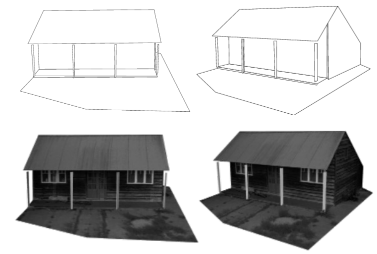
Single-view metrology

Many slides adapted from
S. Seitz, D. Hoiem

R. Magritte, *Personal Values*, 1952

1

Application: 3D from a single image

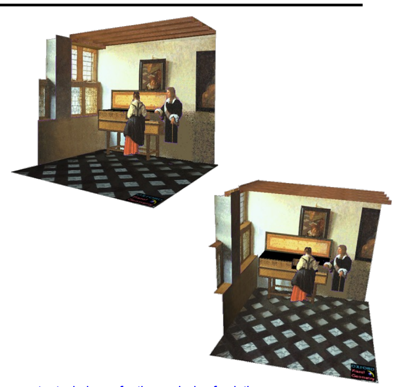


A. Criminisi et al. Single View Metrology, IJCV 2000

2

Application: 3D from a single image

J. Vermeer, Music Lesson, 1662



A. Criminisi et al. Bringing Pictorial Space to Life: computer techniques for the analysis of paintings, *Proc. Computers and the History of Art*, 2002

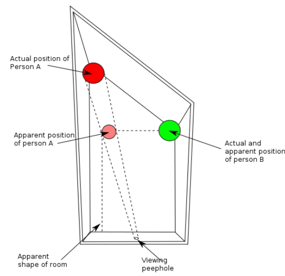
3

Application: Image editing, augmented reality

K. Karsch and V. Hedau and D. Forsyth and D. Hoiem. Rendering Synthetic Objects into Legacy Photographs, *SIGGRAPH Asia* 2011

4

Reminder: Beware!



http://en.wikipedia.org/wiki/Ames_room

5

Outline

- Camera calibration using vanishing points
- Measurements from a single image
- Applications of single-view metrology

6

Camera calibration using vanishing points

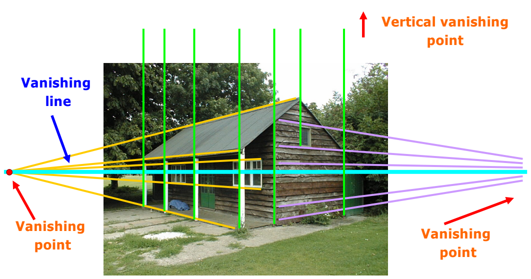
- If world coordinates of reference 3D points are not known, in special cases, we may be able to use vanishing points

Source: A. Efros, A. Criminisi

7

Camera calibration using vanishing points

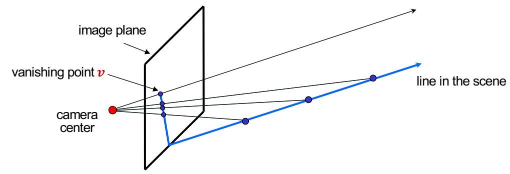
- If world coordinates of reference 3D points are not known, in special cases, we may be able to use vanishing points



Source: A. Efros, A. Criminisi

8

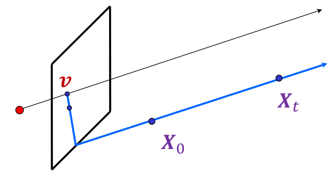
Review: Vanishing points



- All lines having the same direction share the same vanishing point

9

Computing vanishing points



- Let's parameterize the line using point $\boldsymbol{X}_0 = (X_0, Y_0, Z_0, 1)^T$ and direction vector $\boldsymbol{D} = (D_1, D_2, D_3)^T$:

