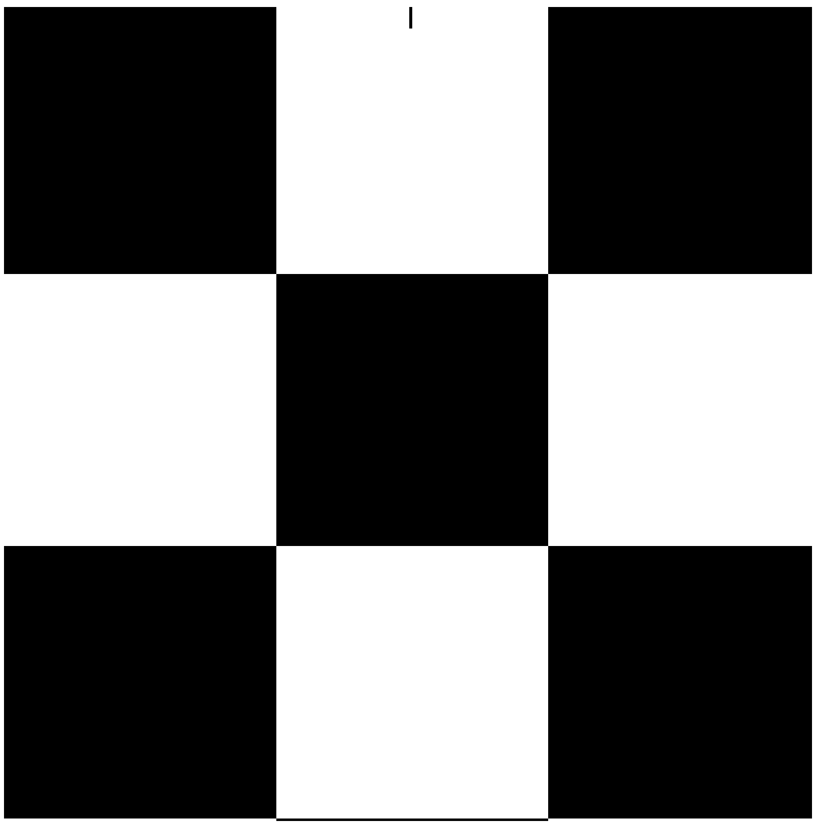
$$l^T p = 0$$

Line equation in image plane $p' \sim Ap$

$$l^T A^{-1} p' = 0$$

implies that $l' = A^{-T} l$.

How can we compute the projective transformation between

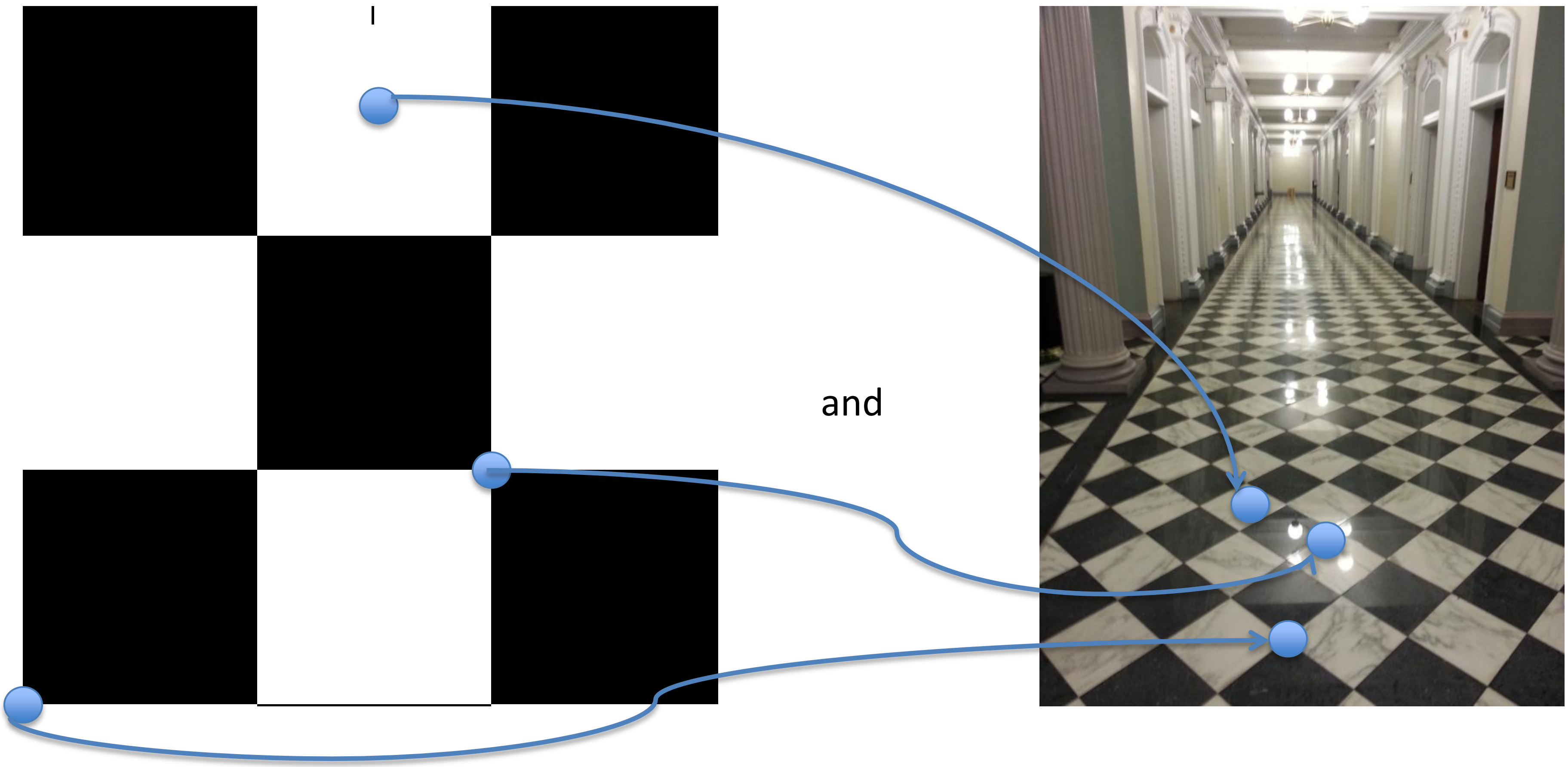


and

Floor tiles measured in [m]

Points in pixel coordinates

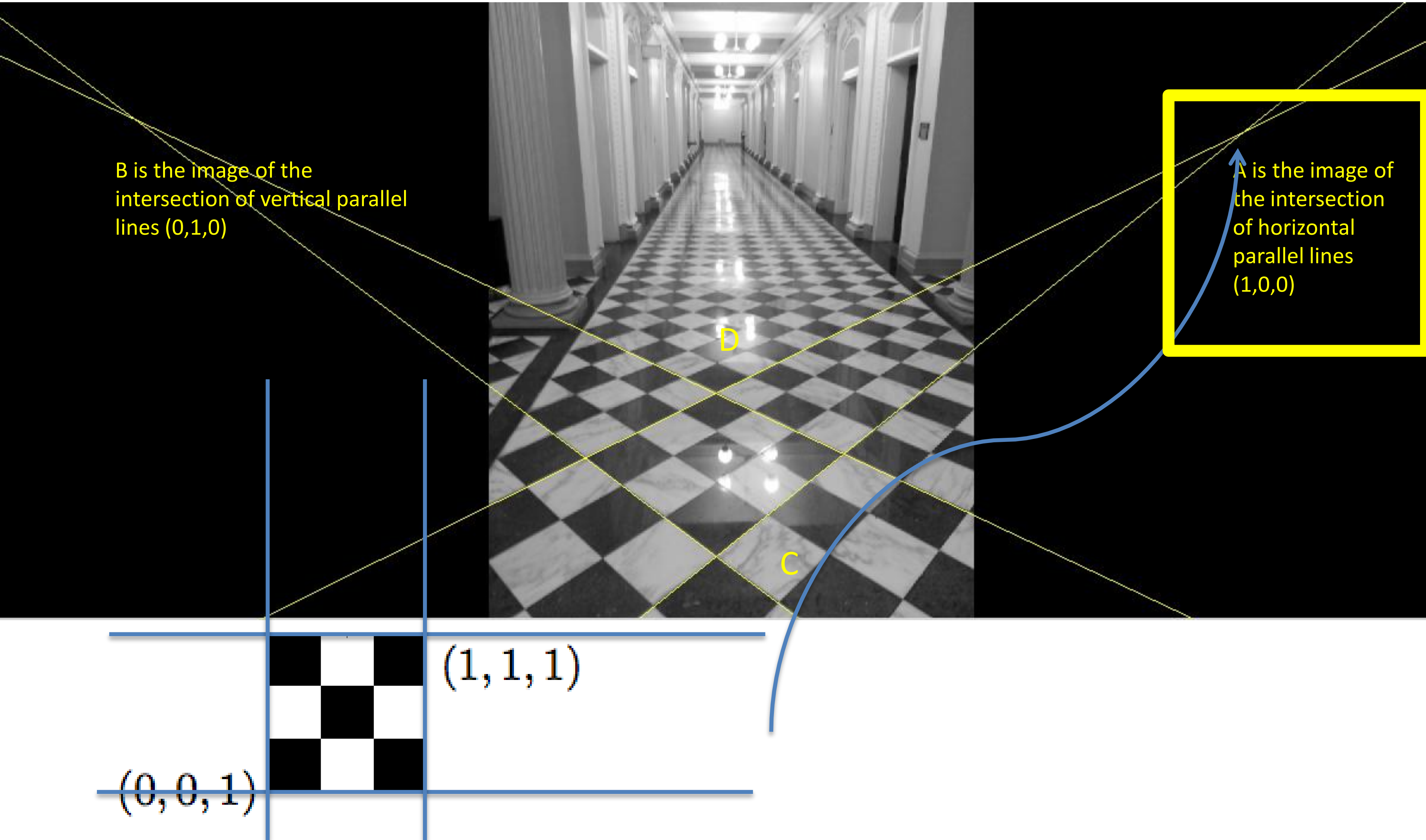
The result of such a transformation would map any point in one plane to the corresponding point in the other



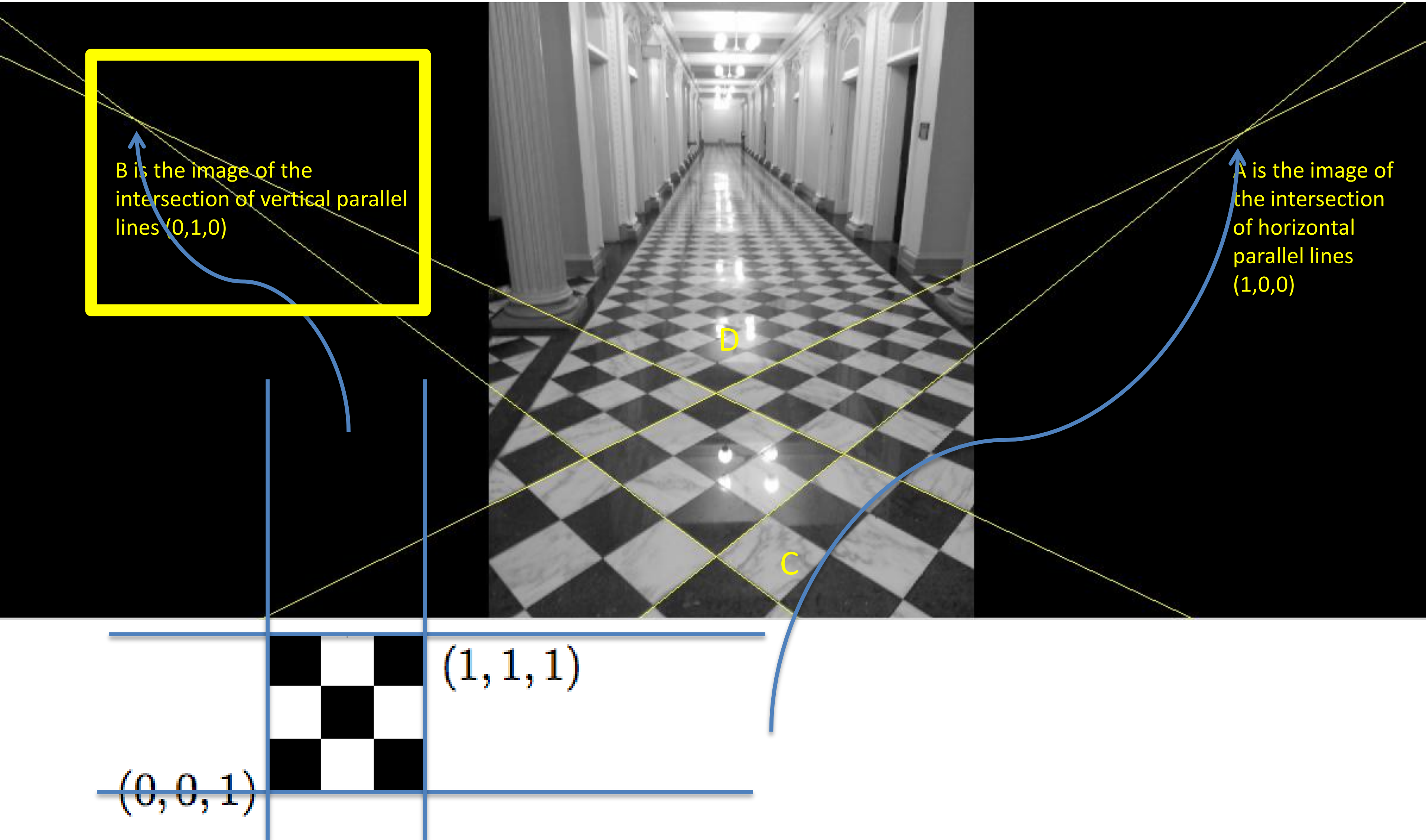
Floor tiles measured in [m]

Points in pixel coordinates

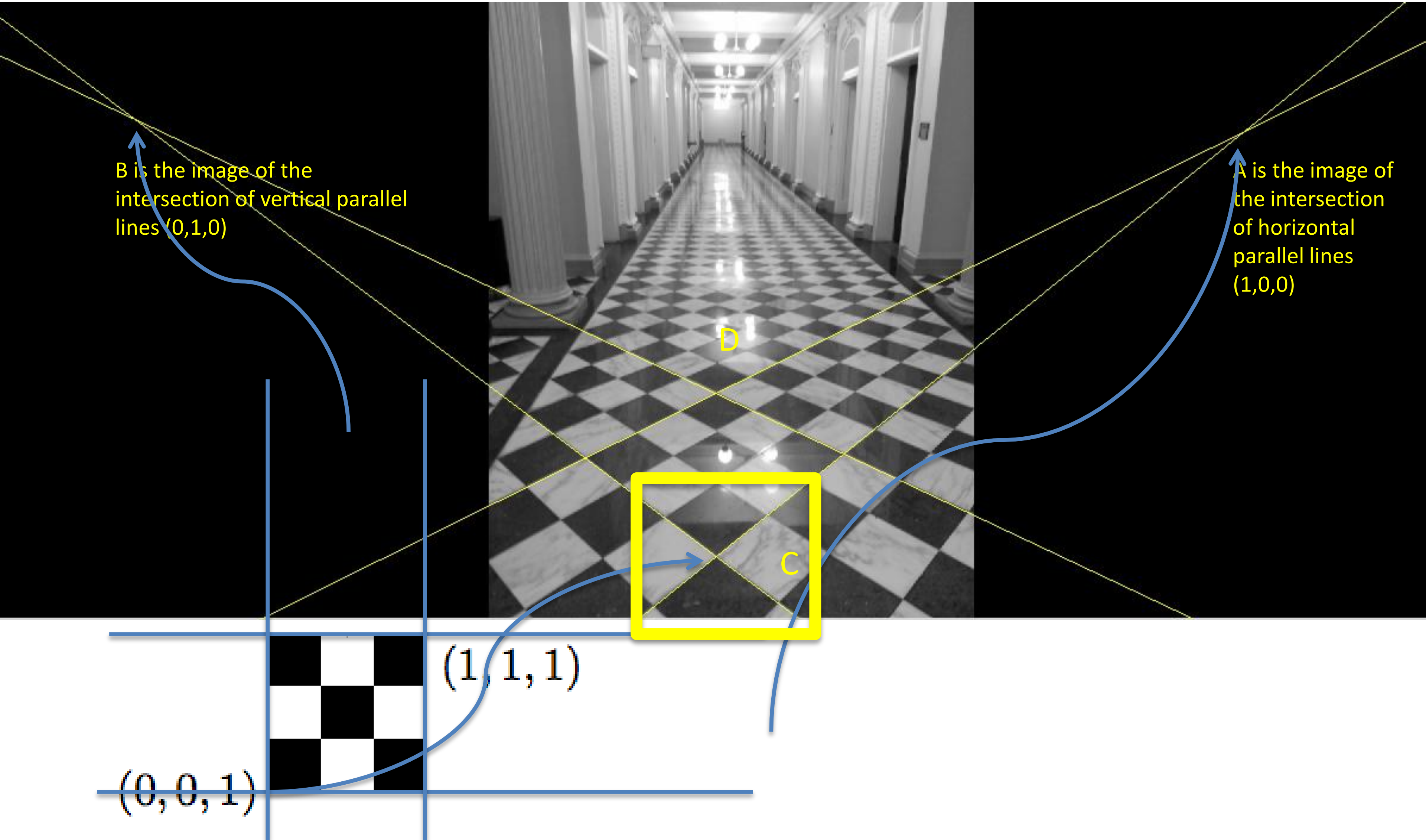
How can we compute the projective transformation between



How can we compute the projective transformation between



How can we compute the projective transformation between



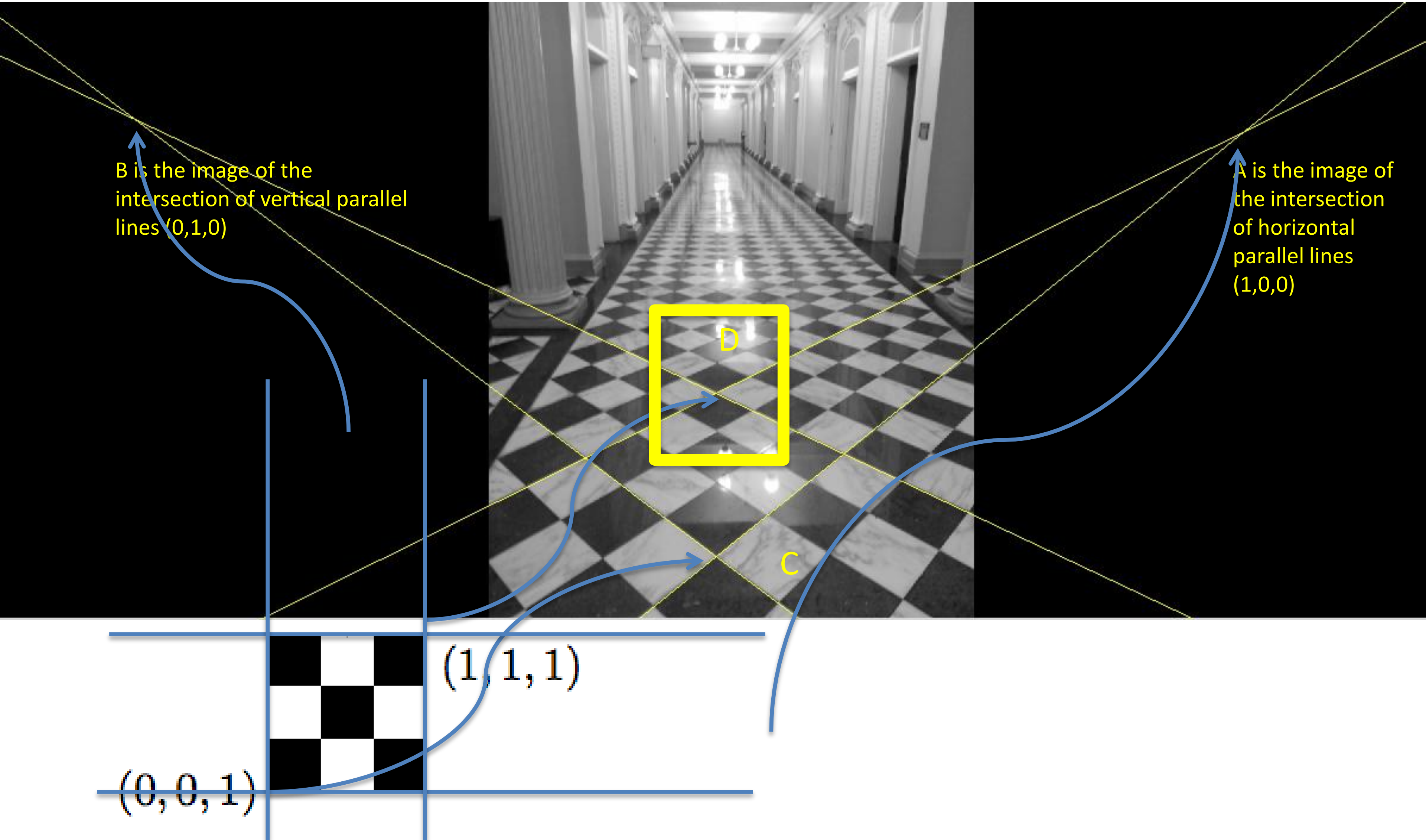
Assume that a mapping A maps the three points $(1,0,0)$, $(0,1,0)$, and $(0,0,1)$ to the non-collinear points A,B,C

with coordinate vectors a, b and $c \in \mathbb{P}^2$. Then the following is a possible projective transformation:

$$\begin{pmatrix} a & b & c \end{pmatrix} = \begin{pmatrix} \alpha a & \beta b & \gamma c \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

with 3 degrees of freedoms α, β and γ. This means 3 points do not suffice to compute a projective transformation.

Let us introduce a 4th point D



Let us assume that the same A maps $(1,1,1)$ to the point d. Then, the following should hold:

$$\lambda d = \begin{pmatrix} \alpha a & \beta b & \gamma c \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

hence

$$\lambda d = \alpha a + \beta b + \gamma c.$$

There always exist such $\lambda, \alpha, \beta, \gamma$ because four elements of $\mathbb{R}^3 \setminus \{(0,0,0)\}$ are always linearly dependent.

Because a, b, c are not collinear, there exist unique $\alpha/\lambda, \beta/\lambda, \gamma/\lambda$ for writing this linear combination.

Since A is the same as A/λ we solve for α, β, γ such that $d = \alpha a + \beta b + \gamma c$, which can be written as a linear system

$$\begin{pmatrix} a & b & c \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = d.$$

Since a, b, c are not collinear we can always find a unique triple α, β, γ. The resulting projective transformation is $A = \begin{pmatrix} \alpha a & \beta b & \gamma c \end{pmatrix}$.

There always exist such $\lambda, \alpha, \beta, \gamma$ because four elements of $\mathbb{R}^3 \setminus \{(0,0,0)\}$ are always linearly dependent.

