Perspective projection

A. Mantegna, *Martyrdom of St. Christopher*, c. 1450

Overview

- Motivation: recovery of 3D structure
- Pinhole projection model
- Properties of projection
- Perspective projection matrix
- Orthographic projection

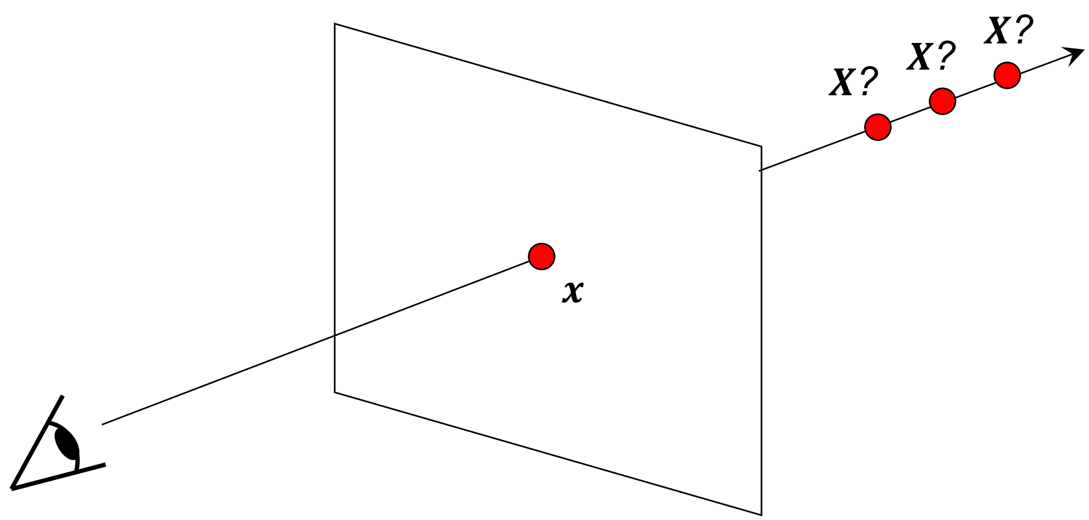
Given an image, can we recover 3D structure?

J. Vermeer, *Music Lesson*, 1662

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

Things aren’t always as they appear…

Single-view ambiguity



Single-view ambiguity

[Rashad Alakbarov shadow sculptures]