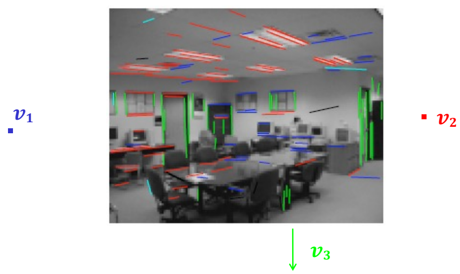
$$\boldsymbol{X}_t = \begin{pmatrix} X_0 + tD_1 \\ Y_0 + tD_2 \\ Z_0 + tD_3 \\ 1 \end{pmatrix} \cong \begin{pmatrix} X_0/t + D_1 \\ Y_0/t + D_2 \\ Z_0/t + D_3 \\ 1/t \end{pmatrix} \qquad \boldsymbol{X}_\infty = \begin{pmatrix} D_1 \\ D_2 \\ D_3 \\ 0 \end{pmatrix}$$

- $\boldsymbol{X}_\infty$ is a *point at infinity*, $\boldsymbol{v}$ is its projection: $\boldsymbol{v} \cong \boldsymbol{P}\boldsymbol{X}_\infty$

10

Calibration from vanishing points

- Consider a scene with three orthogonal vanishing directions:

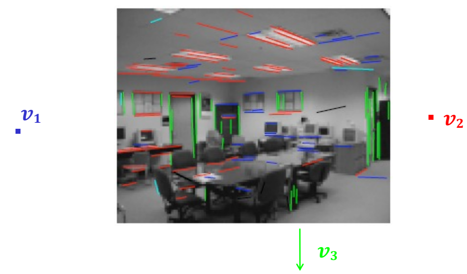


- Note: v_1, v_2 are *finite* vanishing points and v_3 is an *infinite* vanishing point

11

Calibration from vanishing points

- Consider a scene with three orthogonal vanishing directions:



- We can align the world coordinate system with these directions

12

Projection of the world coordinate system

$$\boldsymbol{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix}$$

13

Projection of the world coordinate system

$$\begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \boldsymbol{p}_1$$

$\boldsymbol{p}_1 \quad \boldsymbol{p}_2 \quad \boldsymbol{p}_3 \quad \boldsymbol{p}_4$

14

Projection of the world coordinate system

$$\begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \boldsymbol{p}_2$$

$\boldsymbol{p}_1 \quad \boldsymbol{p}_2 \quad \boldsymbol{p}_3 \quad \boldsymbol{p}_4$

15

Projection of the world coordinate system

$$\begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \boldsymbol{p}_3$$

$\boldsymbol{p}_1 \quad \boldsymbol{p}_2 \quad \boldsymbol{p}_3 \quad \boldsymbol{p}_4$

$\underbrace{\boldsymbol{p}_1 \quad \boldsymbol{p}_2 \quad \boldsymbol{p}_3}$: Vanishing points in x, y, z directions (i.e., $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\boldsymbol{v}_3$)

16

Projection of the world coordinate system

$$\begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \boldsymbol{p}_4$$

$\boldsymbol{p}_1 \quad \boldsymbol{p}_2 \quad \boldsymbol{p}_3 \quad \boldsymbol{p}_4$

Vanishing points in x, y, z directions (i.e., $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\boldsymbol{v}_3$)

Projection of world origin

- Problem: this only gives us the four columns up to independent scale factors, additional constraints needed to solve for them

17

Calibration from vanishing points

- Let us align the world coordinate system with three orthogonal vanishing directions in the scene:

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \boldsymbol{v}_i \cong \boldsymbol{K}[\boldsymbol{R}|\boldsymbol{t}] \begin{pmatrix} \boldsymbol{e}_i \\ 0 \end{pmatrix}$$

18

Calibration from vanishing points

- Let us align the world coordinate system with three orthogonal vanishing directions in the scene:

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \boldsymbol{v}_i \cong \boldsymbol{K}\boldsymbol{R}\boldsymbol{e}_i$$

$$\boldsymbol{e}_i \cong \boldsymbol{R}^T \boldsymbol{K}^{-1} \boldsymbol{v}_i$$

- Orthogonality constraint: $\boldsymbol{e}_i^T \boldsymbol{e}_j = 0$

$$\underbrace{\boldsymbol{v}_i^T \boldsymbol{K}^{-T} \boldsymbol{R}}_{\boldsymbol{e}_i^T} \underbrace{\boldsymbol{R}^T \boldsymbol{K}^{-1} \boldsymbol{v}_j}_{\boldsymbol{e}_j} = 0$$