Because a, b, c are not collinear, there exist unique $\alpha/\lambda, \beta/\lambda, \gamma/\lambda$ for writing this linear combination.

Since A is the same as A/λ we solve for α, β, γ such that $d = \alpha a + \beta b + \gamma c$, which can be written as a linear system

$$\begin{pmatrix} a & b & c \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = d.$$

Since a, b, c are not collinear we can always find a unique triple α, β, γ. The resulting projective transformation is $A = \begin{pmatrix} \alpha a & \beta b & \gamma c \end{pmatrix}$.

There always exist such $\lambda, \alpha, \beta, \gamma$ because four elements of $\mathbb{R}^3 \setminus \{(0,0,0)\}$ are always linearly dependent.

Because a, b, c are not collinear, there exist unique $\alpha/\lambda, \beta/\lambda, \gamma/\lambda$ for writing this linear combination.

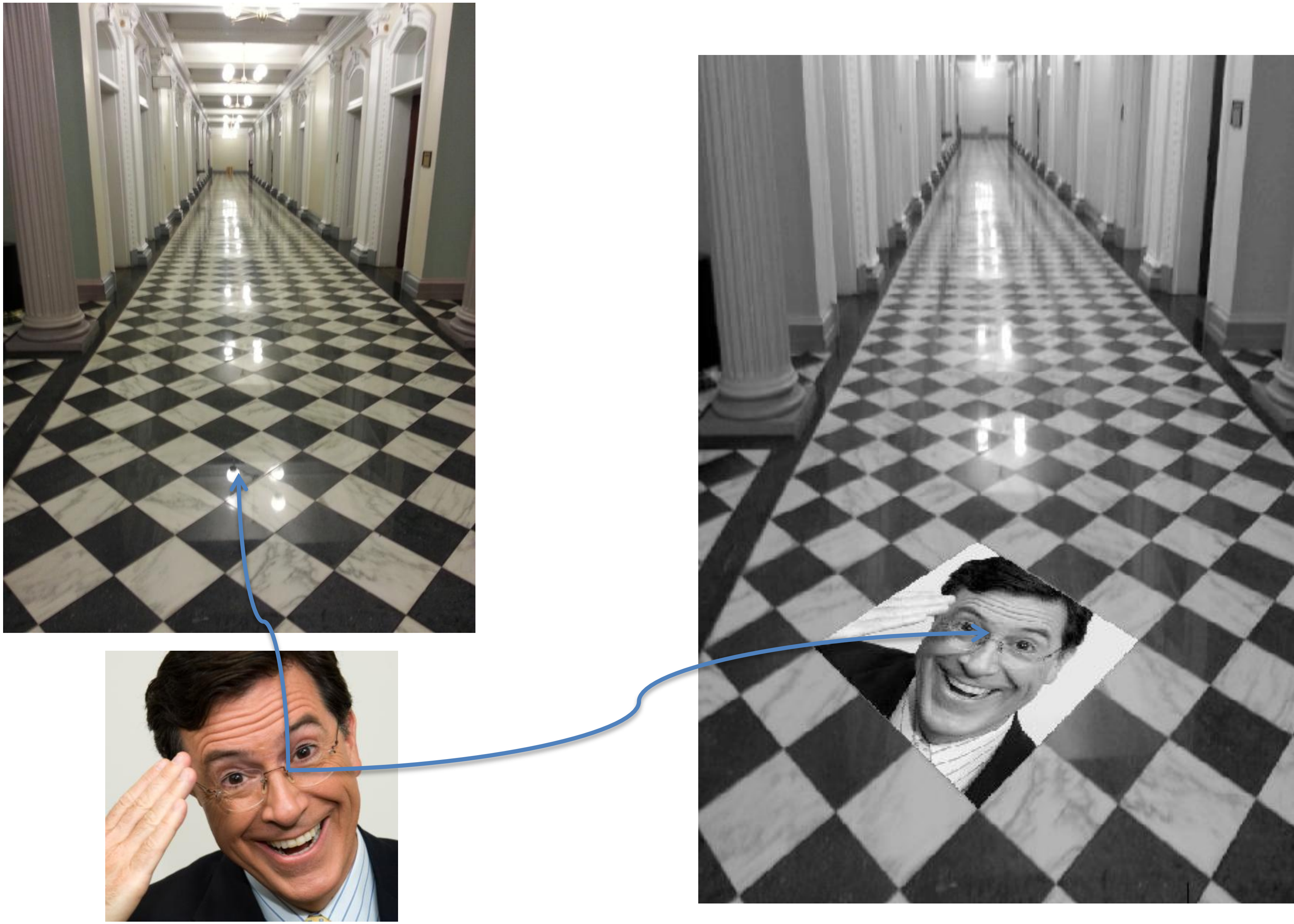
Since A is the same as A/λ we solve for α, β, γ such that $d = \alpha a + \beta b + \gamma c$, which can be written as a linear system

$$\begin{pmatrix} a & b & c \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = d.$$

Since a, b, c are not collinear we can always find a unique triple α, β, γ. The resulting projective transformation is $A = \begin{pmatrix} \alpha a & \beta b & \gamma c \end{pmatrix}$.

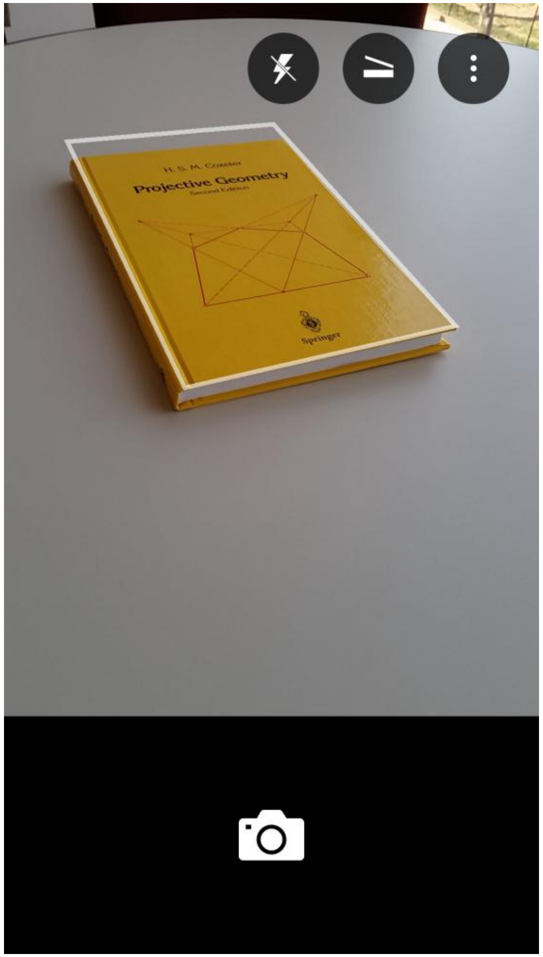
Four points not three of them collinear suffice to recover unambiguously a projective transformation.

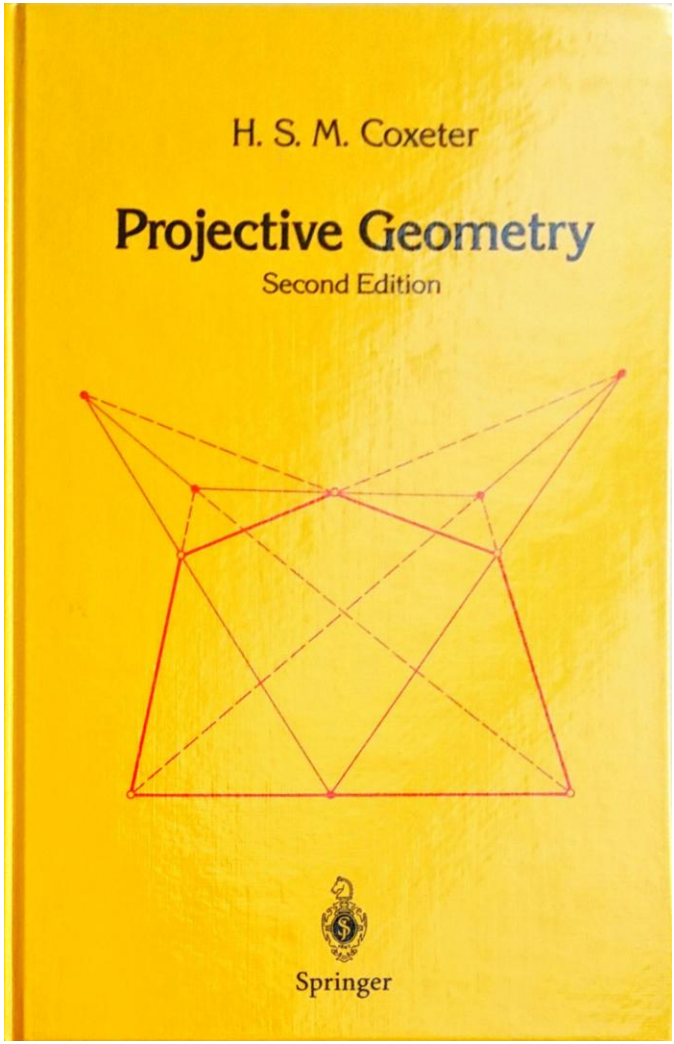
Knowledge of this projective transformation makes Virtual Billboards possible!



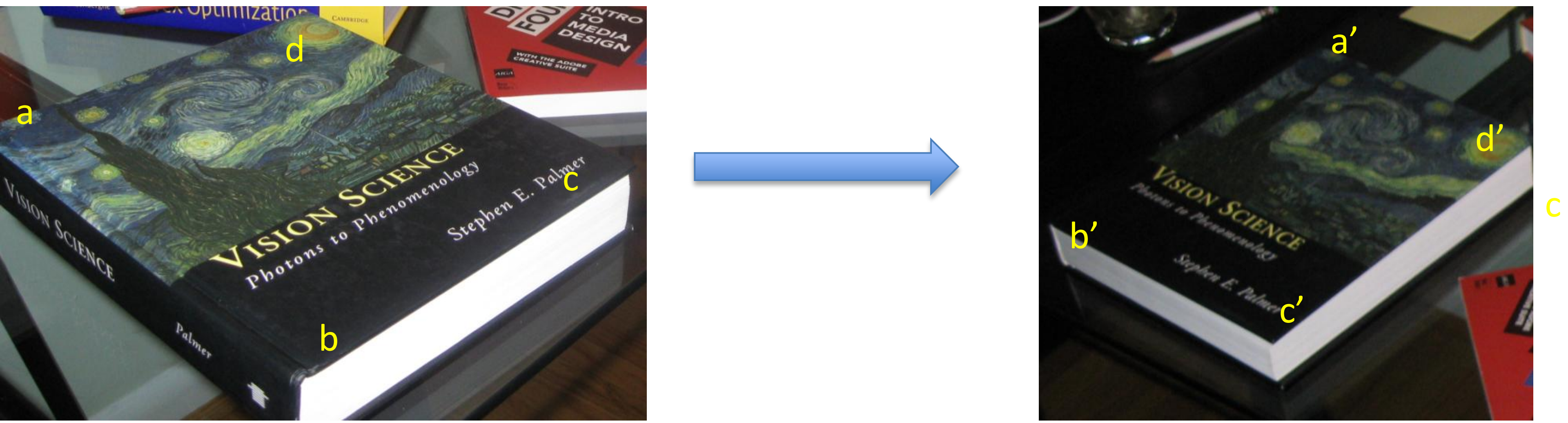
Microsoft Office Lens App

Office Lens


H. S. M. Coxeter
Projective Geometry
Second Edition
Springer


H. S. M. Coxeter
Projective Geometry
Second Edition
Springer

What happens when the original set of points is not a square?



Find projective transformation mapping $(a, b, c, d) \rightarrow (a', b', c', d')$:

To determine this mapping we go through the four canonical points.

We find the mapping from $(1,0,0), etc$ to (a,b,c,d) and we call it T:

$$a \sim T(1,0,0)^T, etc$$

We find the mapping from $(1,0,0), etc$ to (a',b',c',d') and we call it T':

$$a' \sim T'(1,0,0)^T, etc$$

Then, back-substituing $(1,0,0)^T \sim T^{-1}a, etc$ we obtain that

$$a' = T'T^{-1}a, etc$$

To determine this mapping we go through the four canonical points.

We find the mapping from $(1,0,0), etc$ to (a,b,c,d) and we call it T:

$$a \sim T(1,0,0)^T, etc$$

We find the mapping from $(1,0,0), etc$ to (a',b',c',d') and we call it T':

$$a' \sim T'(1,0,0)^T, etc$$

Then, back-substituing $(1,0,0)^T \sim T^{-1}a, etc$ we obtain that

$$a' = T'T^{-1}a, etc$$

To determine this mapping we go through the four canonical points.

We find the mapping from $(1,0,0), etc$ to (a,b,c,d) and we call it T:

$$a \sim T(1,0,0)^T, etc$$

We find the mapping from $(1,0,0), etc$ to (a',b',c',d') and we call it T':

$$a' \sim T'(1,0,0)^T, etc$$

Then, back-substituing $(1,0,0)^T \sim T^{-1}a, etc$ we obtain that

$$a' = T'T^{-1}a, etc$$

Perception: Projective Transformations and Vanishing Points

Kostas Daniilidis

Projective Transformation

- Aka Homography or Collineation
- Represents the perspective projection from a ground plane to an image plane !
- It is an invertible 3x3 matrix but has 8 independent parameters
- For example if (X,Y) measured in meters on the ground and (u,v) in pixels

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \sim H \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}$$

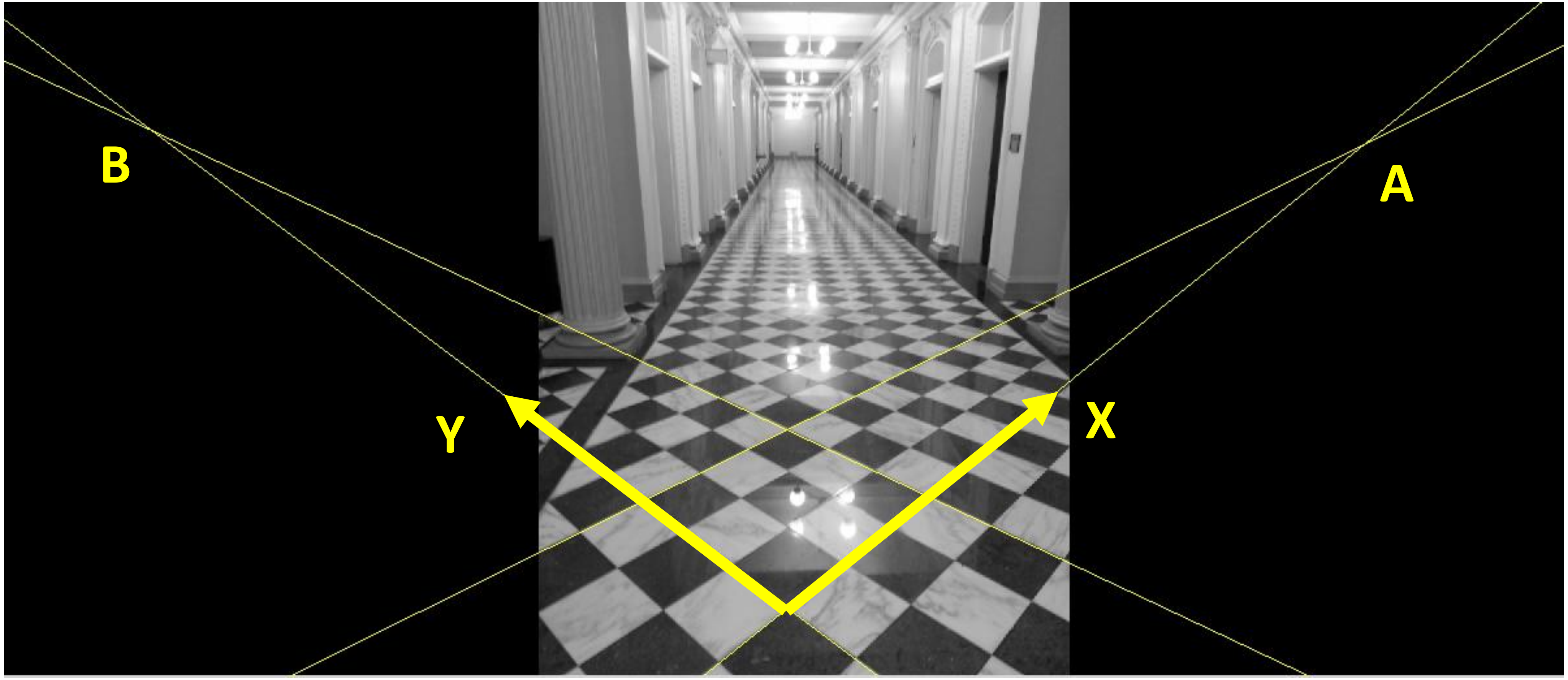
or

$$\begin{aligned} u &= \frac{H_{11}X + H_{12}Y + H_{13}}{H_{31}X + H_{32}Y + H_{33}} \\ v &= \frac{H_{21}X + H_{22}Y + H_{23}}{H_{31}X + H_{32}Y + H_{33}} \end{aligned}$$

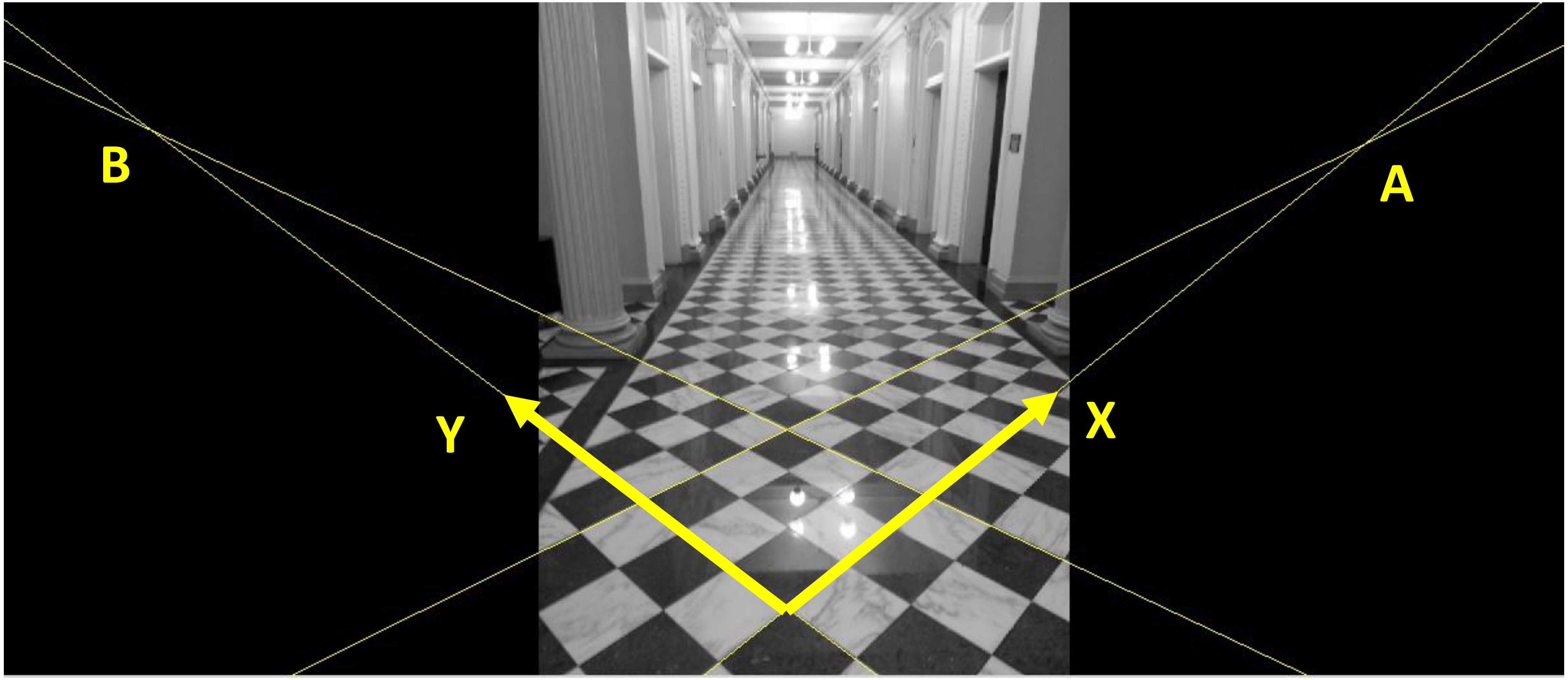
If we know H, we can make back and forth measurements:

10.2,2.3 in meters

Homography columns as vanishing points

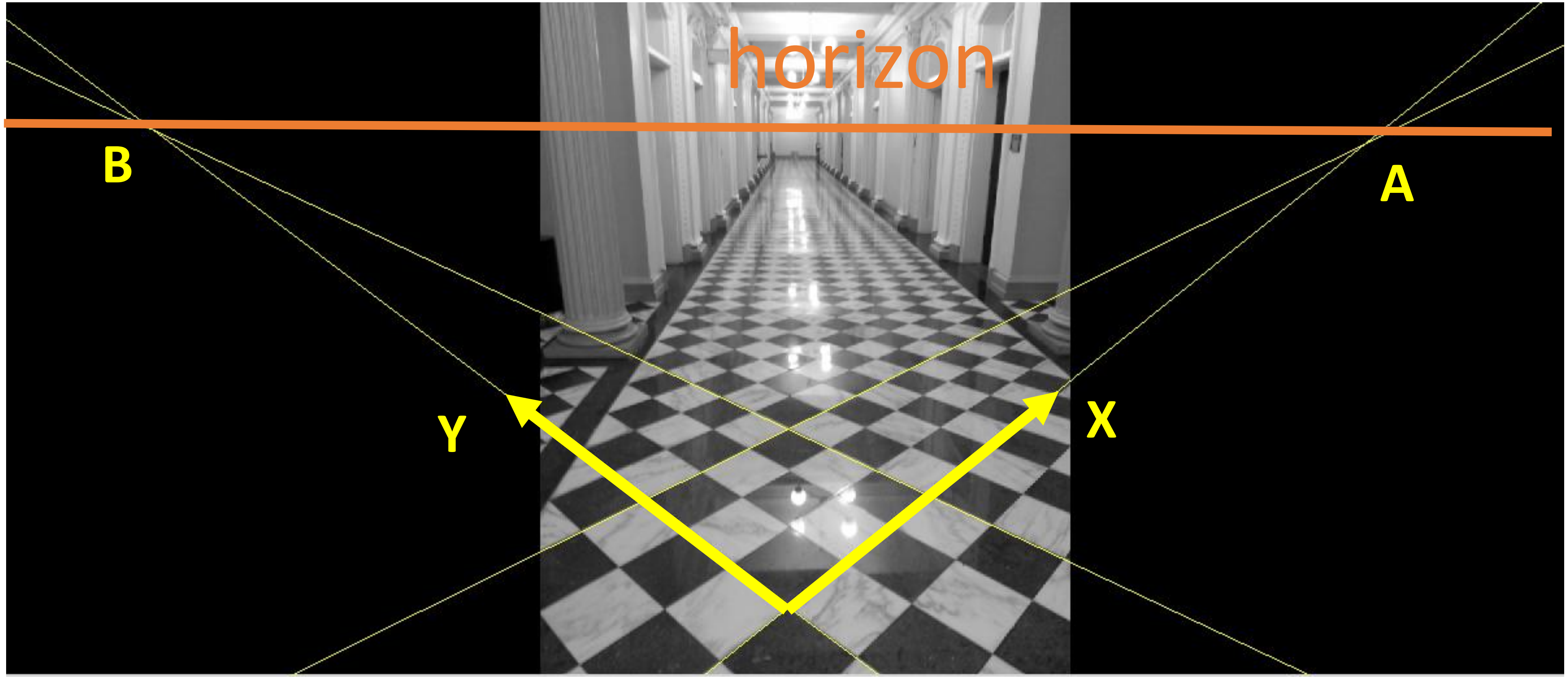


If $H = \begin{pmatrix} h_1 & h_2 & h_3 \end{pmatrix}$ then $h_1 \sim A$ and $h_2 \sim B$.