Anamorphic perspective

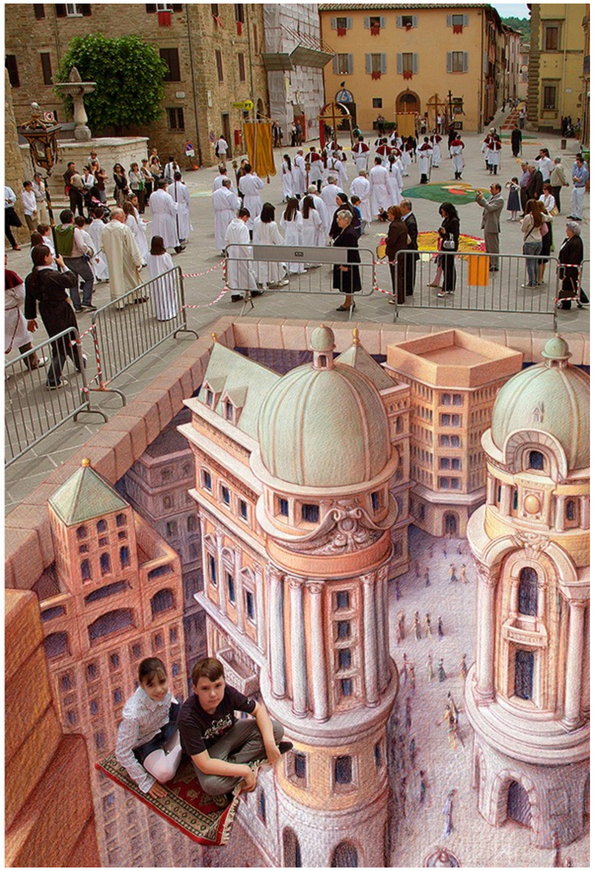


Image source

Anamorphic perspective

H. Holbein The Younger, *The Ambassadors*, 1533

https://en.wikipedia.org/wiki/Anamorphosis

Our goal: Recovery of 3D structure

- When certain assumptions hold, we can recover structure from a single view

- In general, we need *multi-view geometry*

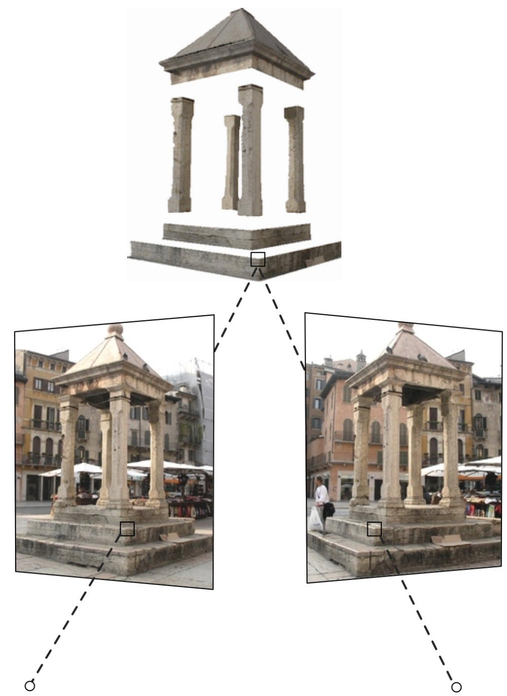


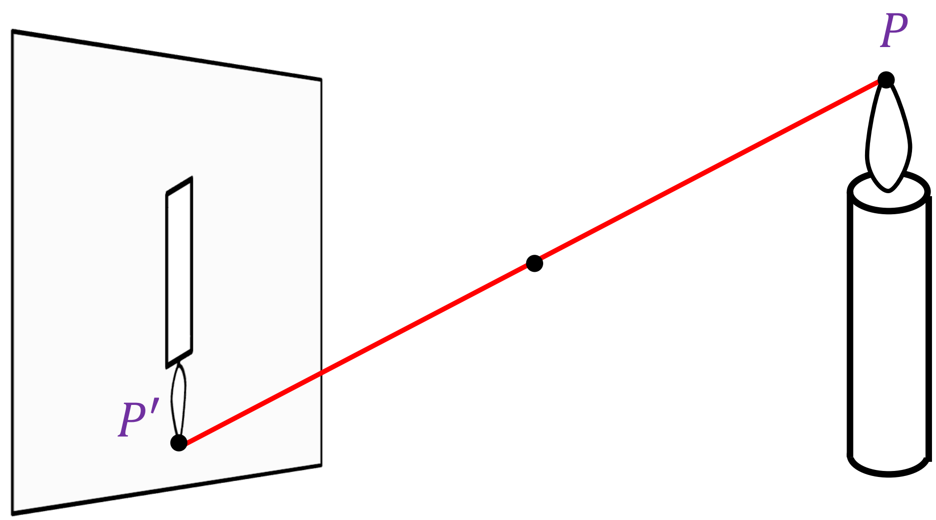
Image source

- But first, we need to understand the geometry of a single camera…

Overview

- Motivation: recovery of 3D structure
- Pinhole projection model

Review: Perspective projection equations



Review: Perspective projection equations

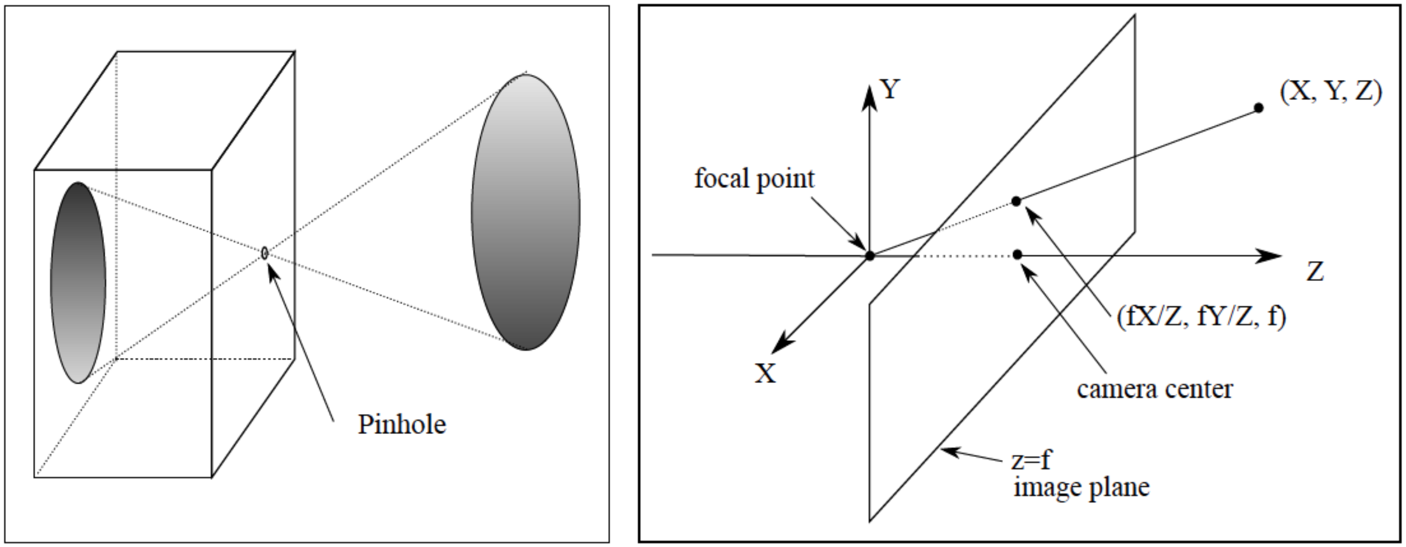
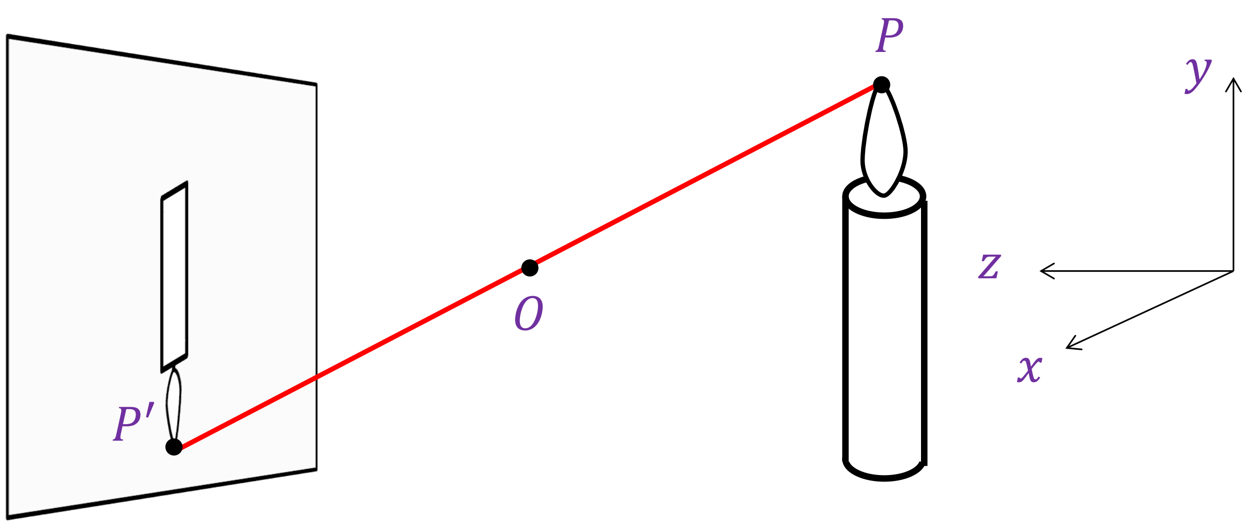


FIGURE 19.1: *The pinhole imaging model. On the* **left***, a light-tight box with a pinhole in it views an object. The only light that a point on the back of the box sees comes through the very small pinhole, so that an inverted image is formed on the back face of the box. On the* **right***, the usual geometric abstraction. The box doesn't affect the geometry, and is omitted. The pinhole has been moved to the back of the box, so that the image is no longer inverted. The image is formed on the plane $z = f$, by convention. Notice the coordinate system is* left-handed, *because the camera looks down the z-axis. This is because most people's intuition is that z* increases *as one moves into the image. The text provides some more detail on this point.*