19

Calibration from vanishing points

- Let us align the world coordinate system with three orthogonal vanishing directions in the scene:

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad \boldsymbol{v}_i \cong \boldsymbol{K}\boldsymbol{R}\boldsymbol{e}_i$$

$$\boldsymbol{e}_i \cong \boldsymbol{R}^T \boldsymbol{K}^{-1} \boldsymbol{v}_i$$

- Orthogonality constraint: $\boldsymbol{e}_i^T \boldsymbol{e}_j = 0$

$$\boldsymbol{v}_i^T \boldsymbol{K}^{-T} \boldsymbol{K}^{-1} \boldsymbol{v}_j = 0$$

- Extrinsic parameter matrix ($\boldsymbol{R}$) disappears and we are left with constraints on the calibration matrix!

20

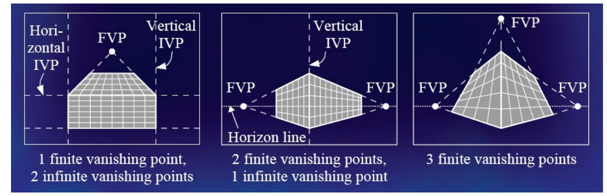
Calibration from vanishing points

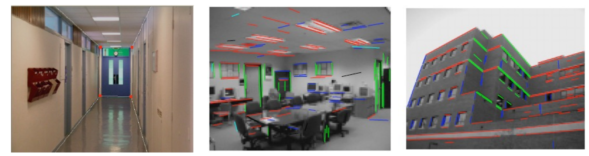
$$v_i^T K^{-T} K^{-1} v_j = 0$$

- How many constraints do we get?
  - Three: one for each pair of vanishing points
- How many unknown parameters does K have?
  - Three: f, p_x, p_y
- A couple of complications:
  - The constraints are nonlinear, but it's not hard to do the algebra (omitted)
  - At least two *finite* vanishing points are needed to solve for both focal length and principal point

21

Calibration from vanishing points





Cannot recover focal length, principal point is the third vanishing point

Can solve for focal length, principal point

22

Rotation from vanishing points

- Constraints on vanishing points: $v_i \cong K R e_i$
- We just used orthogonality constraints to solve for K
- Now we have:

$$K^{-1} v_i \cong R e_i$$

Notice: $R e_1 = [r_1 \quad r_2 \quad r_3] \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = r_1$

Thus, $r_i \cong K^{-1} v_i$

- The scale ambiguity goes away since we require $\|r_i\|^2 = 1$

23

Calibration from vanishing points: Summary

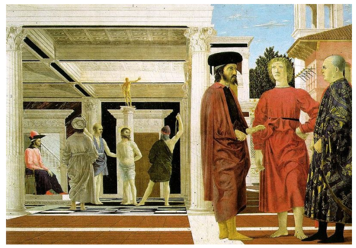
1. Solve for intrinsic parameters (focal length, principal point) using three orthogonal vanishing points
2. Get extrinsic parameters (rotation) directly from vanishing points once calibration matrix is known

- Advantages
  - No need for calibration chart, 2D-3D correspondences
  - Could be completely automatic
- Disadvantages
  - Only applies to certain kinds of scenes
  - It is tricky to accurately localize vanishing points
  - Need at least two finite vanishing points