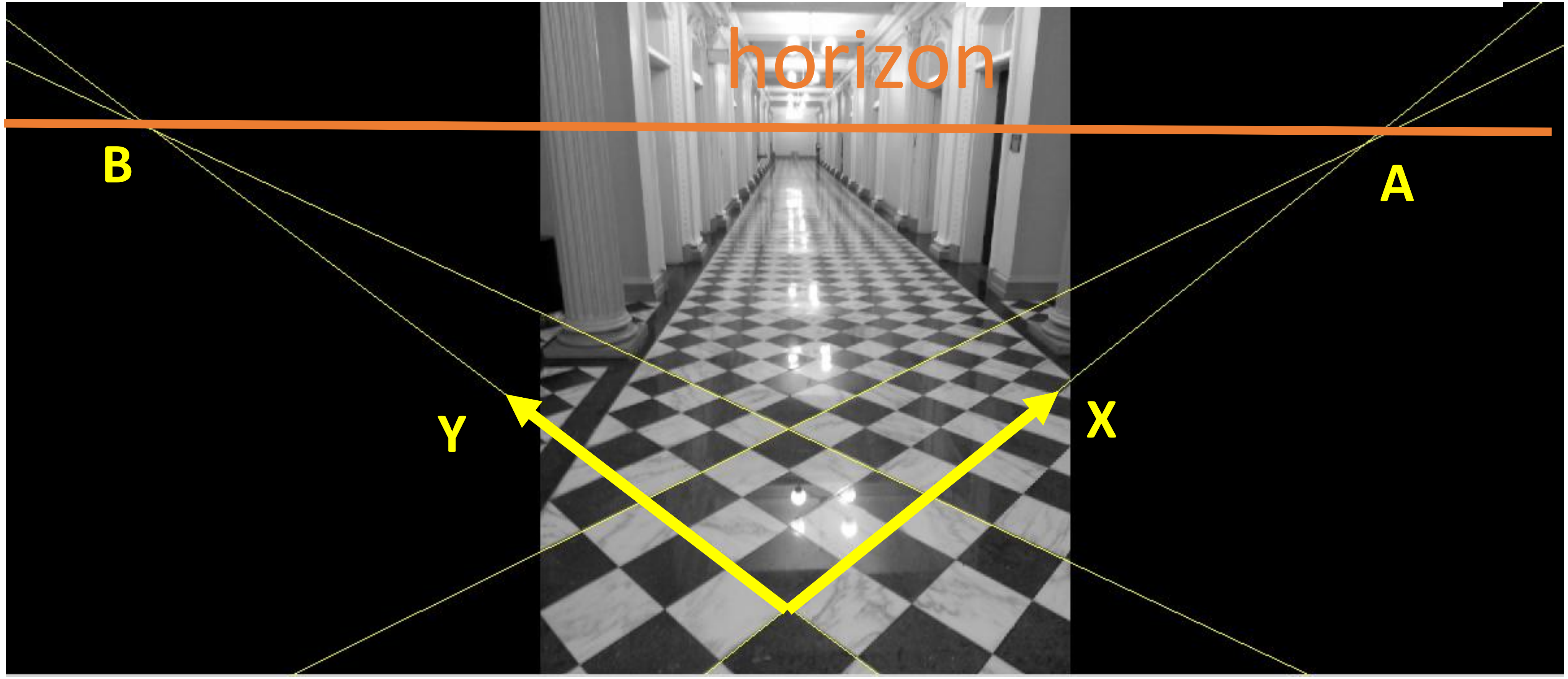


So the first two columns are two Orthogonal vanishing points

If we connect two vanishing points we obtain the horizon!

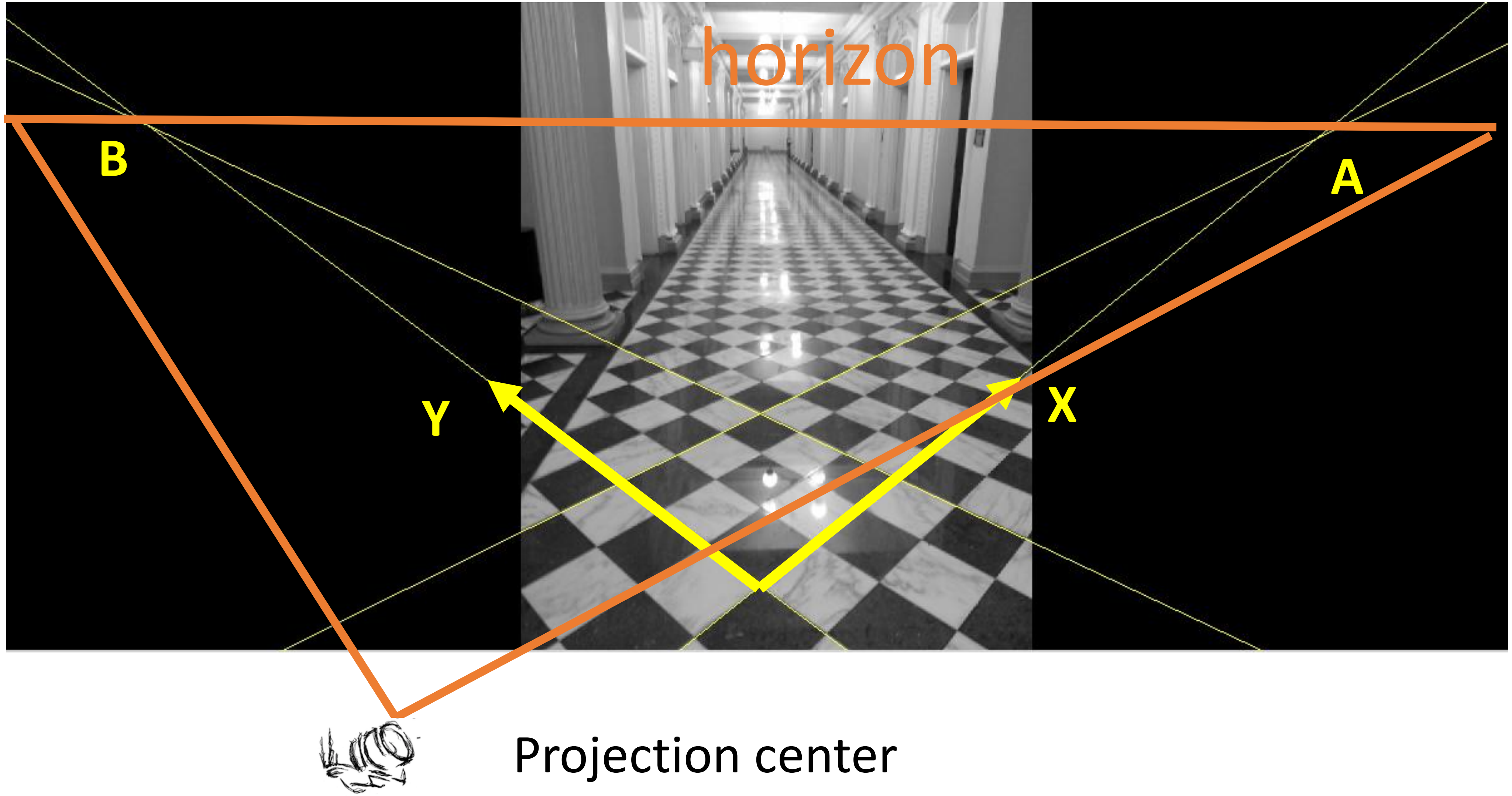


Equation of horizon: $$(h_1 \times h_2)^T \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0$$



Horizon with projection center build a horizon plane with normal

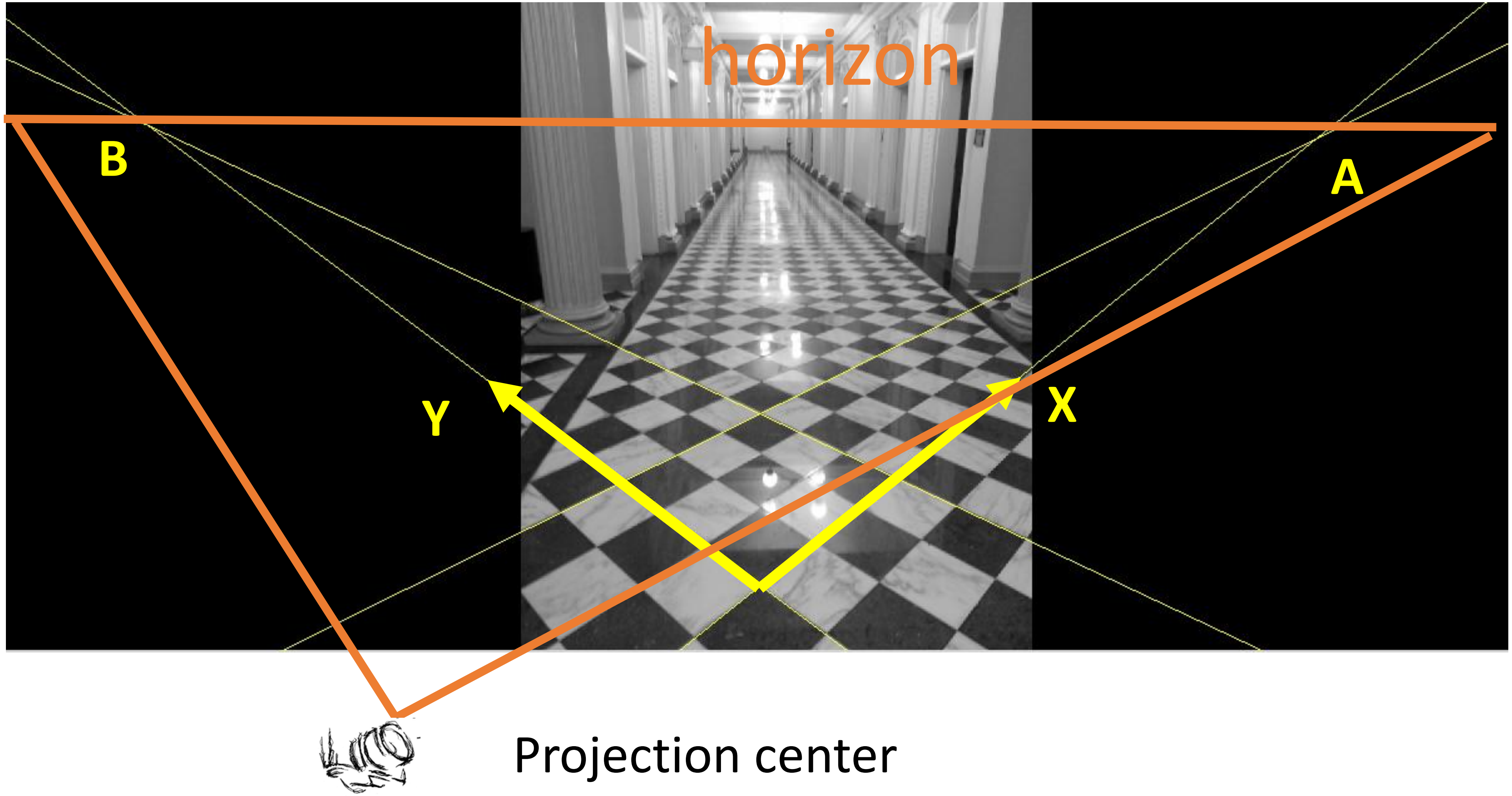
$$(h_1 \times h_2)$$



Projection center

The horizon plane is parallel to the ground plane and hence

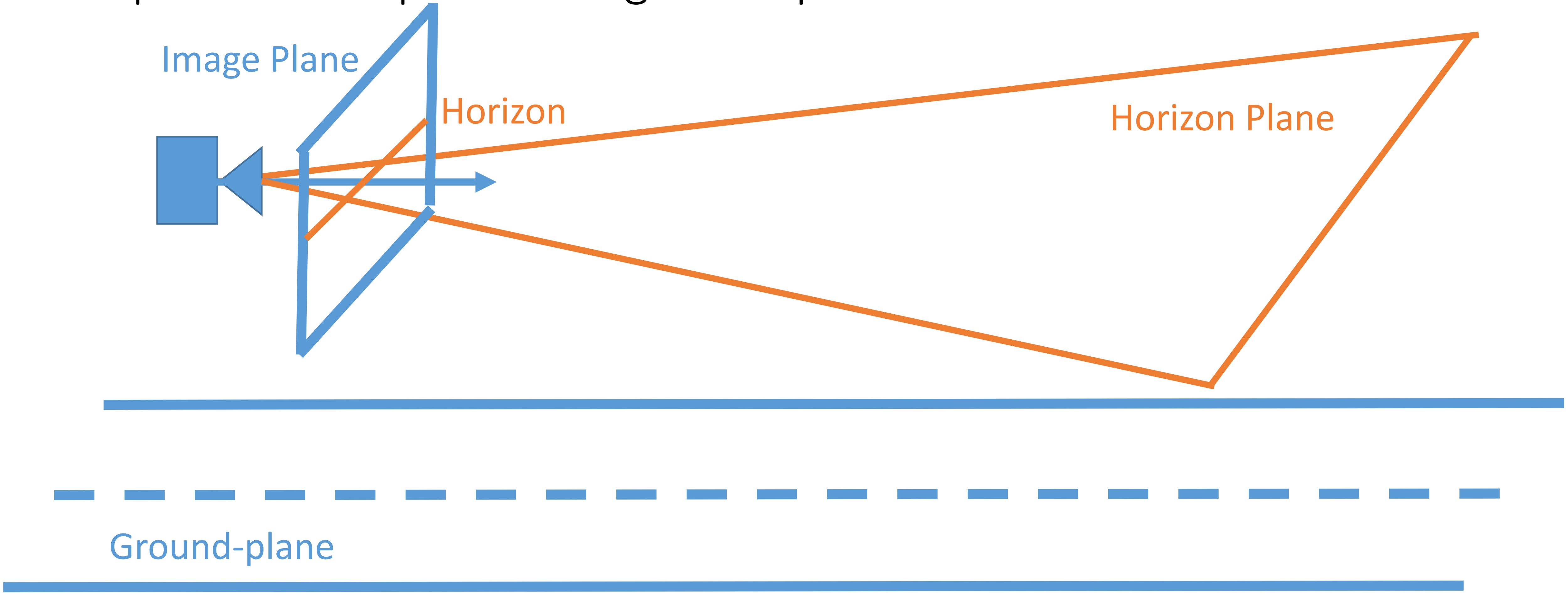
$(h_1 \times h_2)$ is the normal to the ground plane expressed via pixels!



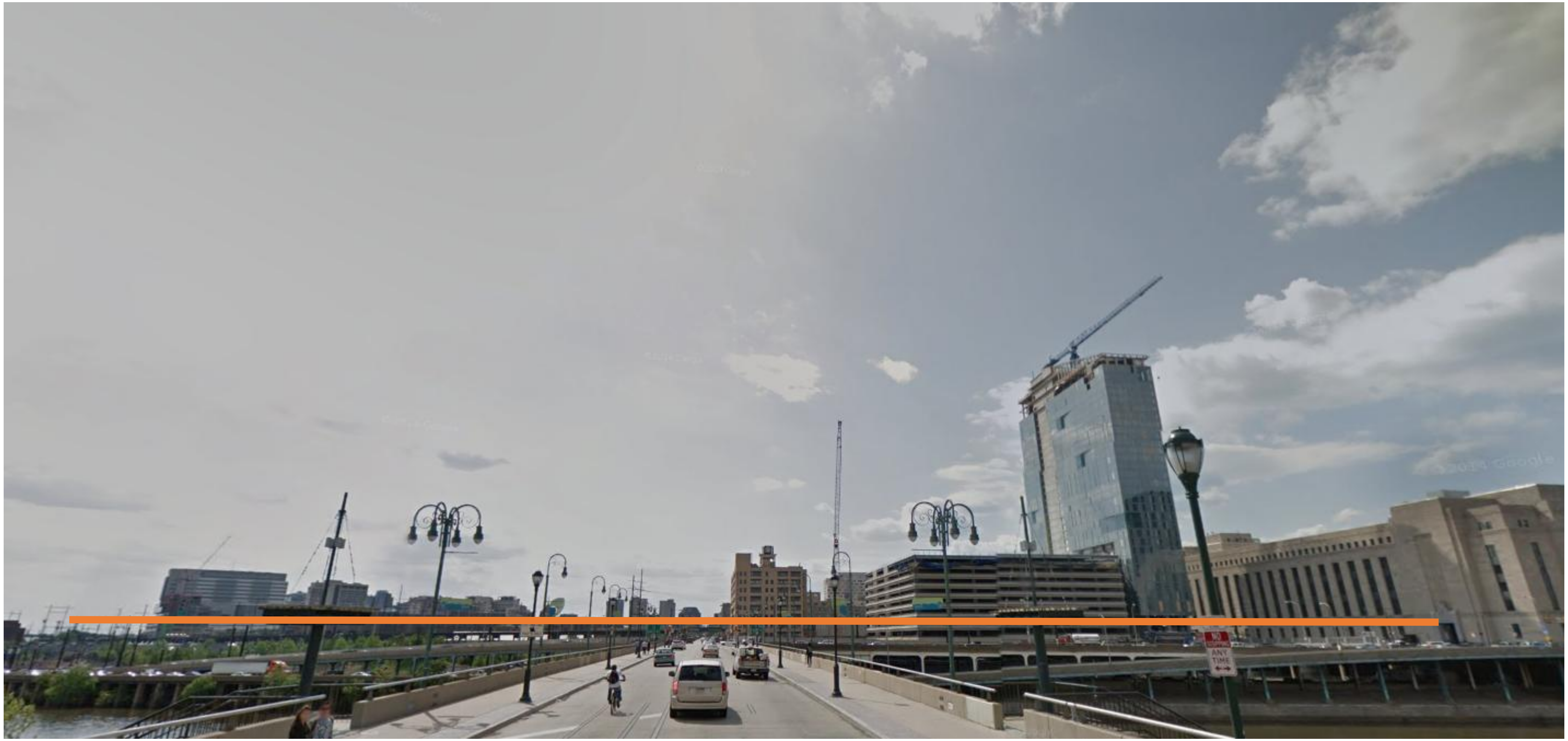
Projection center

Horizon gives complete info about how ground plane is oriented! If horizon is horizontal the center it means that optical axis is parallel to groundplane!

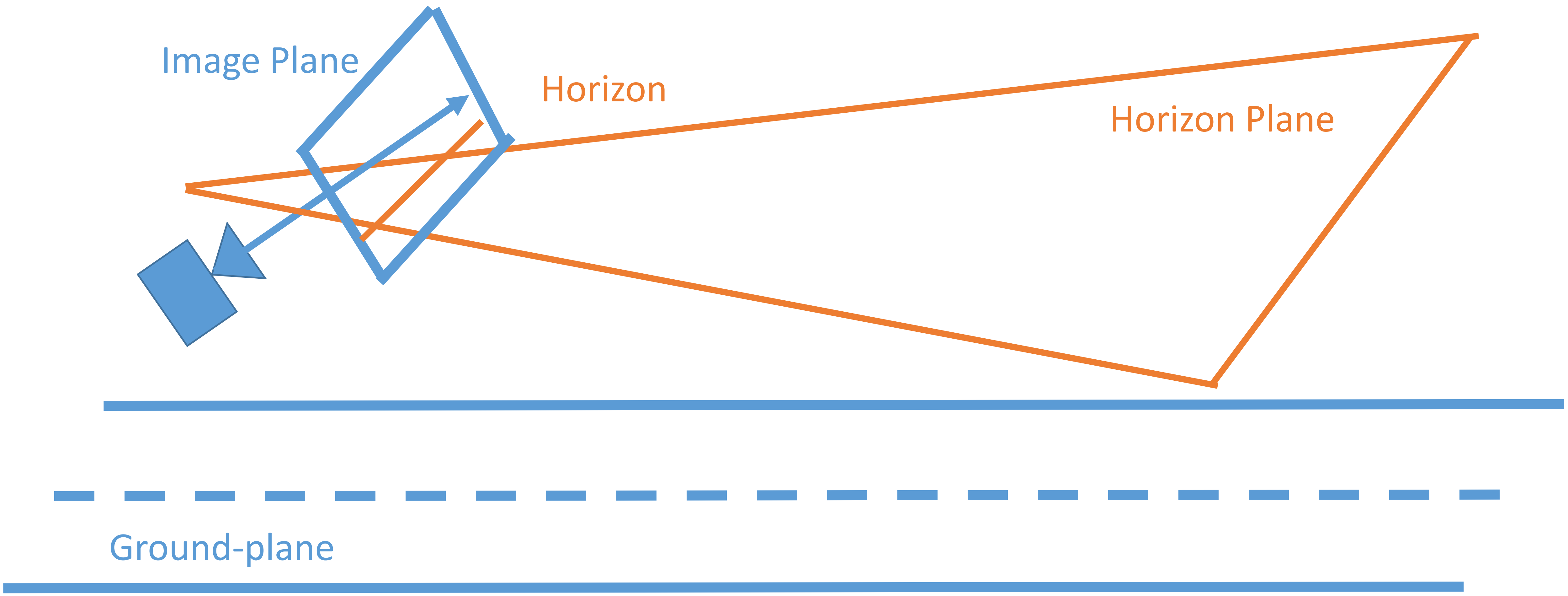
Horizon gives complete info about how ground plane is oriented! If horizon is horizontal the center it means that optical axis is parallel to ground-plane!



If horizon moves to the bottom it means we look upwards!

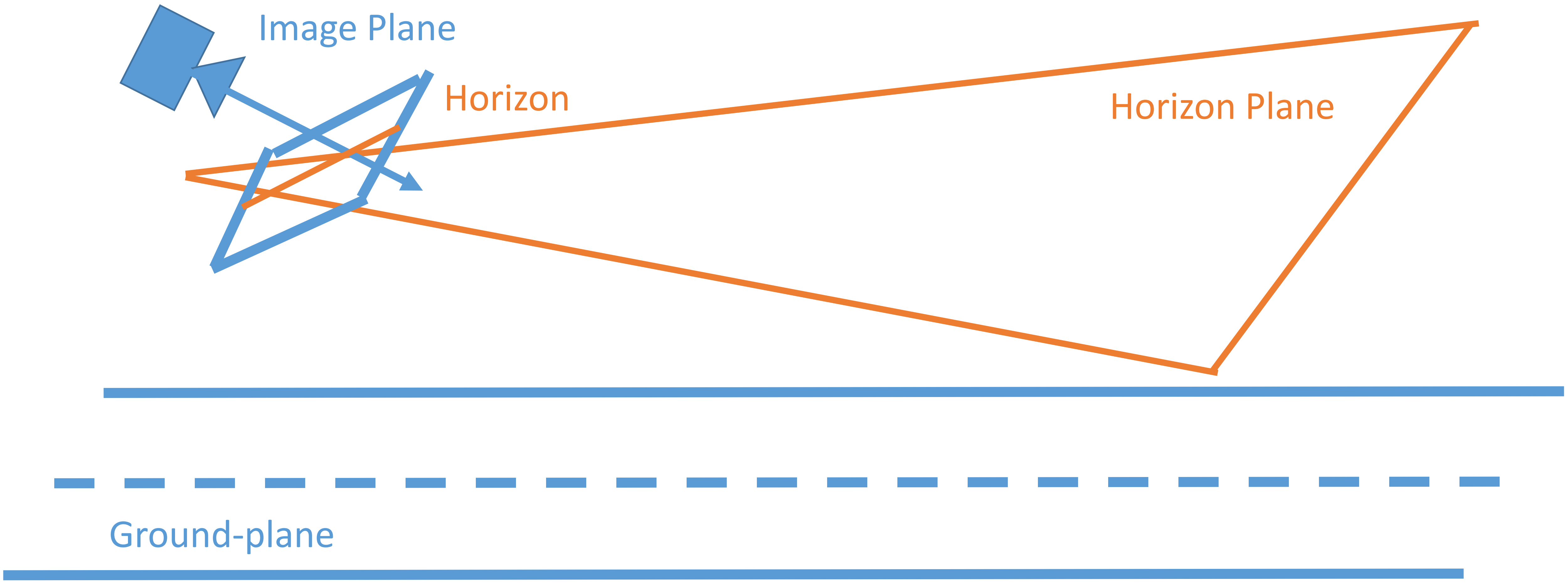


If horizon moves to the bottom it means we look upwards!

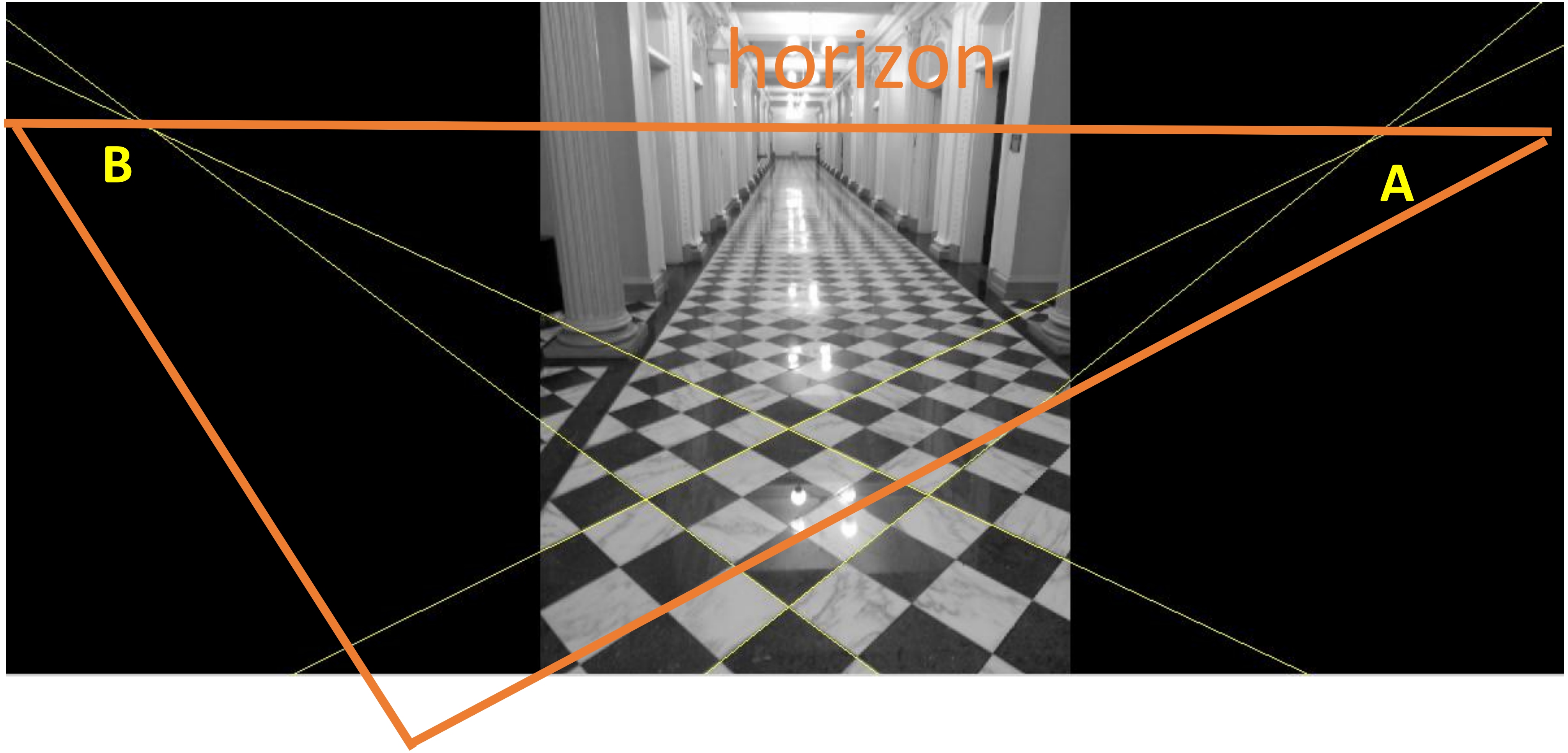


If horizon moves to the top it means we look downwards !

If horizon moves to the bottom it means we look upwards!



Horizon tells us how camera is oriented.
Constrains the collineation !



Perception: Projective Transformations and Vanishing Points

Kostas Daniilidis

Projective Transformation

- Aka Homography or Collineation
- Represents the perspective projection from a ground plane to an image plane !
- It is an invertible 3x3 matrix but has 8 independent parameters
- For example if (X,Y) measured in meters on the ground and (u,v) in pixels

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \sim H \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}$$

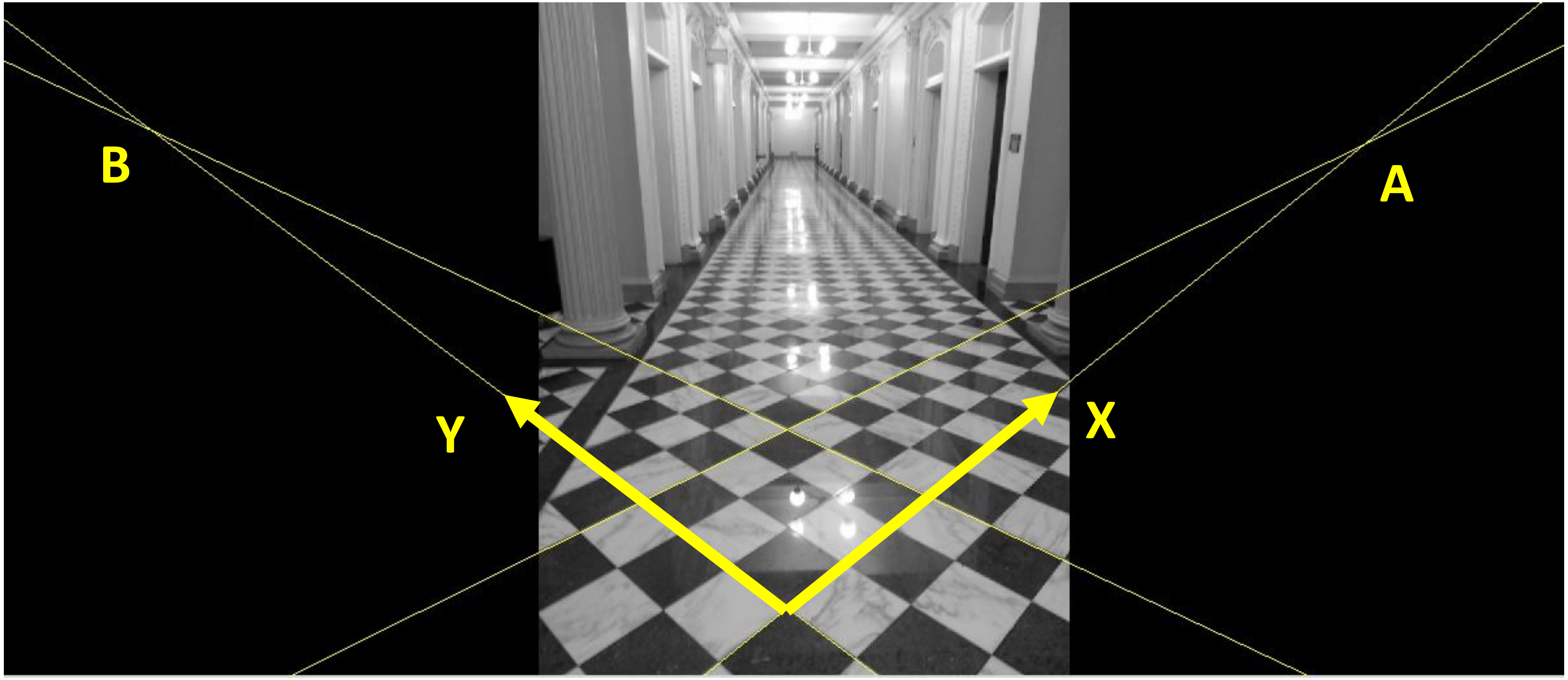
or

$$u = \frac{H_{11}X + H_{12}Y + H_{13}}{H_{31}X + H_{32}Y + H_{33}}$$

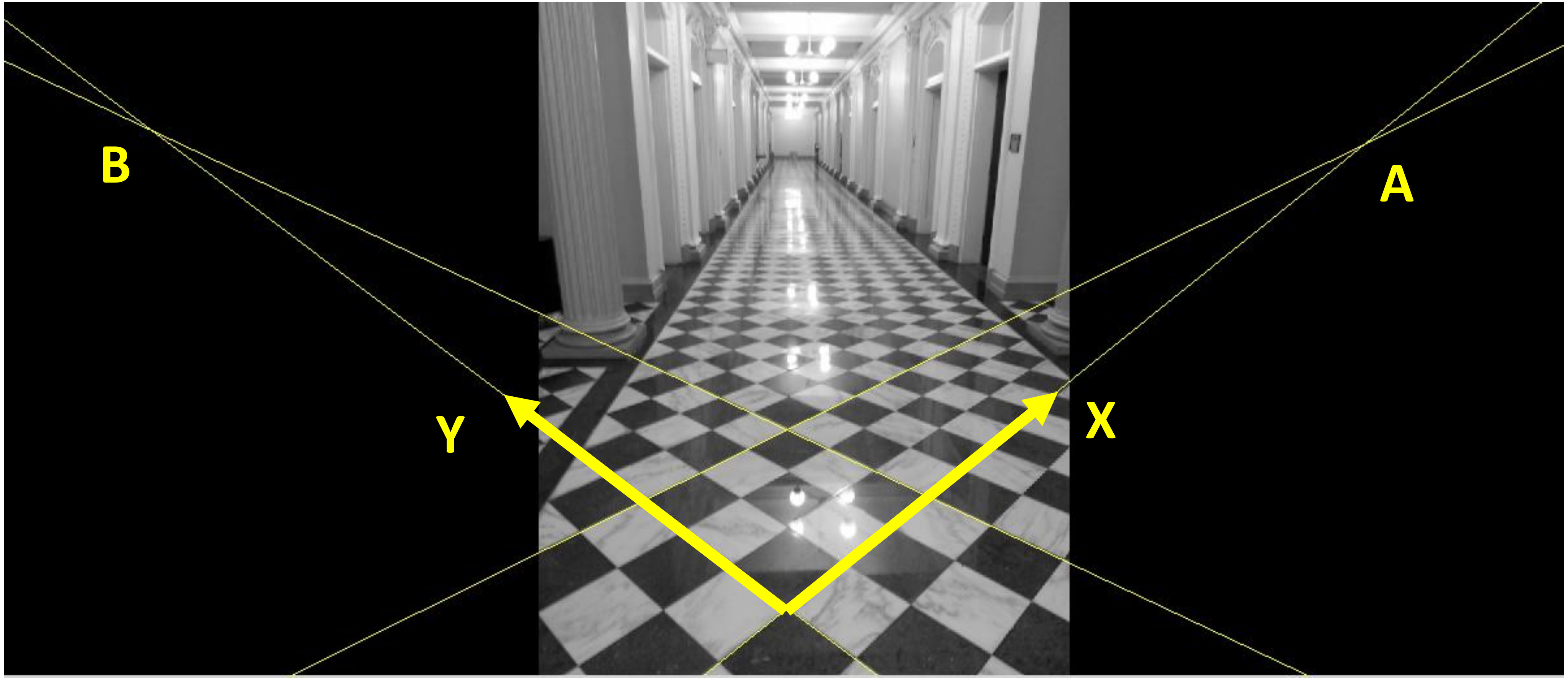
$$v = \frac{H_{21}X + H_{22}Y + H_{23}}{H_{31}X + H_{32}Y + H_{33}}$$

If we know H, we can make back and forth measurements:

Homography columns as vanishing points

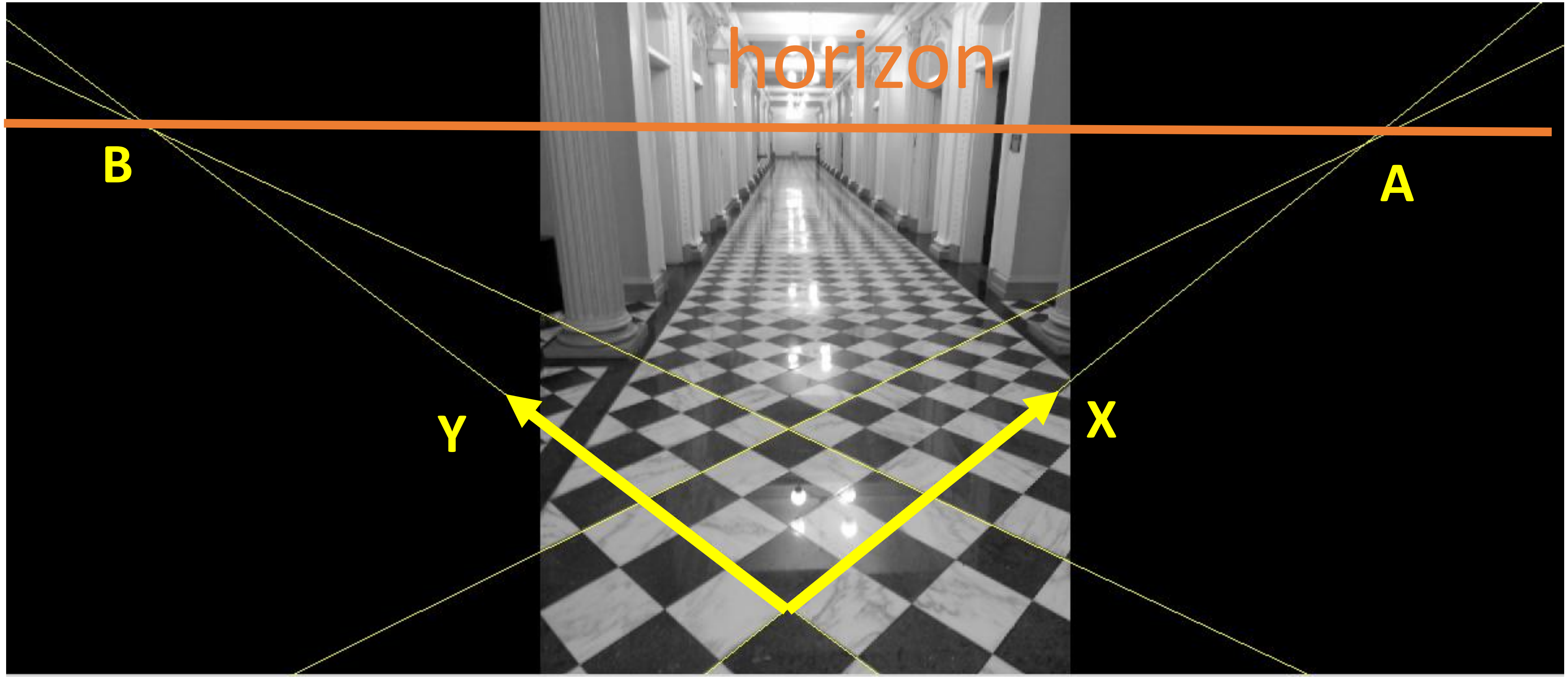


If $H = \begin{pmatrix} h_1 & h_2 & h_3 \end{pmatrix}$ then $h_1 \sim A$ and $h_2 \sim B$.

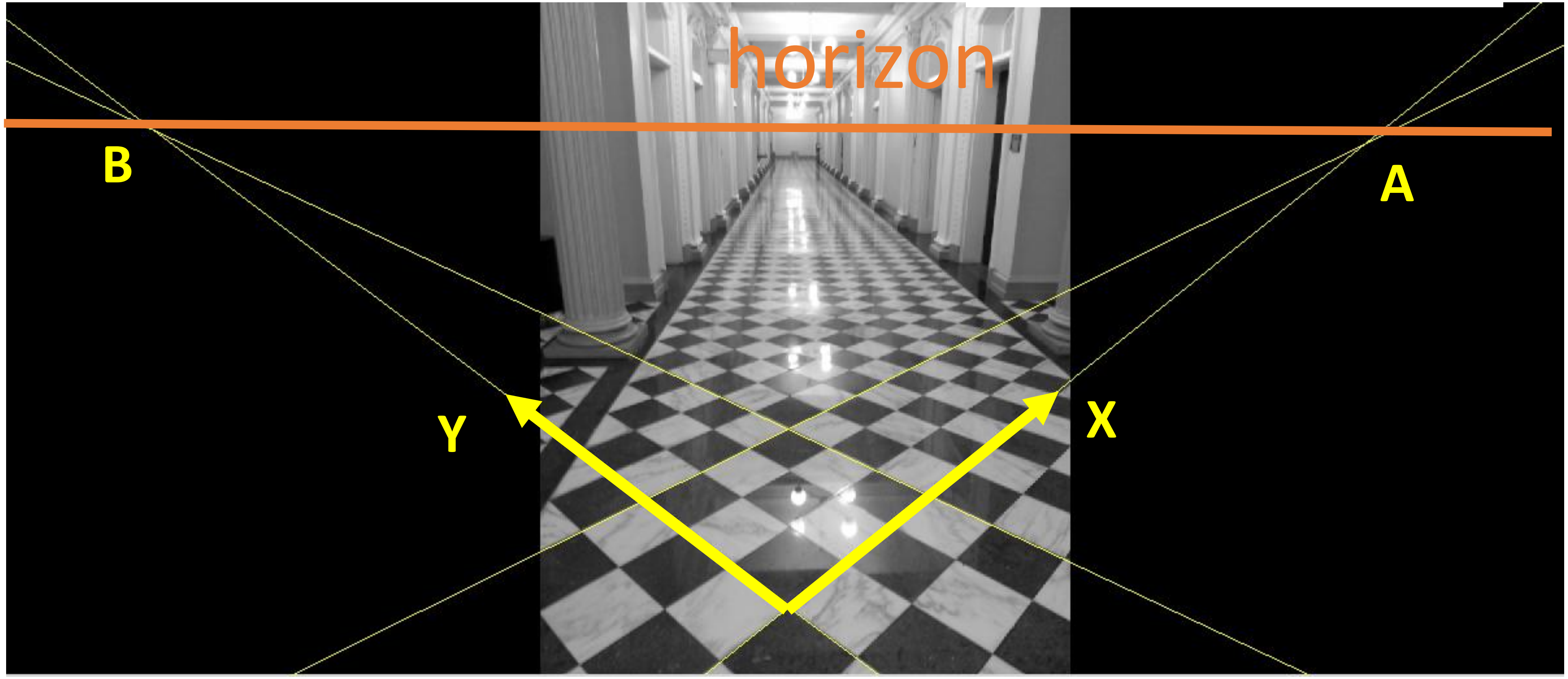


So the first two columns are two Orthogonal vanishing points

If we connect two vanishing points we obtain the horizon!