24

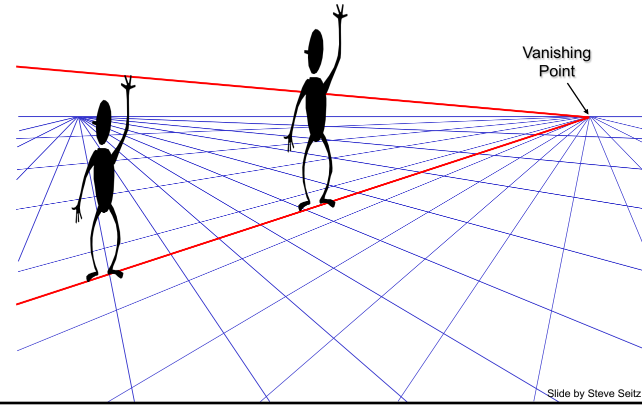
Outline

- Camera calibration using vanishing points
- Measurements from a single image
  - Measuring height above the ground plane
  - Measuring within planes



25

Comparing heights



26

Using a ruler

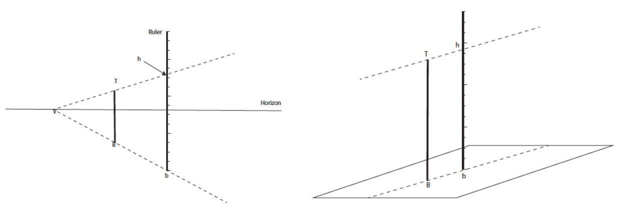


FIGURE 24.3: **Left**, *an image of a ruler and an object, which just happen to be standing perpendicular to a ground plane. In an* uncalibrated *image like this, we can measure the height of the object. Construct the line bB, and intersect that with the horizon to get the point V. The line from the top of the object T to the true height of the object on the ruler (h) is parallel in 3D to bB. In turn, the line Th must intersect the horizon at V. So if you construct VT, it will intersect the ruler at h yielding the height of the object.* **Right** *shows a 3D view; the line Th must be parallel to bB, and so in the image these two lines intersect at the horizon.*

27

Using a ruler

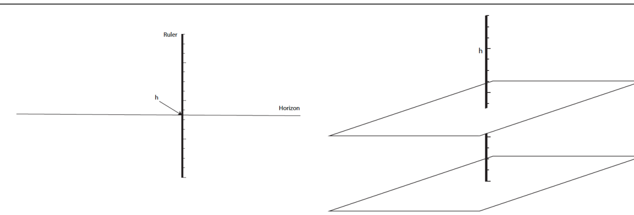


FIGURE 24.4: **Left**, *an image of a ruler which just happens to be standing perpendicular to a ground plane. In an* uncalibrated *image like this, we can measure the height of the camera focal point above the ground plane. The plane through the focal point parallel to the ground plan (and so the same height above the ground plane as the focal point) must form the horizon, so the intersection between horizon and ruler yields the height of the focal point.* **Right** *shows a 3D view; the bottom plane is the ground plane, and the top plane is the plane through the focal point parallel to the ground plane.*

28

Working without a ruler is harder than might seem

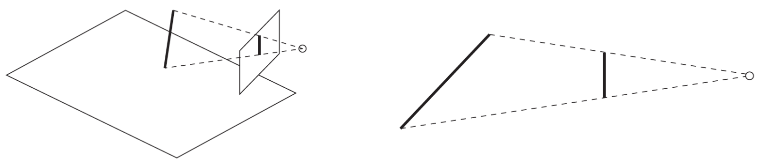


FIGURE 24.5: **Left**, *a perspective camera views a reference object perpendicular to a ground plane. This produces a line segment in the image plane.* **Right** *shows the reference object and the line segment in the image plane.*

29

Working without a ruler is harder than might seem

Parametrize the reference line segment in 3D

using affine coordinates to get $\mathbf{p} + t\mathbf{d}$, where $\mathbf{d}$ is a unit vector (so a step of 1 in t is a step of length 1 along the reference segment). Write c_{ij} for the i, j'th component of the 3×4 camera matrix. Then the homogeneous coordinates for the image line will be

$$\begin{pmatrix} (c_{11}p_1 + c_{12}p_2 + c_{13}p_3 + c_{14}) + t(c_{11}d_1 + c_{12}d_2 + c_{13}d_3 + c_{14}) \\ (c_{21}p_1 + c_{22}p_2 + c_{23}p_3 + c_{24}) + t(c_{21}d_1 + c_{22}d_2 + c_{23}d_3 + c_{24}) \\ (c_{31}p_1 + c_{32}p_2 + c_{33}p_3 + c_{34}) + t(c_{31}d_1 + c_{32}d_2 + c_{33}d_3 + c_{34}) \end{pmatrix} = \begin{pmatrix} a + bt \\ c + dt \\ e + ft \end{pmatrix}.$$