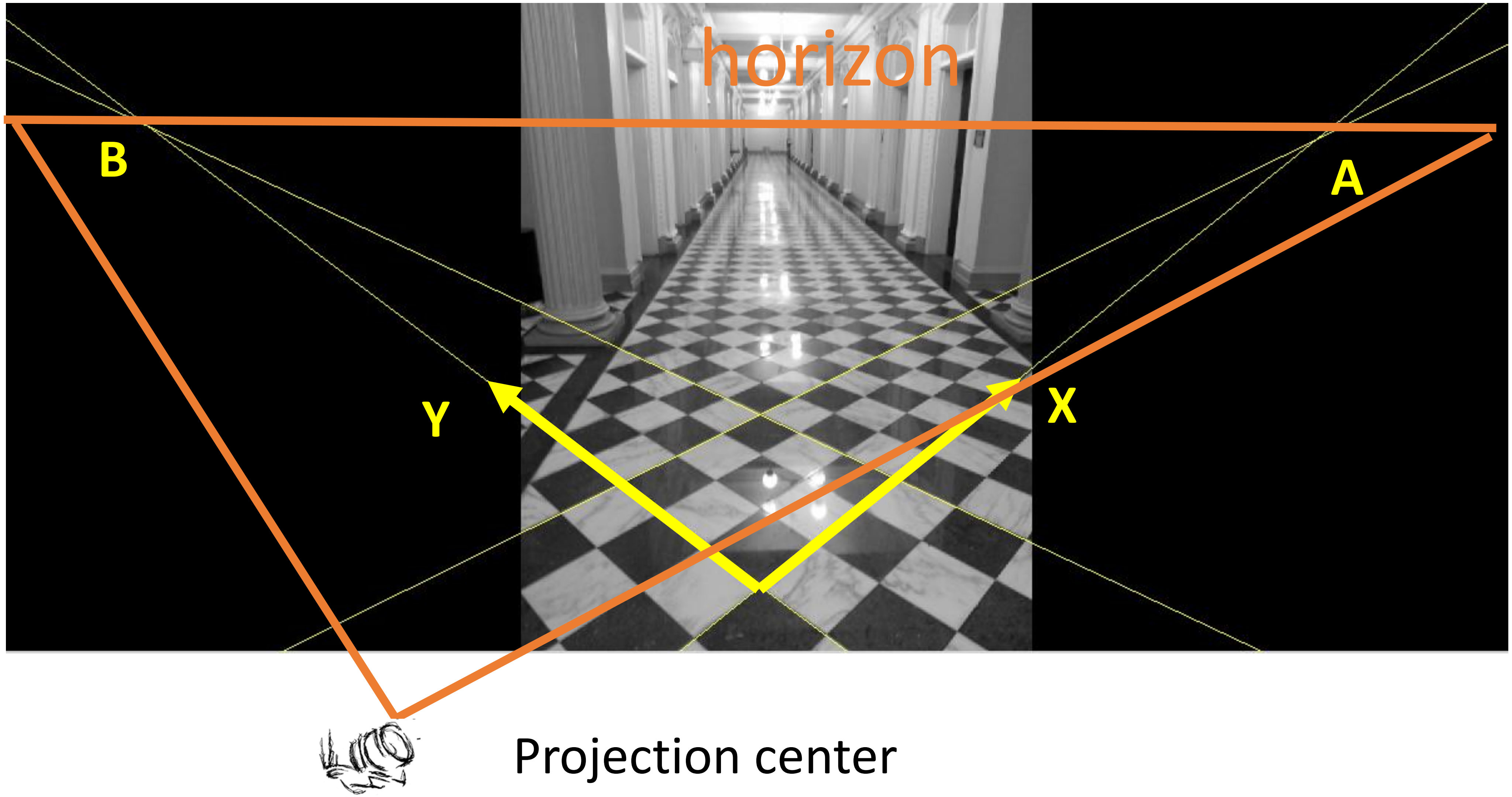


Equation of horizon: $(h_1 \times h_2)^T \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0$



Horizon with projection center build a horizon plane with normal

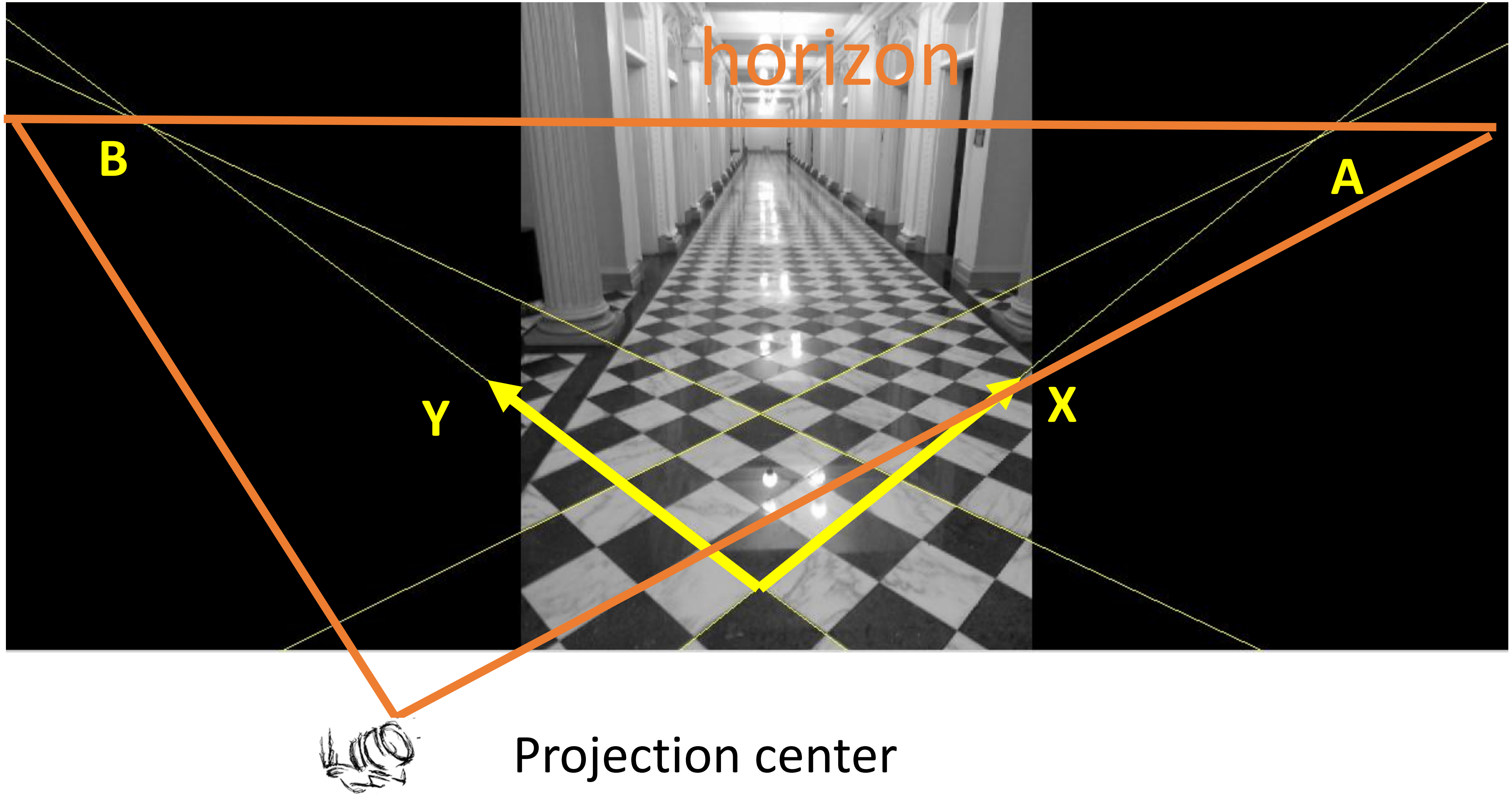
$$(h_1 \times h_2)$$



Projection center

The horizon plane is parallel to the ground plane and hence

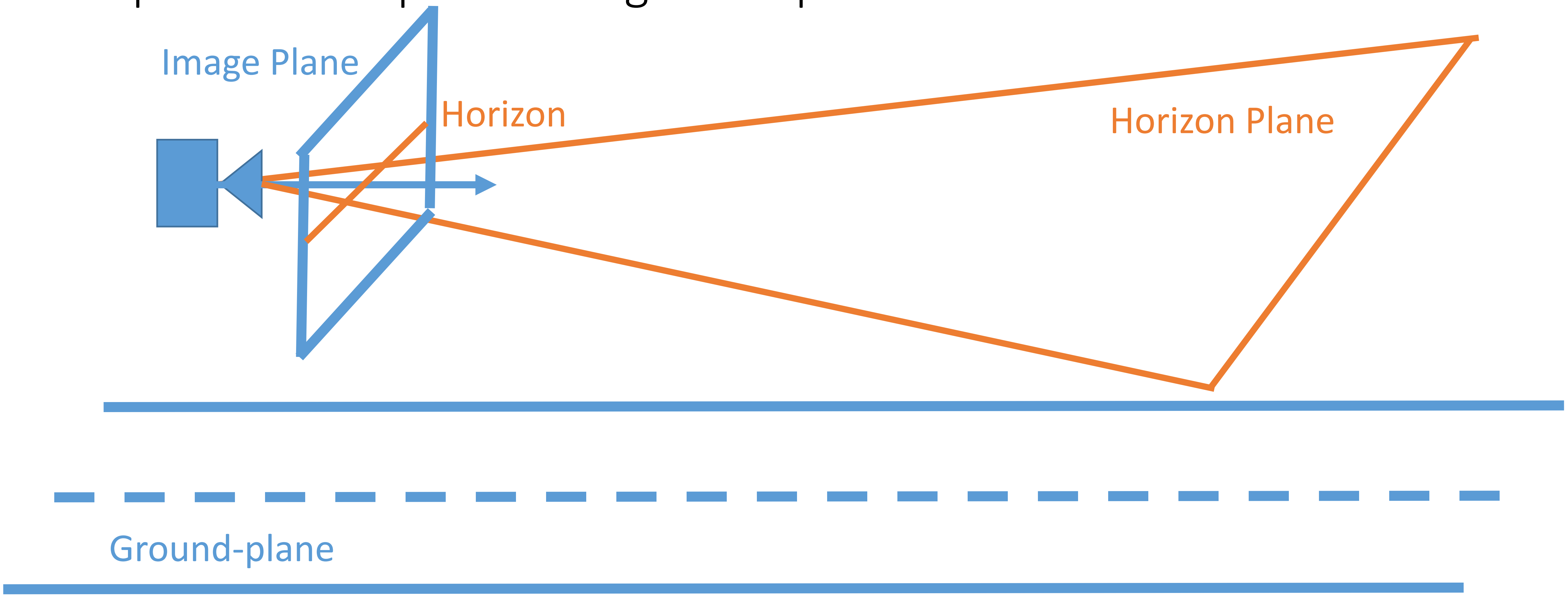
$(h_1 \times h_2)$ is the normal to the ground plane expressed via pixels!



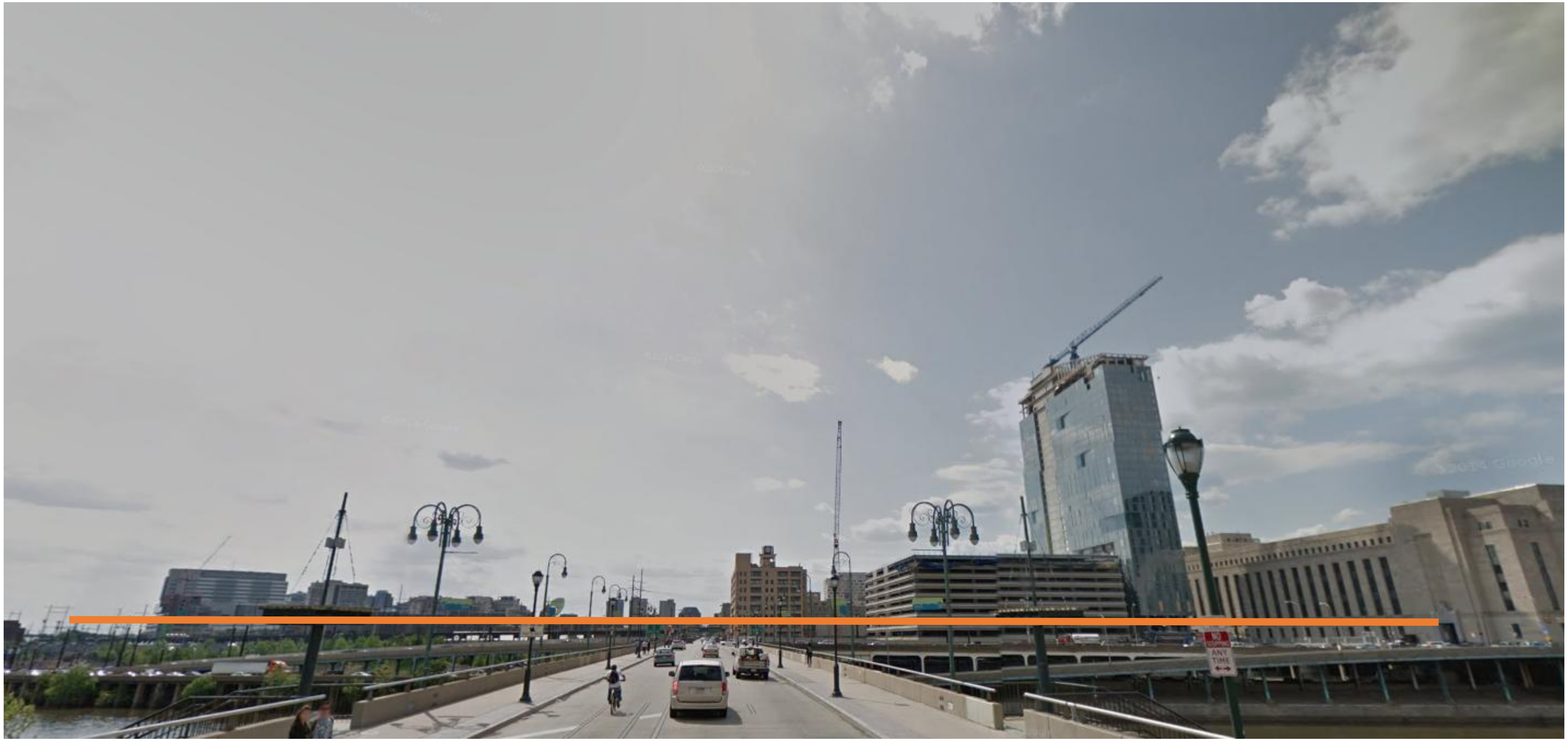
Projection center

Horizon gives complete info about how ground plane is oriented! If horizon is horizontal the center it means that optical axis is parallel to groundplane!

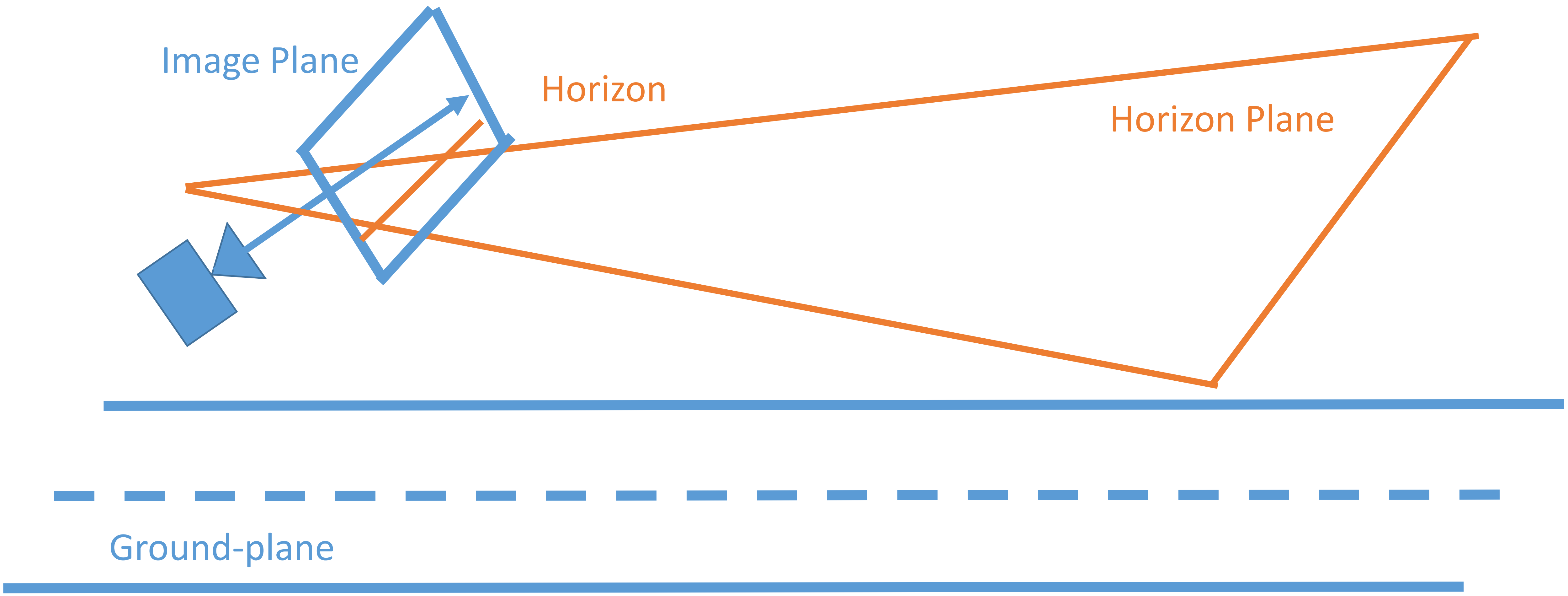
Horizon gives complete info about how ground plane is oriented! If horizon is horizontal the center it means that optical axis is parallel to ground-plane!



If horizon moves to the bottom it means we look upwards!

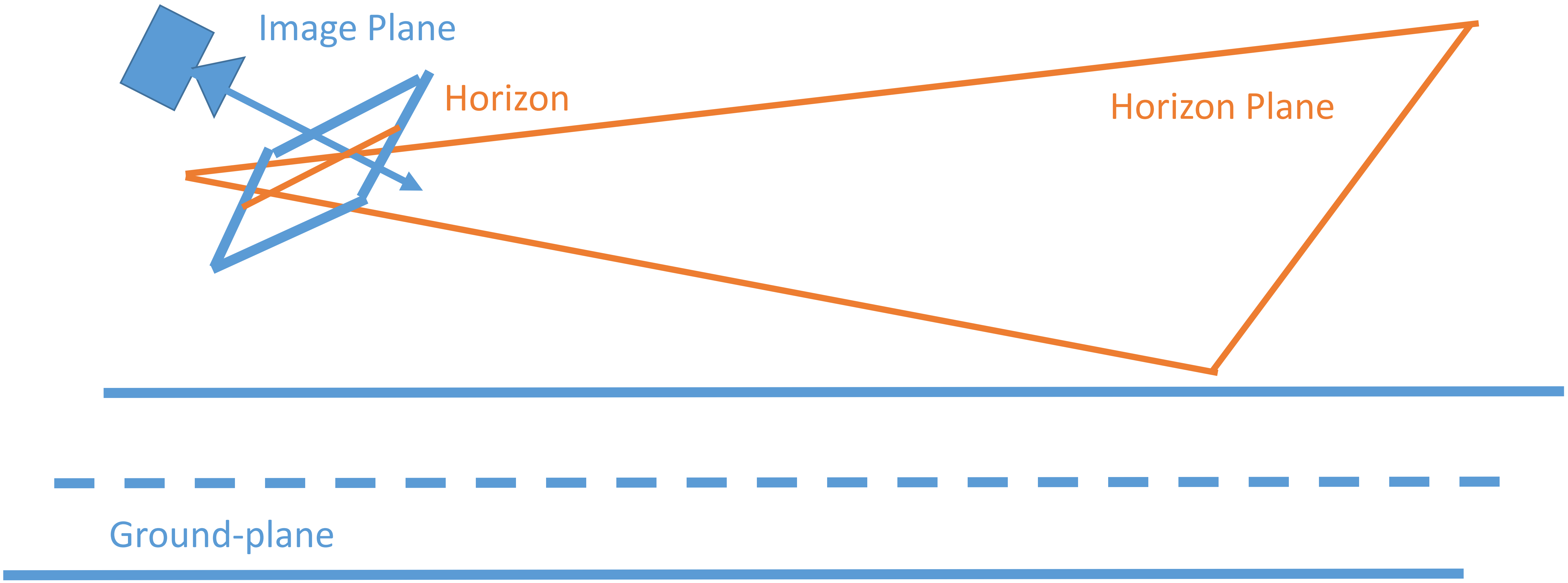


If horizon moves to the bottom it means we look upwards!

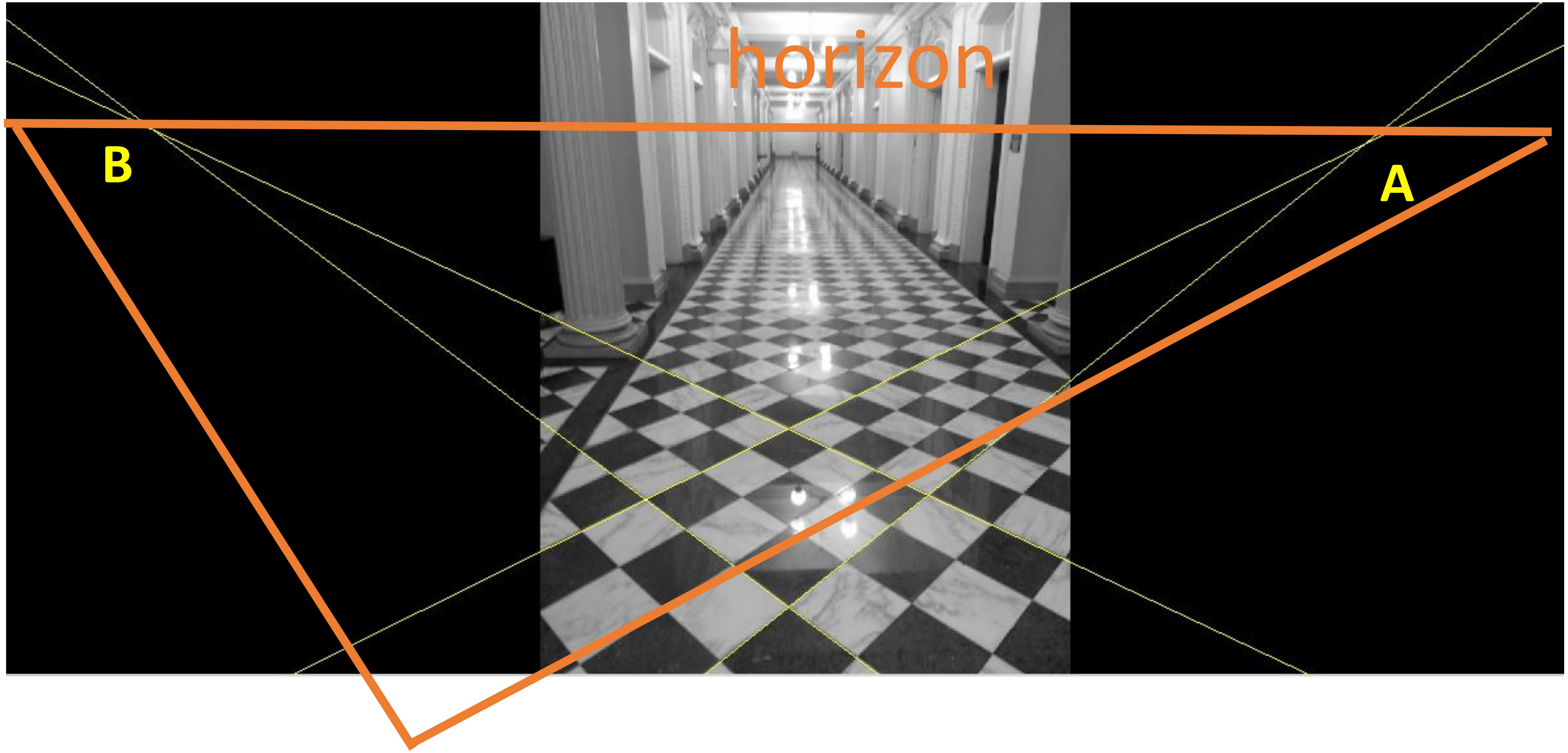


If horizon moves to the top it means we look downwards !

If horizon moves to the bottom it means we look upwards!



Horizon tells us how camera is oriented.
Constrains the collineation !