SInce we know the image is a line, we can ignore one of these three homogeneous coordinates, so the transformation is a projective transformation. Now on the 3D

30

Working without a ruler is harder than might seem

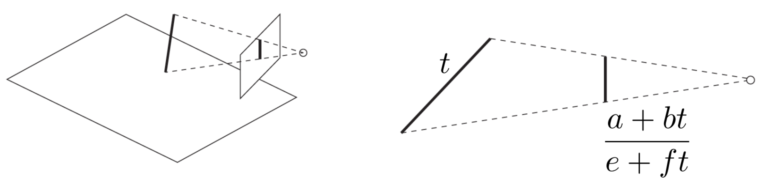


FIGURE 24.5: **Left**, *a perspective camera views a reference object perpendicular to a ground plane. This produces a line segment in the image plane.* **Right** *shows the reference object and the line segment in the image plane.*

31

Working without a ruler is harder than might seem

Now on the 3D reference line segment, the points $t = 0$ and $t = 1$ are the same distance apart as the points $t = 1$ and $t = 2$. But in the image line, using affine coordinates, these points are

$$\frac{a}{c}, \frac{a+b}{c+d}, \frac{a+2b}{c+2d}$$

which are not, in general, evenly spaced (check this with, for example, $a = 0$, $b = 1$, $c = 1$, $d = 1$).

32

The Cross-ratio

A clever trick from projective geometry allows us to use a reference object to measure heights. Write $\mathbf{P}_1, \ldots, \mathbf{P}_4$ for the coordinates of four points on a projective line, written in homogeneous coordinates. Write $\mathcal{M}$ for a projective transformation of the line to itself (so a 2×2 matrix with non-zero determinant. Finally, write

$$d(\mathbf{P}_i, \mathbf{P}_j) = \det\left([\mathbf{P}_i \mathbf{P}_j]\right).$$

Notice that

$$\det\left([\mathcal{M}\mathbf{P}_i \mathcal{M}\mathbf{P}_j)]\right) = \det\left(\mathcal{M}[\mathbf{P}_i \mathbf{P}_j]\right) = \det\left(\mathcal{M}\right)\det\left([\mathbf{P}_i \mathbf{P}_j]\right)$$

33

The Cross-ratio

which means that

$$\frac{d(\mathbf{P}_1, \mathbf{P}_2)d(\mathbf{P}_3, \mathbf{P}_4)}{d(\mathbf{P}_1, \mathbf{P}_3)d(\mathbf{P}_2, \mathbf{P}_4)}$$

is a *projective invariant* — computing the value of this *cross ratio* using $\mathbf{P}_1, \ldots, \mathbf{P}_4$ or using $\mathcal{M}\mathbf{P}_1, \ldots, \mathcal{M}\mathbf{P}_4$ will yield the same number, as long as $\mathcal{M}$ is a projective transformation.

34

The Cross-ratio

Now check that the cross-ratio of the four points $(0, 1)$, $(a, 1)$, $(b, 1)$ and $(1, 0)$ is a/b (notice the last point is the point at infinity). We can use this observation to measure height relative to a reference object. Using the notation of Figure 24.6, we

35

The Cross-ratio

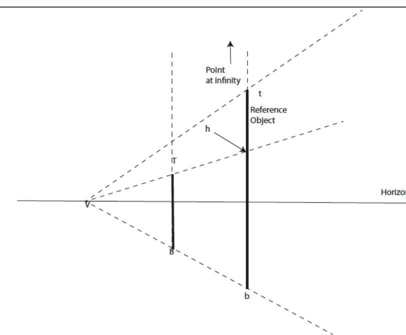


FIGURE 24.6: *A perspective camera views a reference object and another object perpendicular to a ground plane. This produces a line segment in the image plane. Constructing appropriate lines in the figure and taking a cross ratio yields the height of the object.*