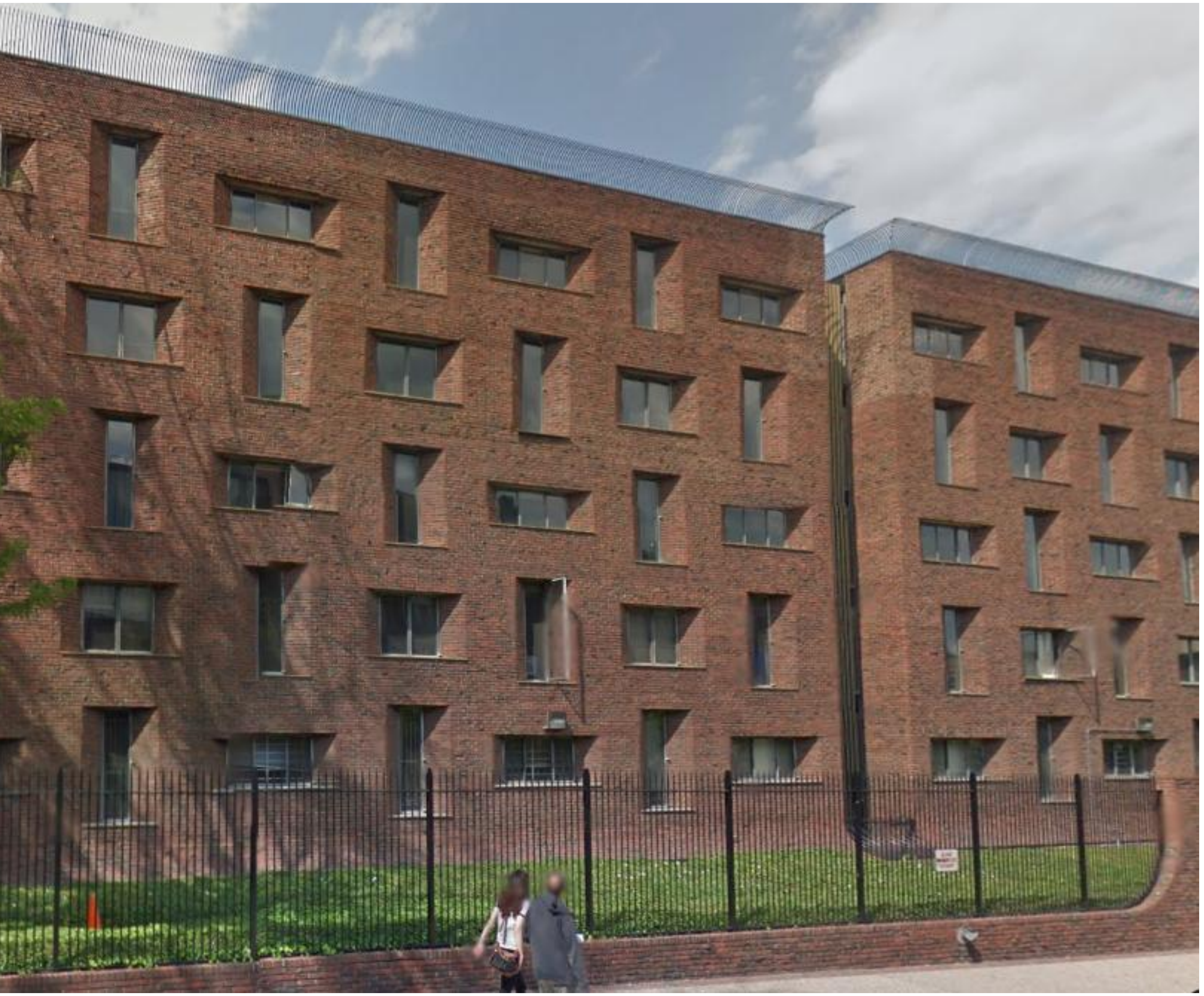


Two view metrology

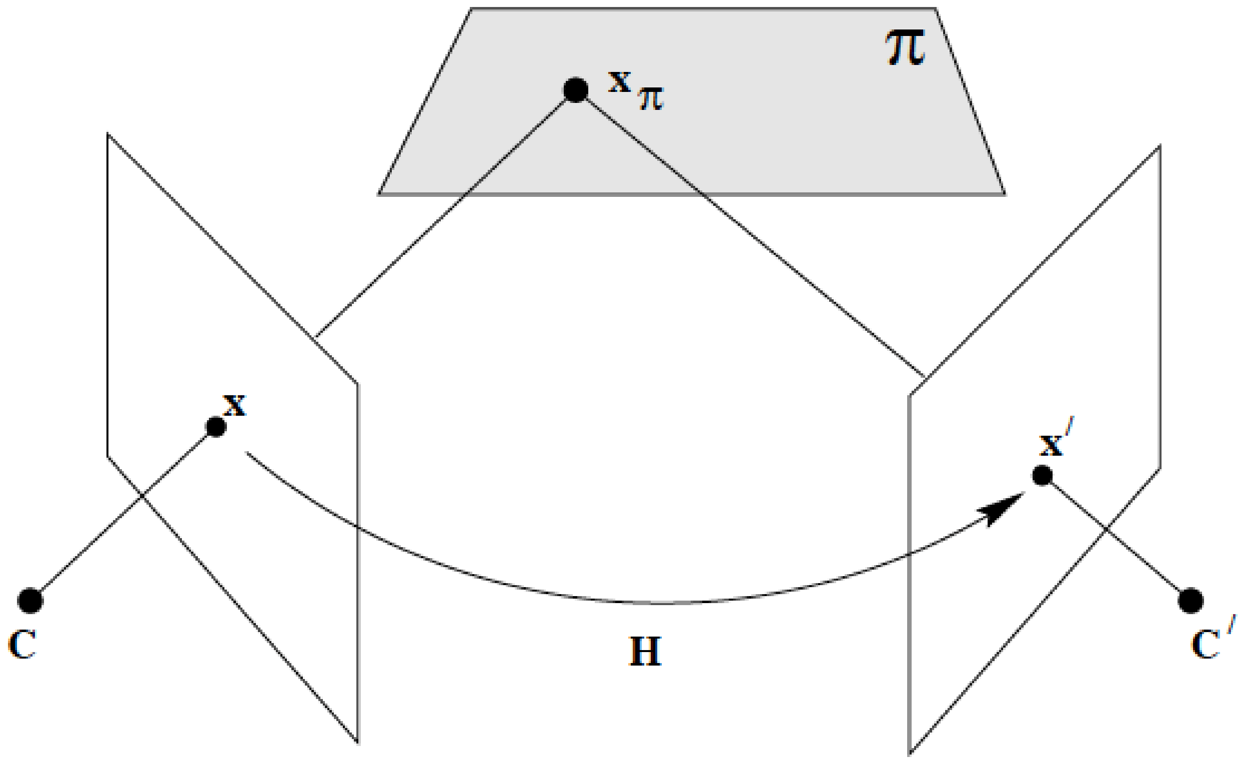
Perception: Kostas Daniilidis

We have learnt that the transformation from a world plane (like the road plane) to the image is a projective transformation !

What about two views of the same plane (façade)?



Projective transformation between two images of the same plane!



1966 World Cup: England-Germany

Was it a goal?

Goal-directed Video Metrology

Ian Reid and Andrew Zisserman

Dept of Engineering Science, University of Oxford, Oxford, OX1 3PJ

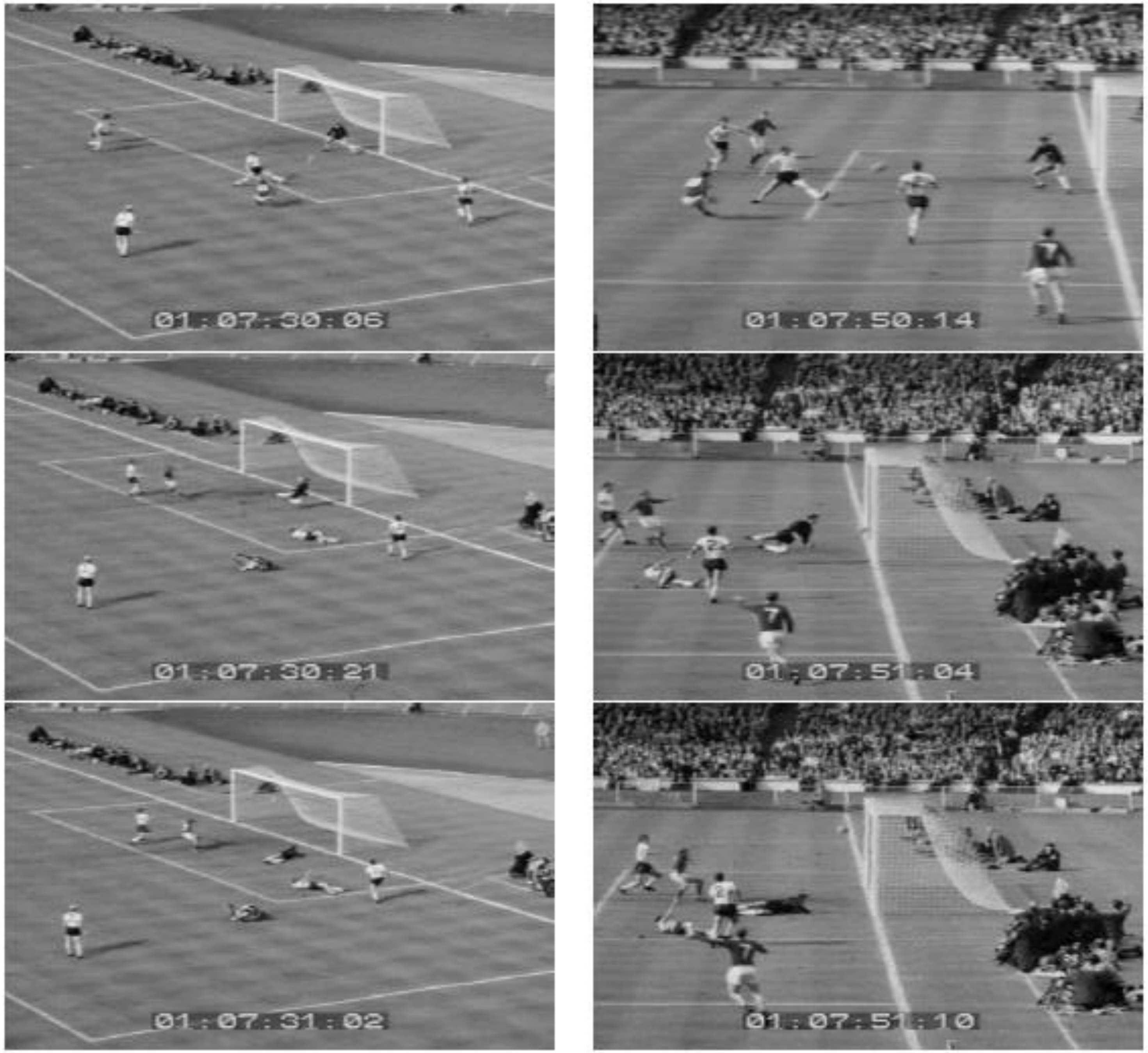
Abstract. We investigate the general problem of accurate metrology from uncalibrated video sequences where only partial information is available. We show, via a specific example – plotting the position of a goal-bound soccer ball – that accurate measurements can be obtained, and that both qualitative and quantitative questions about the data can be answered.

From two video sequences of an incident captured from different viewpoints, we compute a novel (overhead) view using pairs of corresponding images. Using projective constructs we determine the point at which the vertical line through the ball pierces the ground plane in each frame.

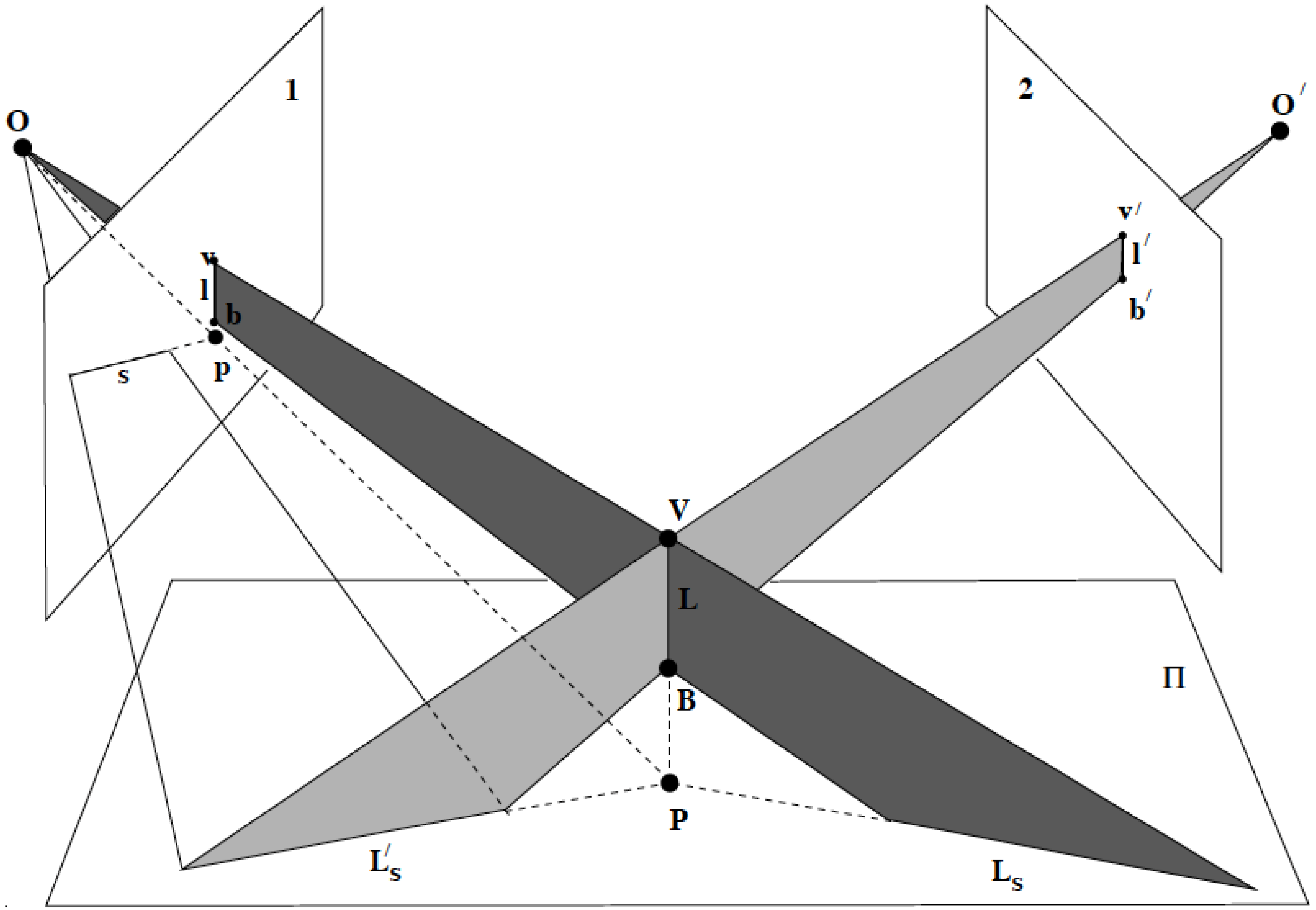
Throughout we take care to consider possible sources of error and show how these may be eliminated, neglected, or we derive appropriate uncertainty measures which are propagated via a first-order analysis.

Can we infer from two different viewpoints whether the ball was inside the goal ?

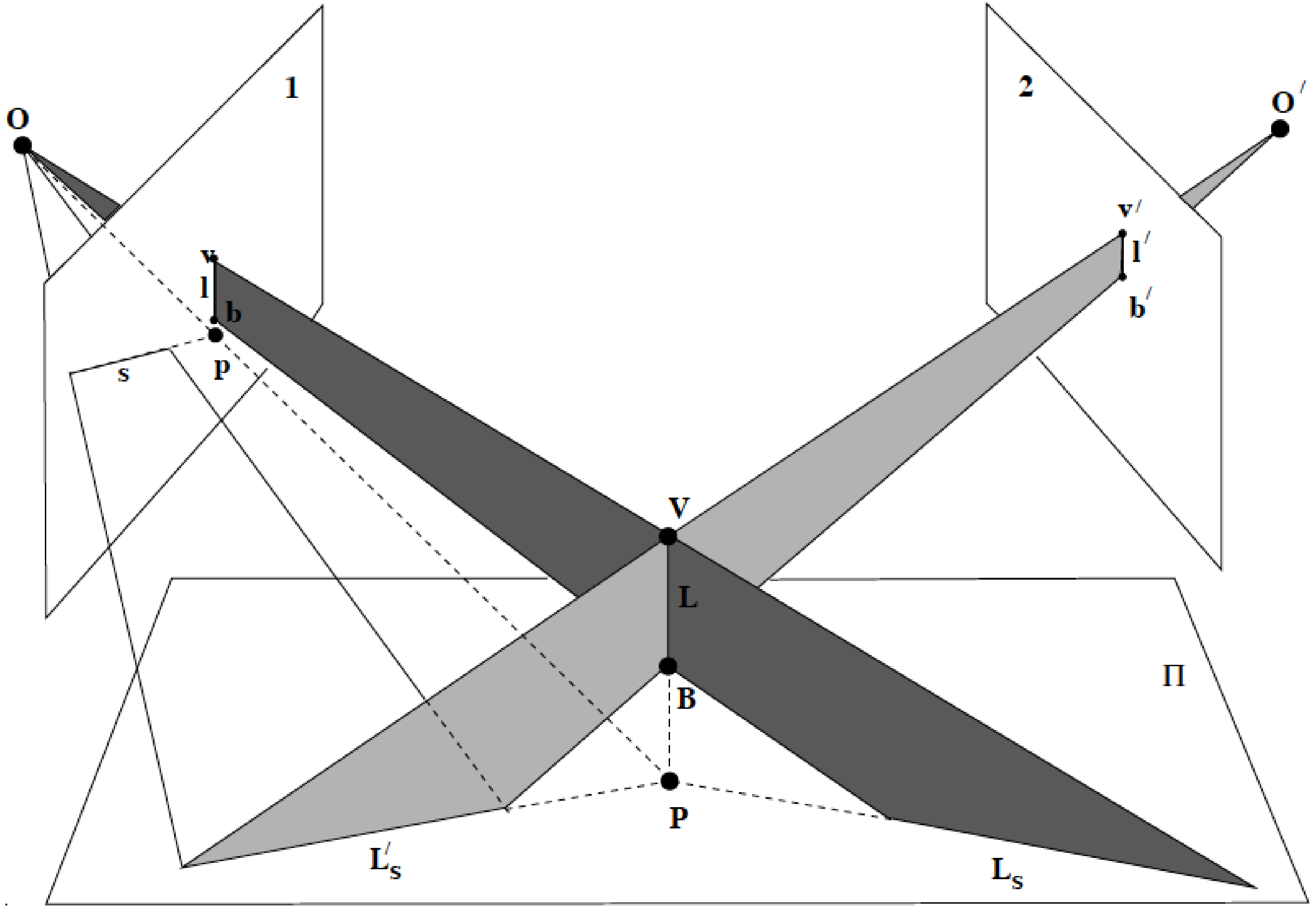
Today, this is done with what is called Goal Line Technology!



Let B be the ball ! The question is whether the vertical projection of the ball P is behind the goal-line or not!

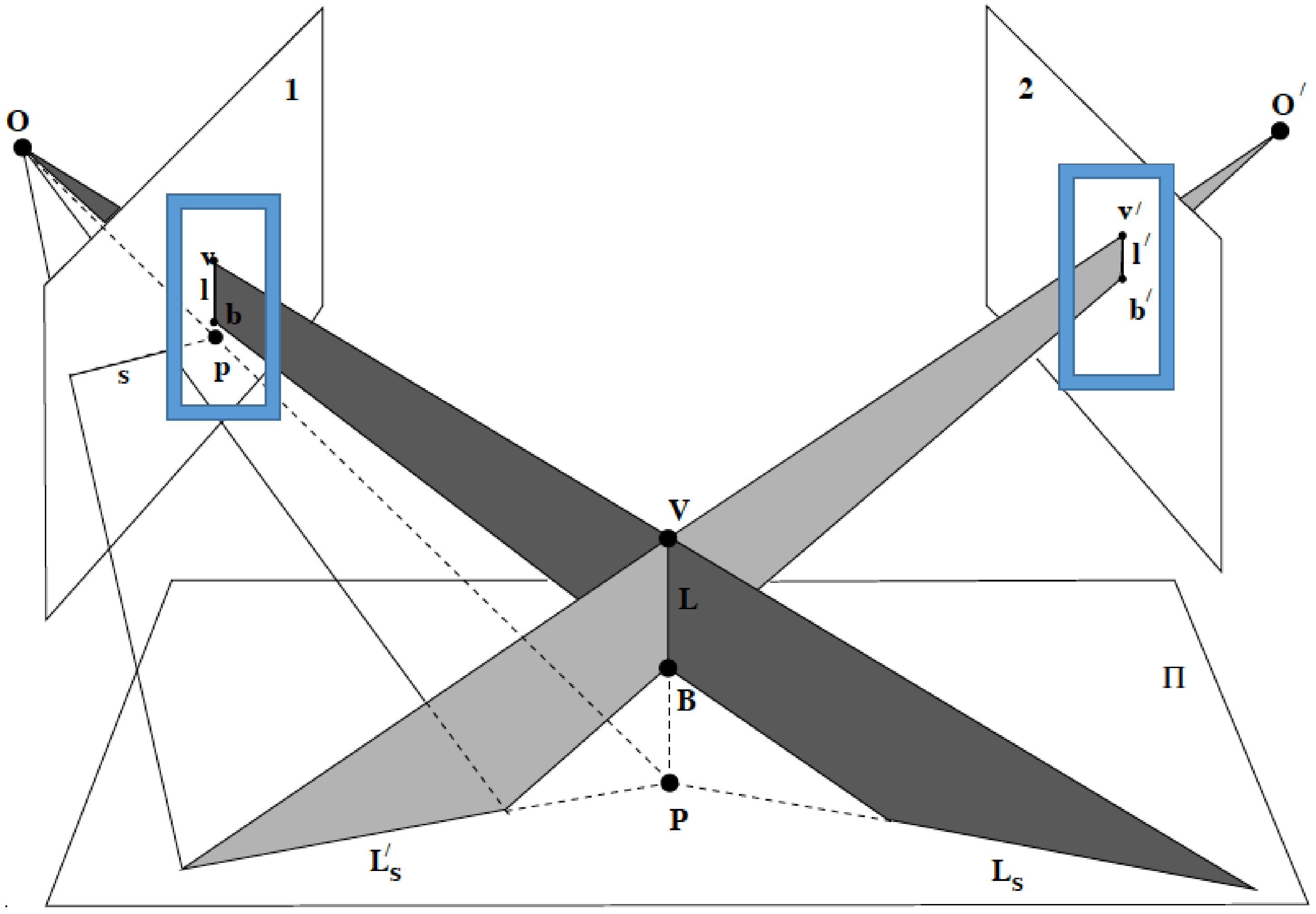


Let us also assume that we can find the vertical vanishing points v and v' in both images !