36

The Cross-ratio

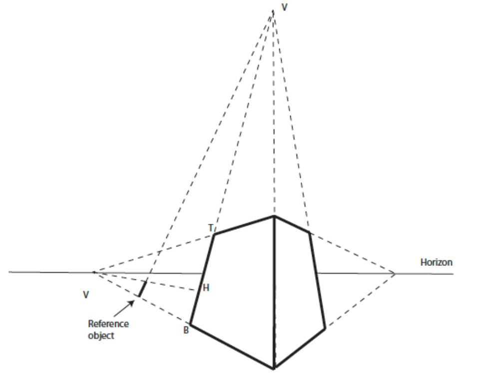
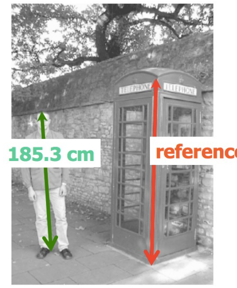


FIGURE 24.7: *A building and a person viewed in a more extreme perspective view than that of 24.6. The person has known height, and can act as reference object. The same construction as in that figure yields the height of the building relative to the person.*

37

Single-view measurement examples

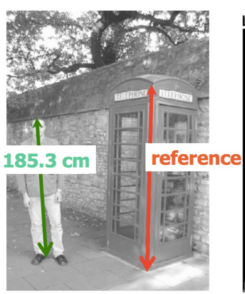


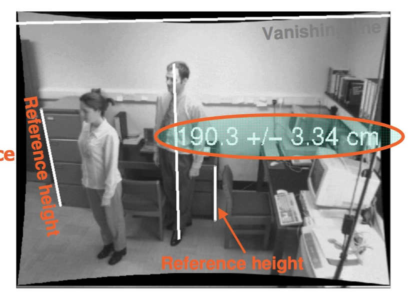
That booth is still there! (Oxford, September 2022)

A. Criminisi, I. Reid, and A. Zisserman, Single View Metrology, IJCV 2000
Figure from UPenn CIS580 slides

50

Single-view measurement examples





A. Criminisi, I. Reid, and A. Zisserman, Single View Metrology, IJCV 2000
Figure from UPenn CIS580 slides

51

Another example

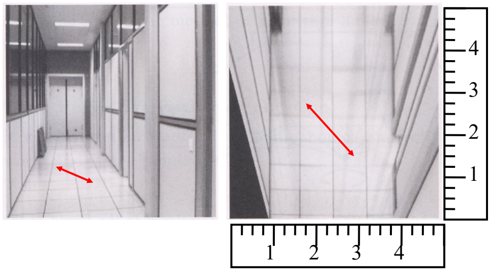
- Are the heights of the two groups of people consistent with one another?
  - Measure heights using Christ as reference

Piero della Francesca, *Flagellation,* ca. 1455

A. Criminisi, M. Kemp, and A. Zisserman, Bringing Pictorial Space to Life: computer techniques for the analysis of paintings, *Proc. Computers and the History of Art*, 2002

52

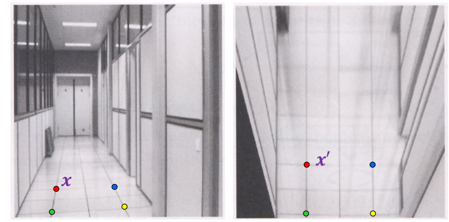
Measurements within planes



- Simplest approach: unwarp then measure
- What kind of warp is this?