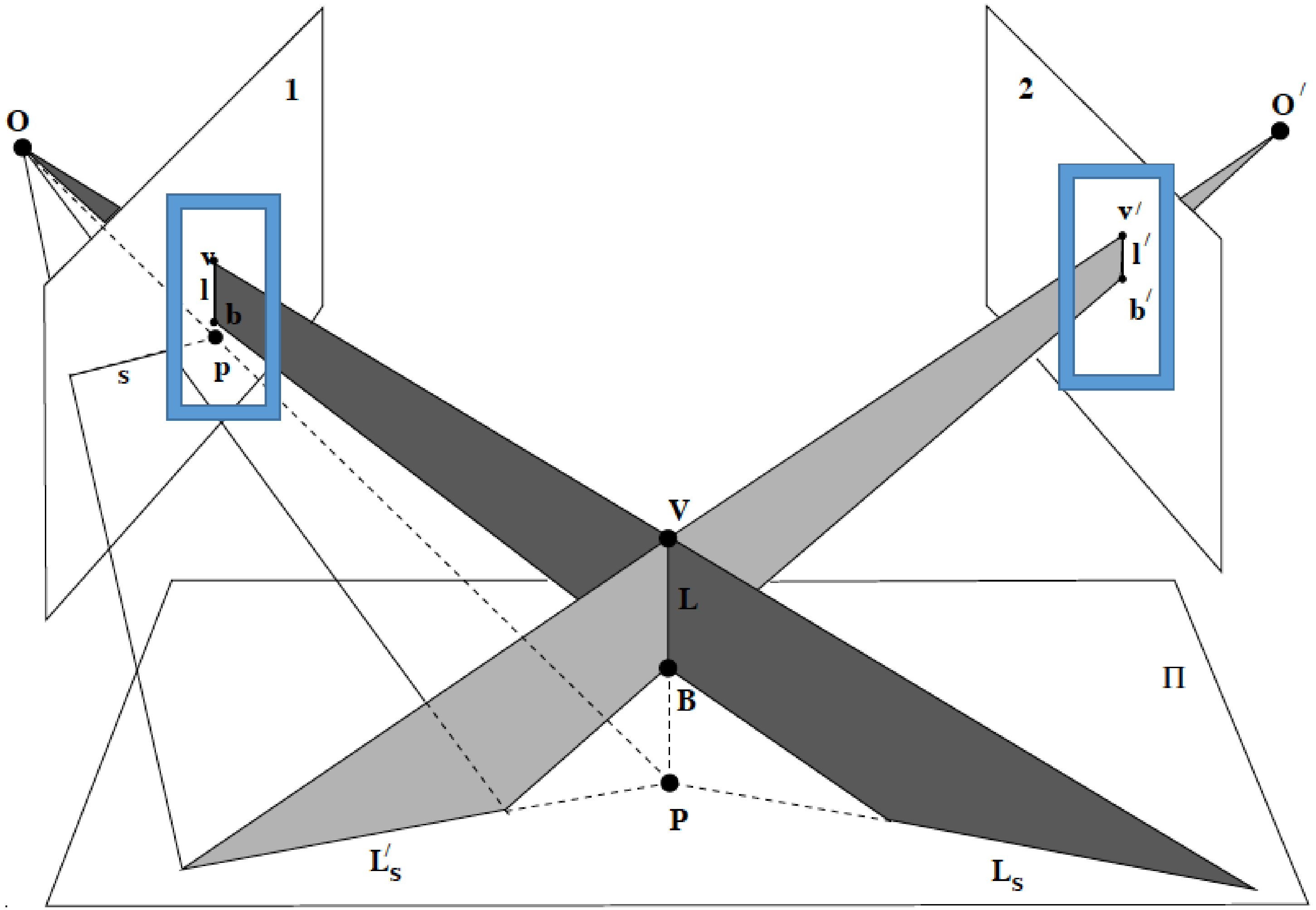


Vertical vanishing points

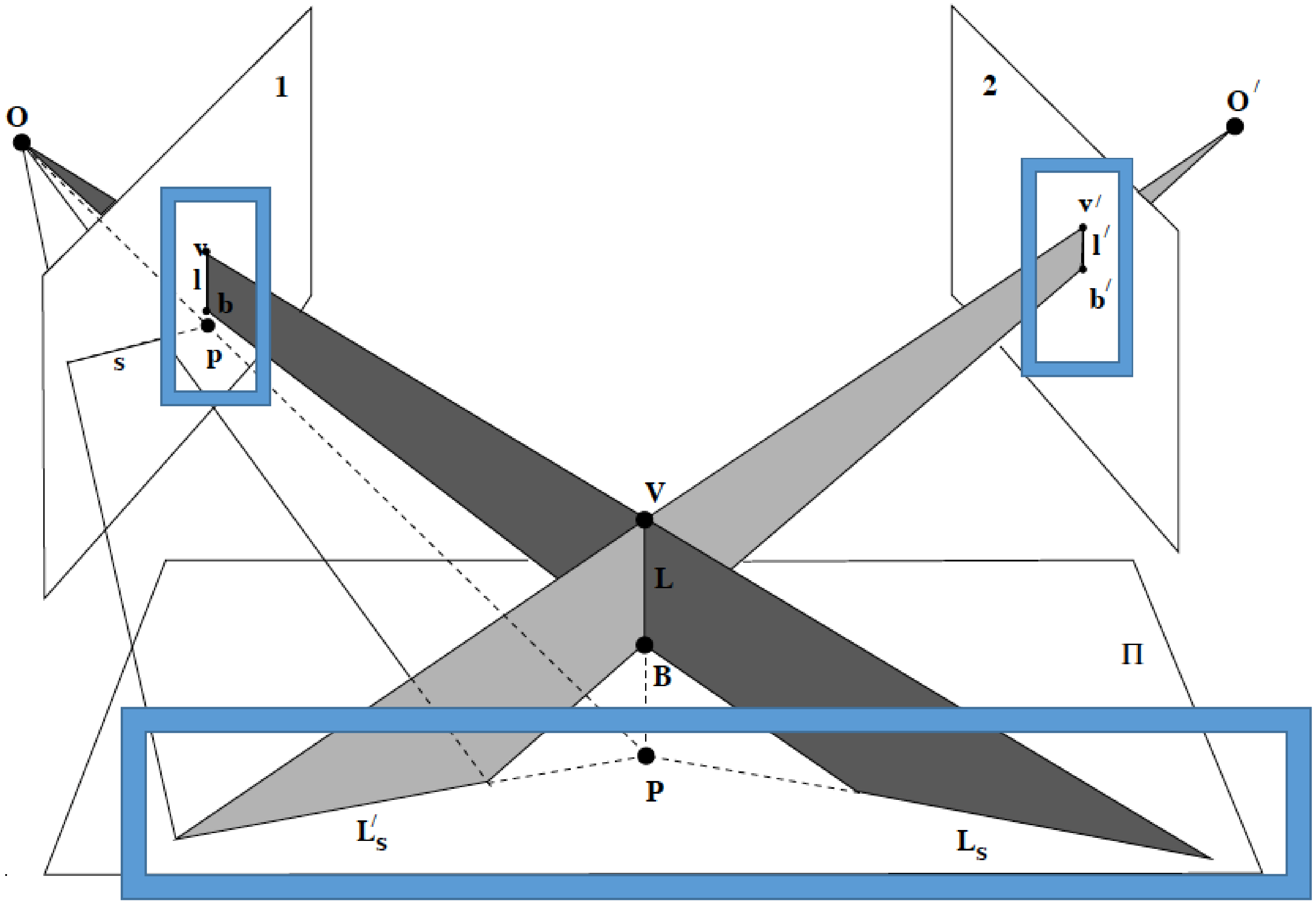
Then we know in both images the projections of vertical lines l and l'.



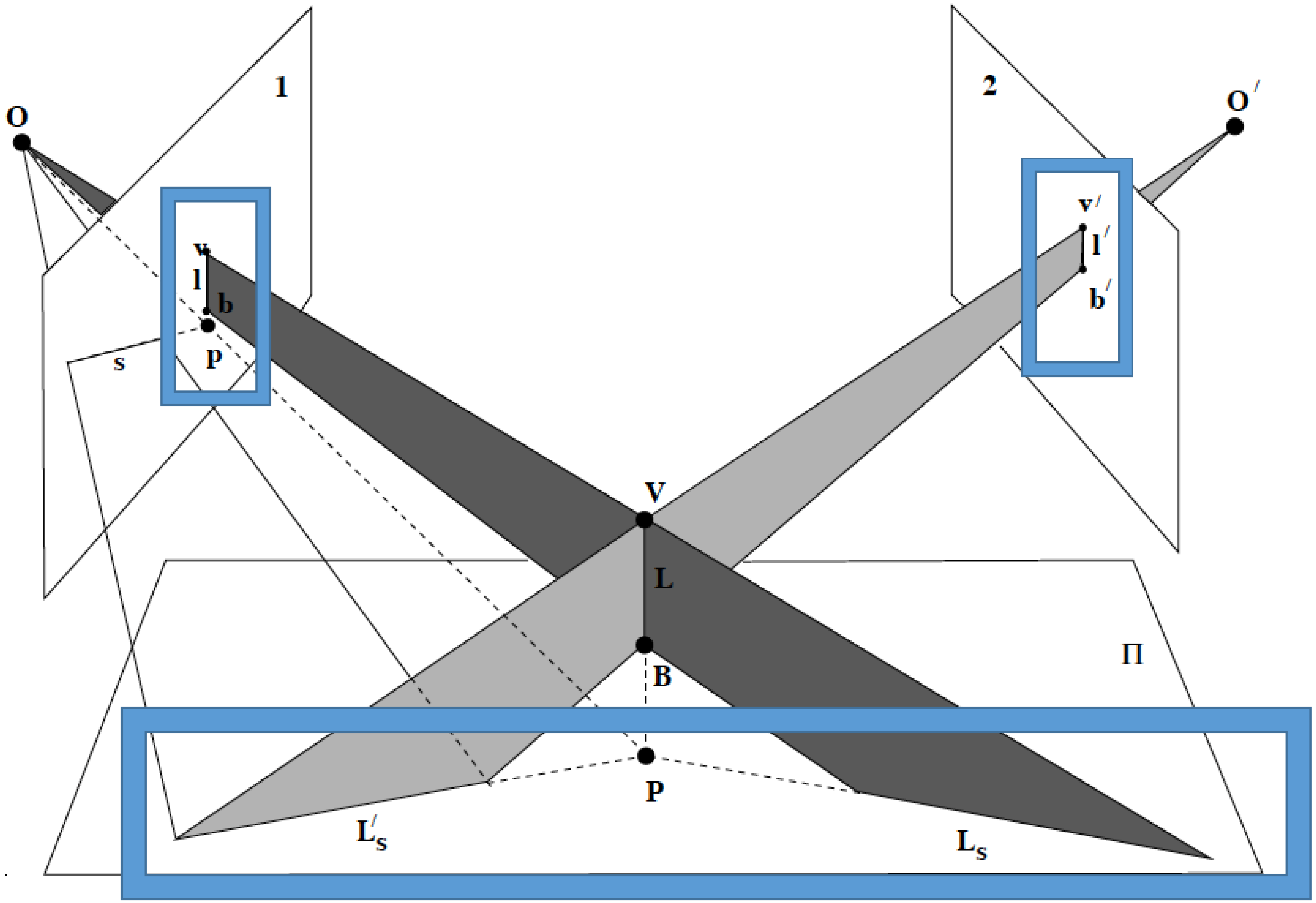
If we project them back on the soccer field we will find the shadows of the vertical through the ball.



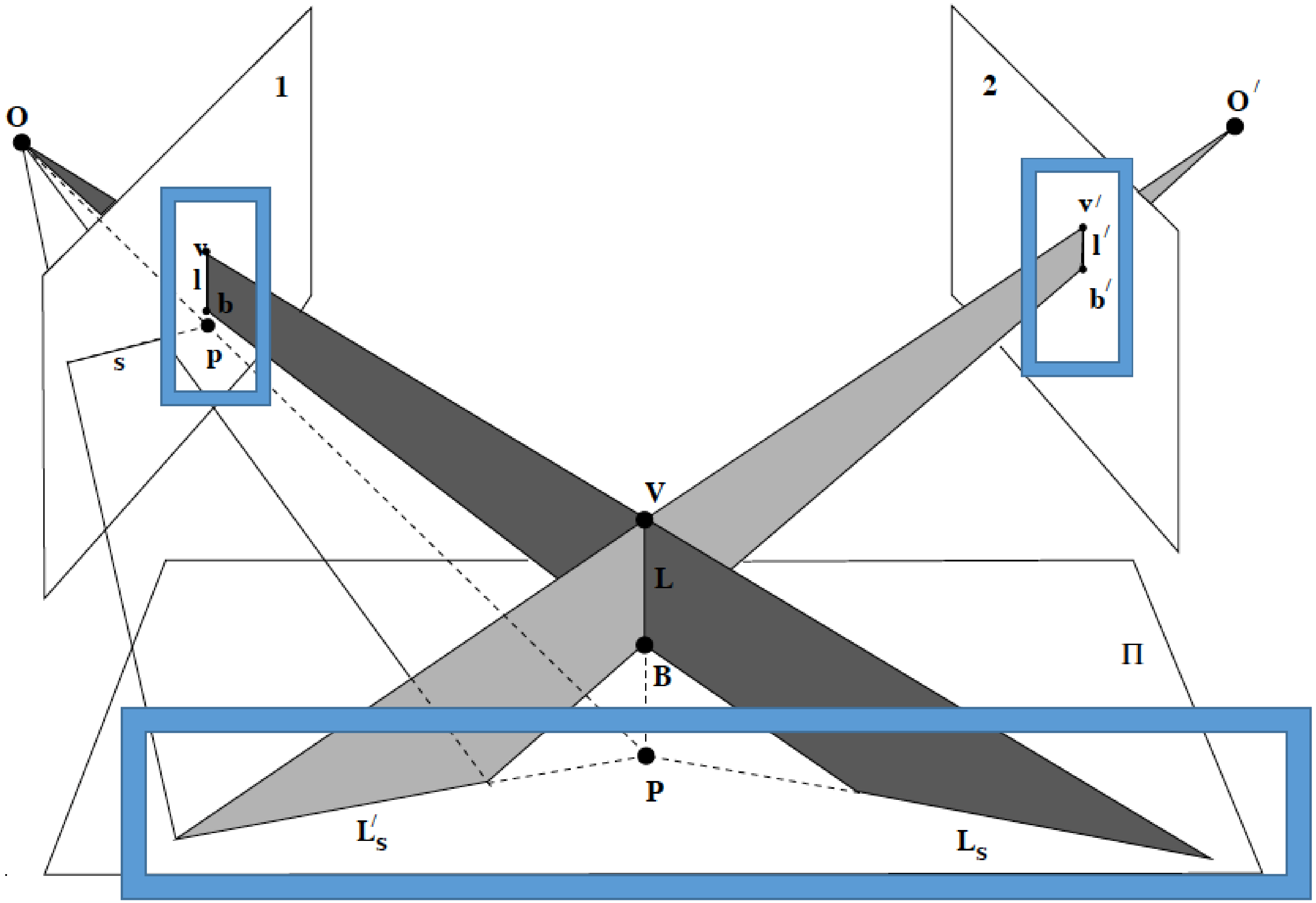
If we project them back on the soccer field we will find the shadows of the vertical through the ball.



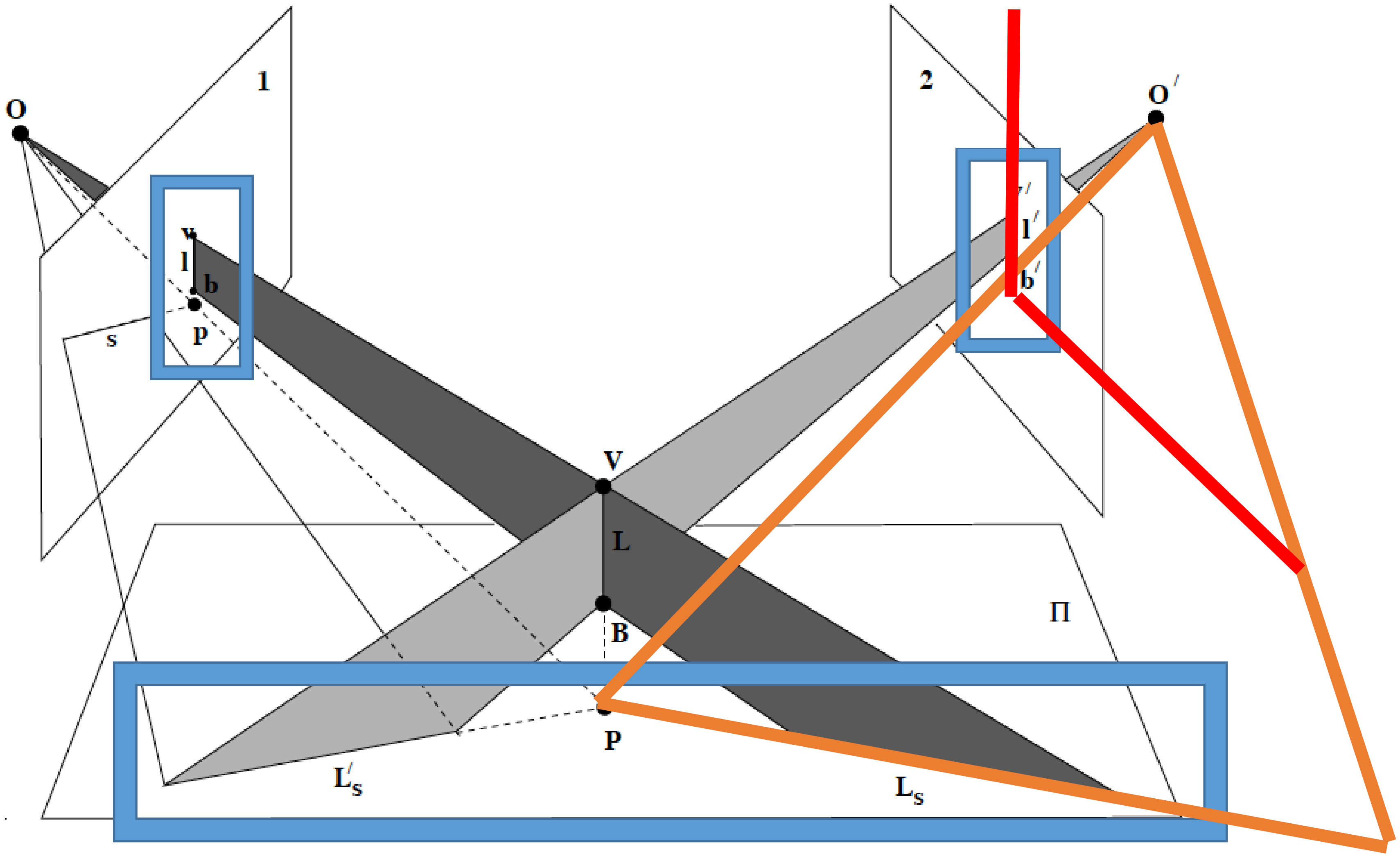
Shadows intersect at the desired point P !



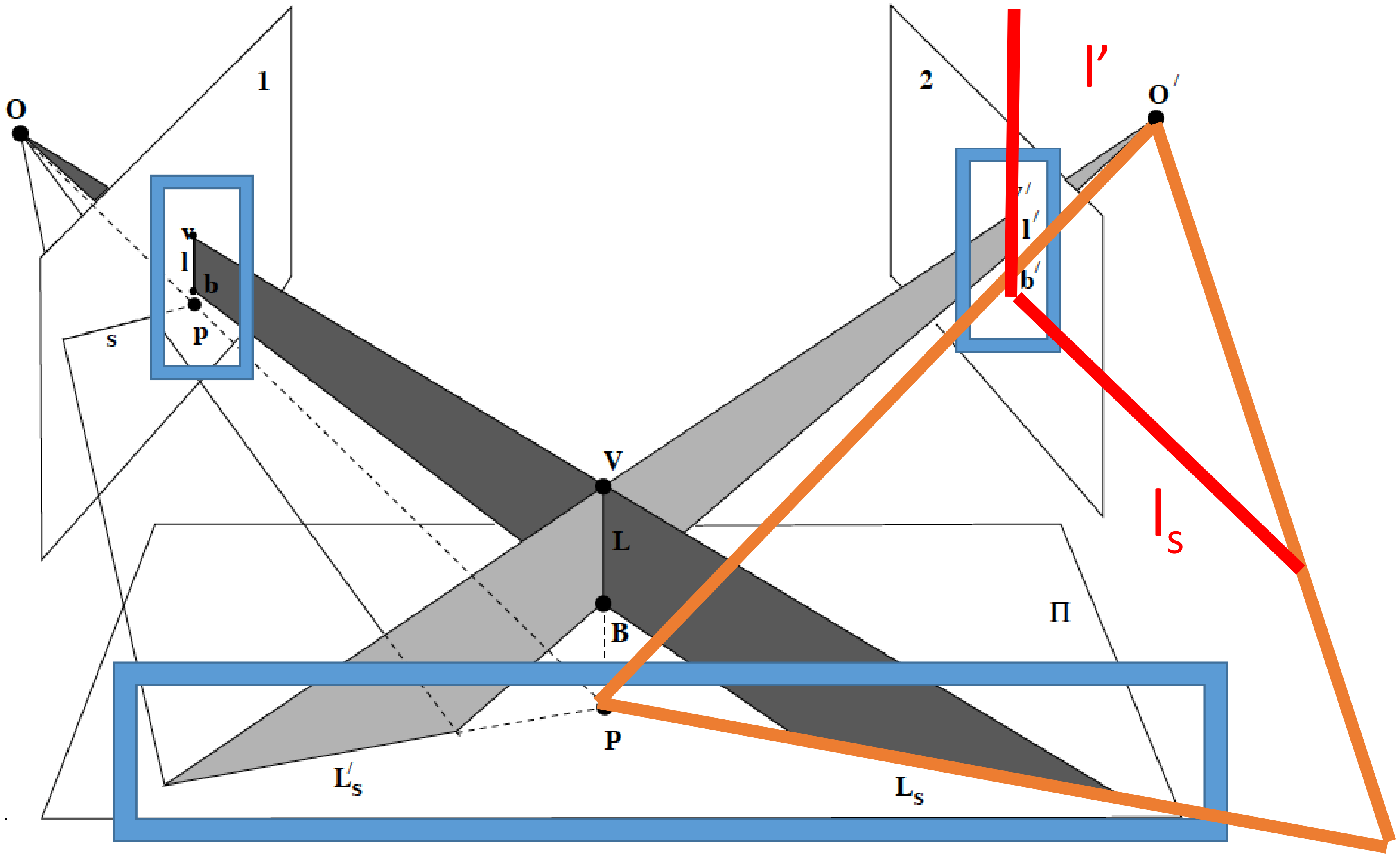
How can we find the projection of P in one of the image planes?



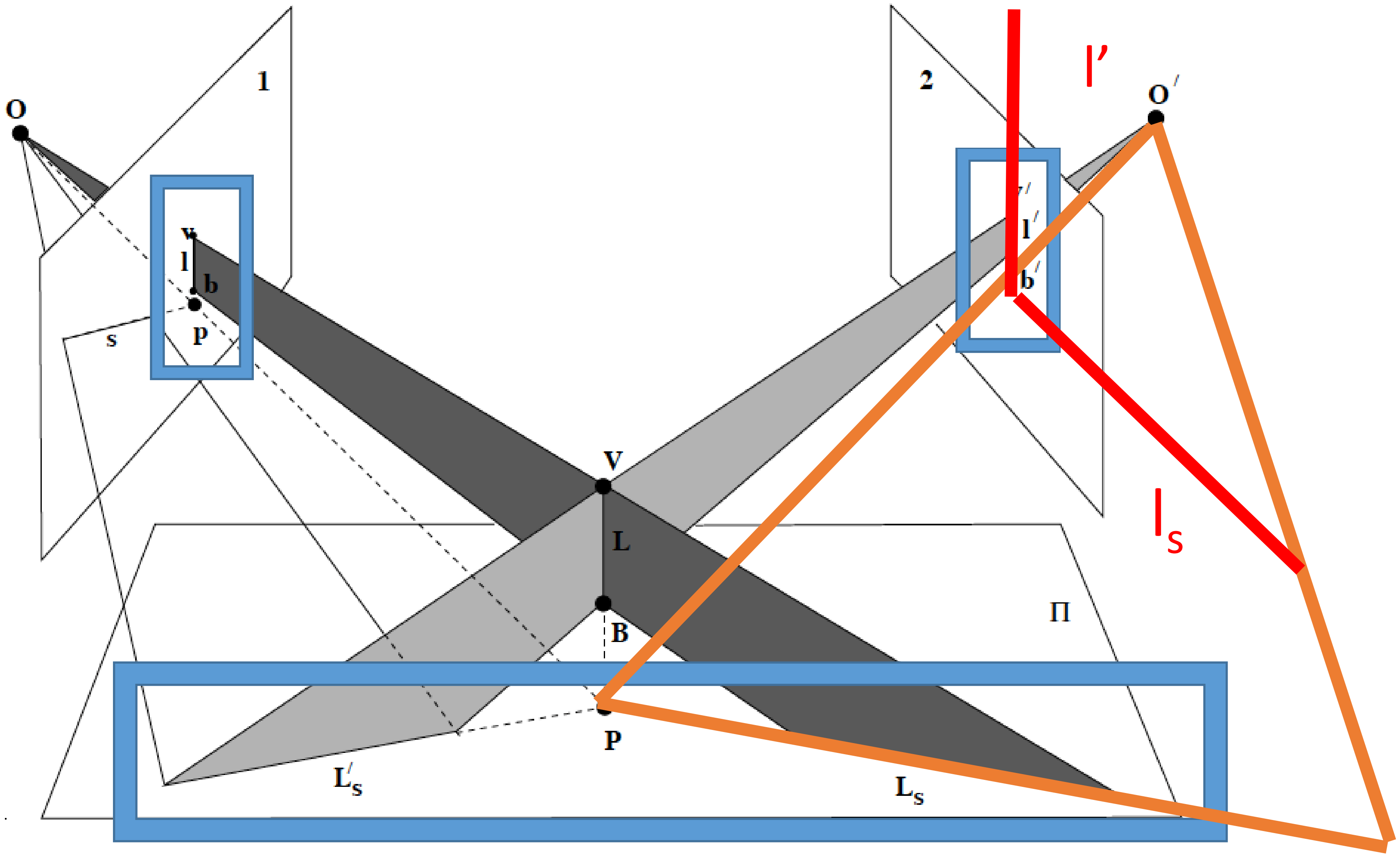
By computing the projections of the shadows! The red lines in the right image plane!



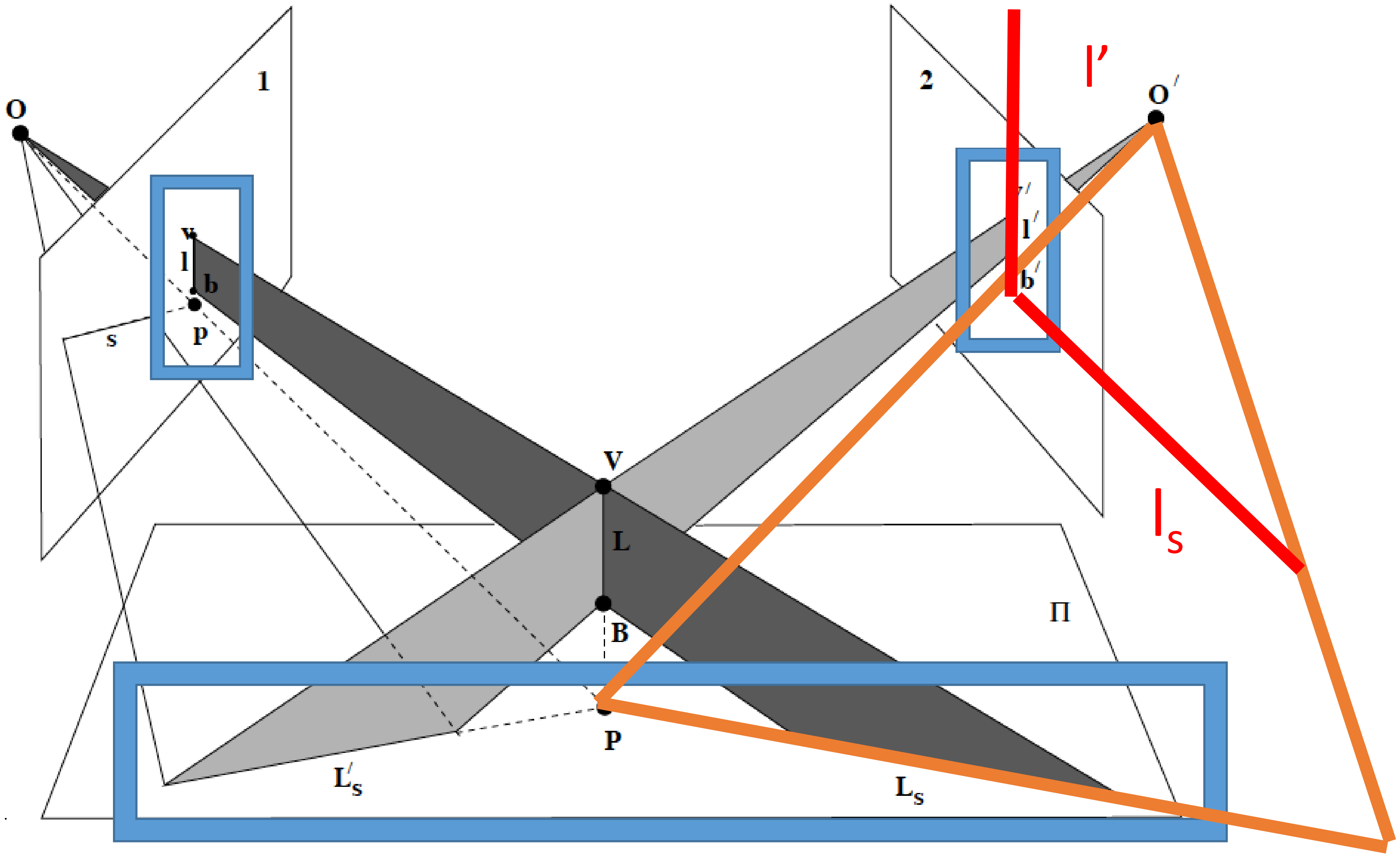
By computing the projections of the shadows l' and l_s, the red lines in the right image plane!



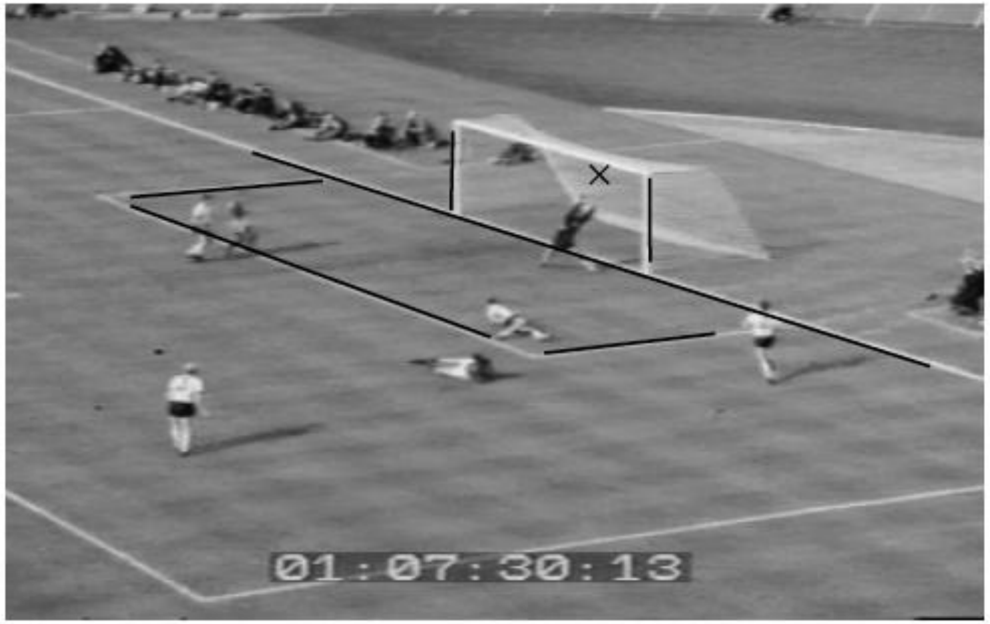
But we only know l’. But l_s and l (left) are the projections from the same shadow line on the soccer field.

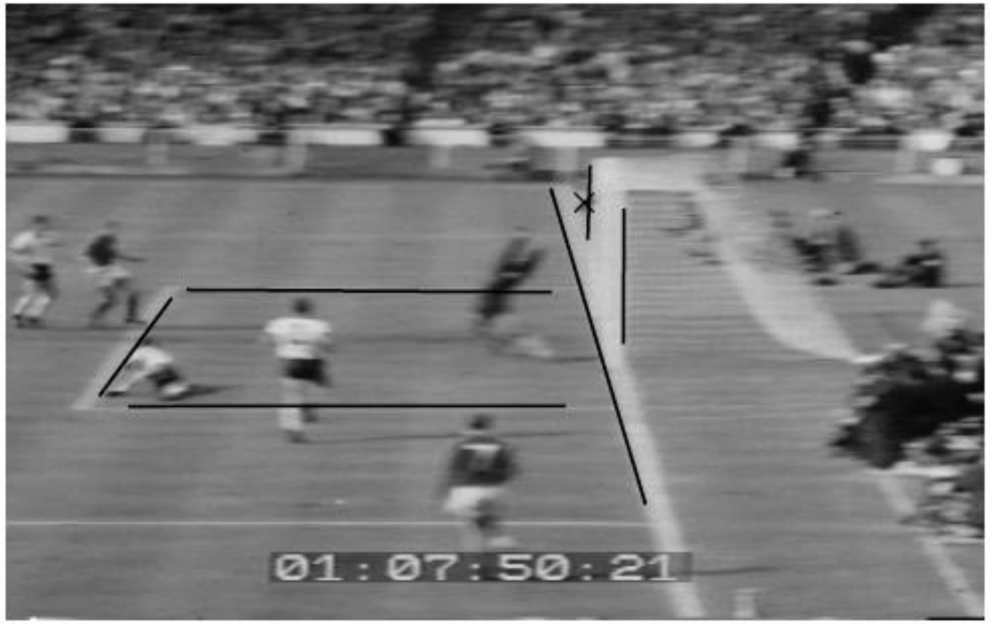


Because the soccer field is planar, they are related by a homography (collineation).

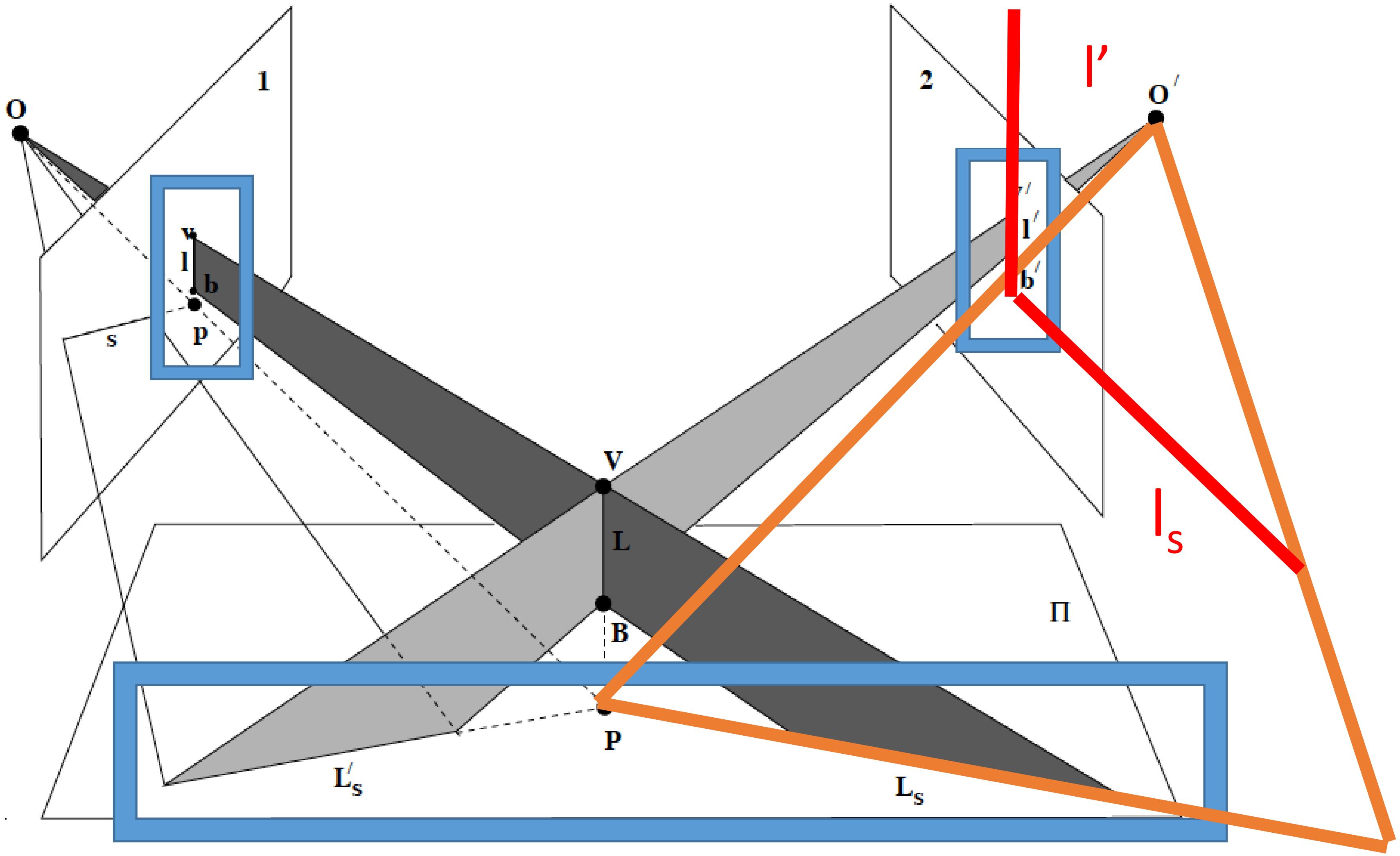


Homography of soccer field can be computed because we see the same features:



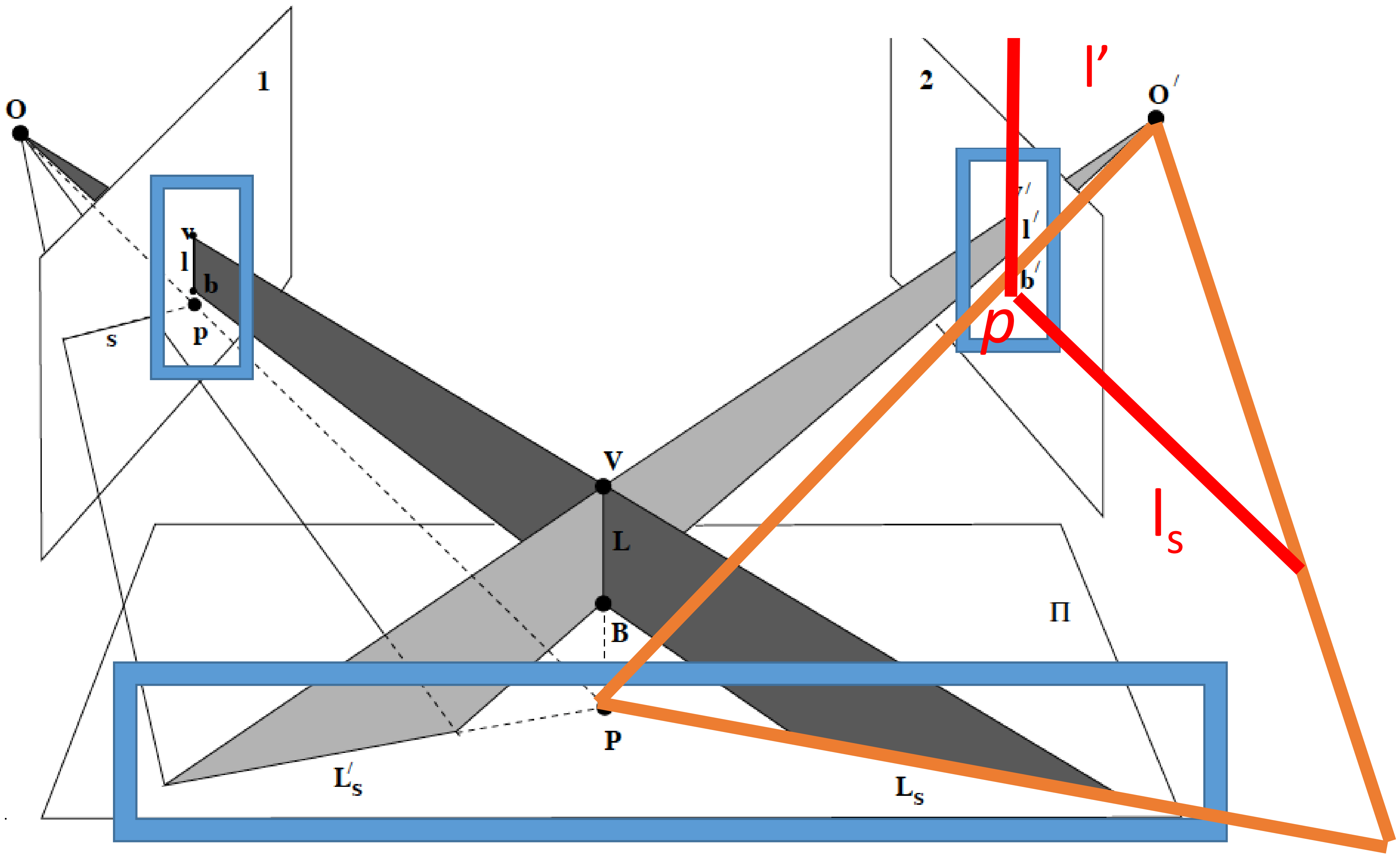


If H is the homography between points, then $l_s \sim H^{-T} l$



And the projection p is intersection of two lines

$$p \sim H^{-T} l \times l'$$

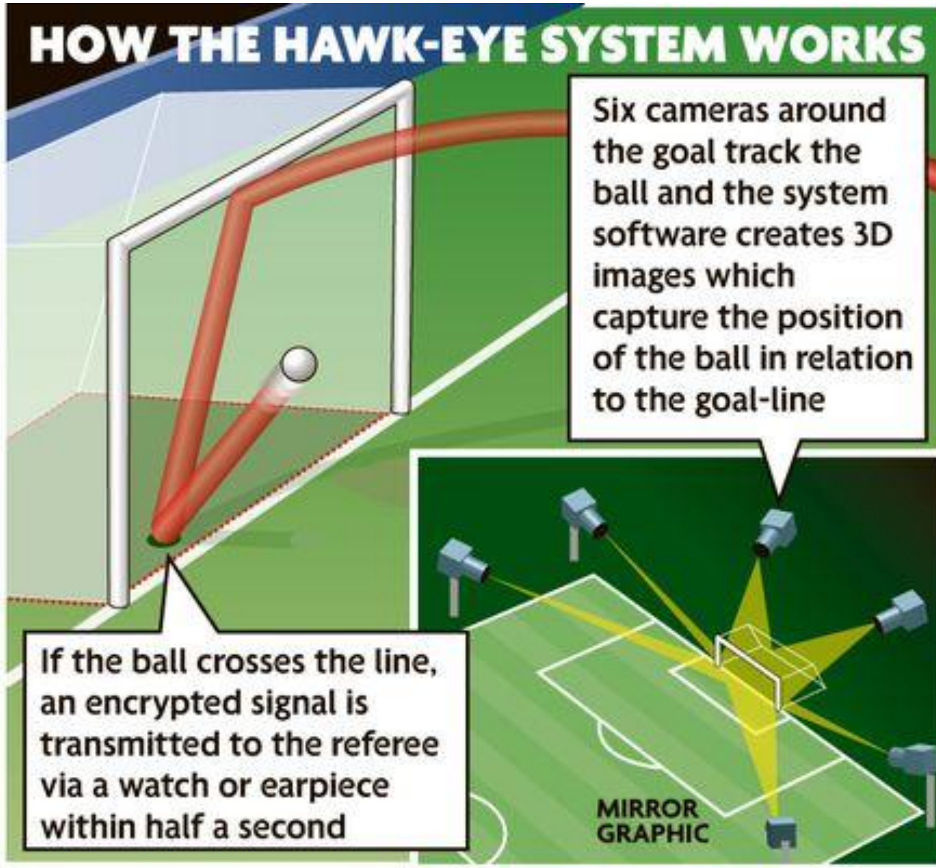


So, was it a goal ?

- No, the ball did not cross the line in full.
- A lot more engineering behind geometry:
  - Synchronizing frames cross cameras
  - Size of the ball
  - Motion blure

England – Germany 2010 !

Today, multiple cameras with very accurate relative pose tracked! The Hawk-Eye.



Thanks to Reid and Zisserman!

http://www.learnopencv.com/how-computer-vision-solved-the-greatest-soccer-mystery-of-all-times/

Goal-directed Video Metrology

Ian Reid and Andrew Zisserman

Dept of Engineering Science, University of Oxford, Oxford, OX1 3PJ

Abstract. We investigate the general problem of accurate metrology from uncalibrated video sequences where only partial information is available. We show, via a specific example – plotting the position of a goal-bound soccer ball – that accurate measurements can be obtained, and that both qualitative and quantitative questions about the data can be answered.
From two video sequences of an incident captured from different viewpoints, we compute a novel (overhead) view using pairs of corresponding images. Using projective constructs we determine the point at which the vertical line through the ball pierces the ground plane in each frame.
Throughout we take care to consider possible sources of error and show how these may be eliminated, neglected, or we derive appropriate uncertainty measures which are propagated via a first-order analysis.