53

Image rectification



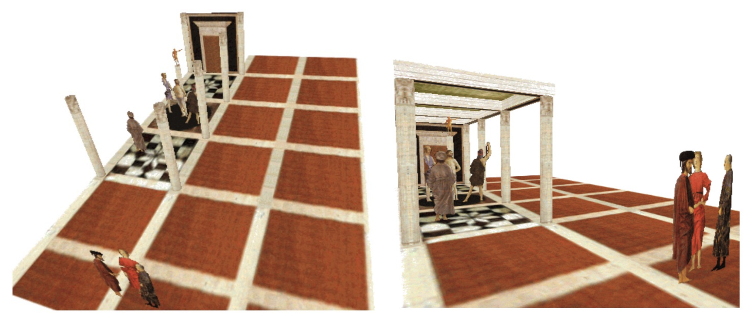
- To unwarp (rectify) an image, solve for homography $\boldsymbol{H}$ given four pairs of matches assumed to have a known configuration (e.g., square)

54

Image rectification: Example

55

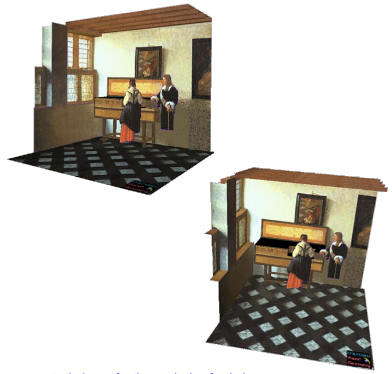
Putting everything together: Single-view modeling



A. Criminisi et al. Bringing Pictorial Space to Life: computer techniques for the analysis of paintings, *Proc. Computers and the History of Art*, 2002

56

Putting everything together: Single-view modeling



J. Vermeer, Music Lesson, 1662

A. Criminisi et al. Bringing Pictorial Space to Life: computer techniques for the analysis of paintings, *Proc. Computers and the History of Art*, 2002

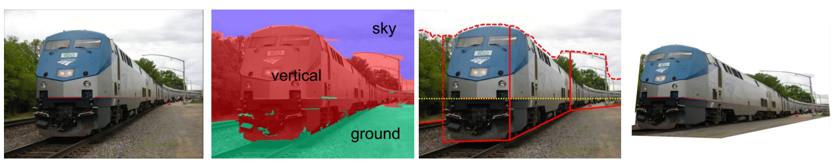
57

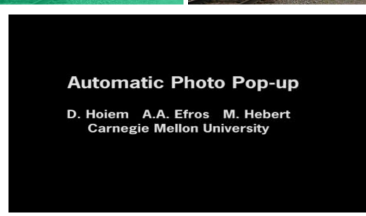
Outline

- Camera calibration using vanishing points
- Measurements from a single image
- Applications of single-view metrology

58

Application: Fully automatic modeling





D. Hoiem, A.A. Efros, and M. Hebert, Automatic Photo Pop-up, SIGGRAPH 2005

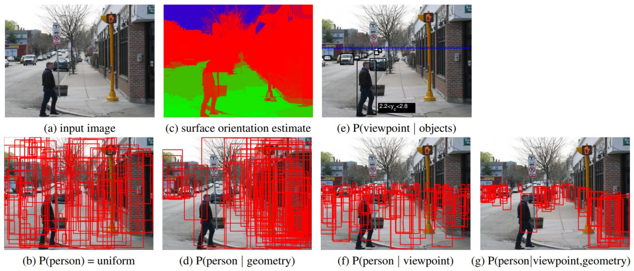
59

Application: Object detection

D. Hoiem, A.A. Efros, and M. Hebert. Putting Objects in Perspective, CVPR 2006

60

Application: Object detection



D. Hoiem, A.A. Efros, and M. Hebert. Putting Objects in Perspective. CVPR 2006

62

Application: Image editing

- Inserting synthetic objects into images

K. Karsch and V. Hedau and D. Forsyth and D. Hoiem. Rendering Synthetic Objects into Legacy Photographs. *SIGGRAPH Asia* 2011

63

Application: Image editing

- Inserting synthetic objects into images:

K. Karsch and V. Hedau and D. Forsyth and D. Hoiem. Rendering Synthetic Objects into Legacy Photographs. *SIGGRAPH Asia* 2011

64