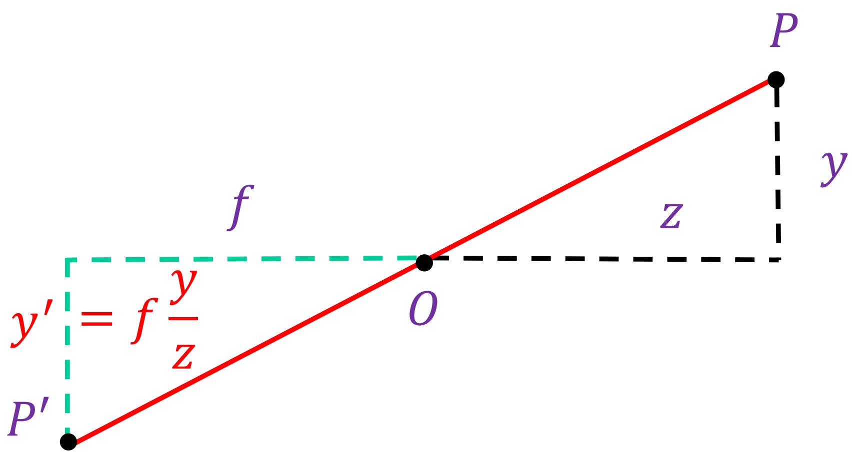
Review: Perspective projection equations



Canonical coordinate system

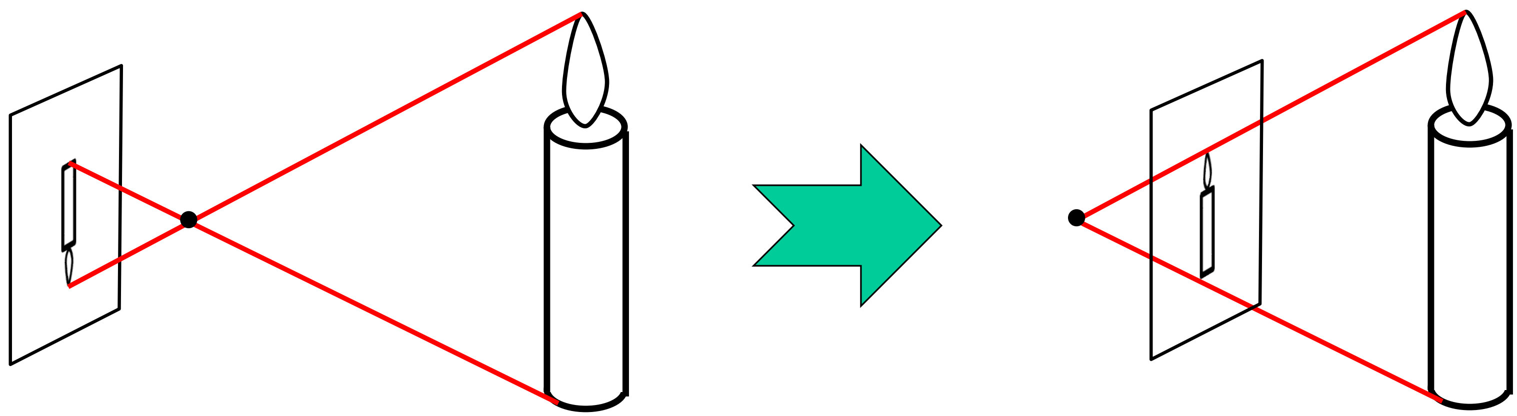
- The optical center (O) is at the origin
- The z axis is the *optical axis* perpendicular to the image plane
- The xy plane is parallel to the image plane, x and y axes are horizontal and vertical directions of the image plane

Review: Perspective projection equations



$$\boxed{(x, y, z) \rightarrow \left(f\,\frac{x}{z}, f\,\frac{y}{z}\right)}$$

Review: Perspective projection equations



Note: instead of dealing with an image that is upside down, most of the time we will pretend that the image plane is *in front* of the camera center
This will get us a minus sign in these coordinates {(X, Y, Z)-> (-fX/Z, -fY/Z)}

Overview

- Motivation: recovery of 3D structure
- Pinhole projection model
- Properties of projection
  - What happens to:
    - Lines
    - Planes
    - General 3D shapes

Lines project to lines; distant objects are smaller

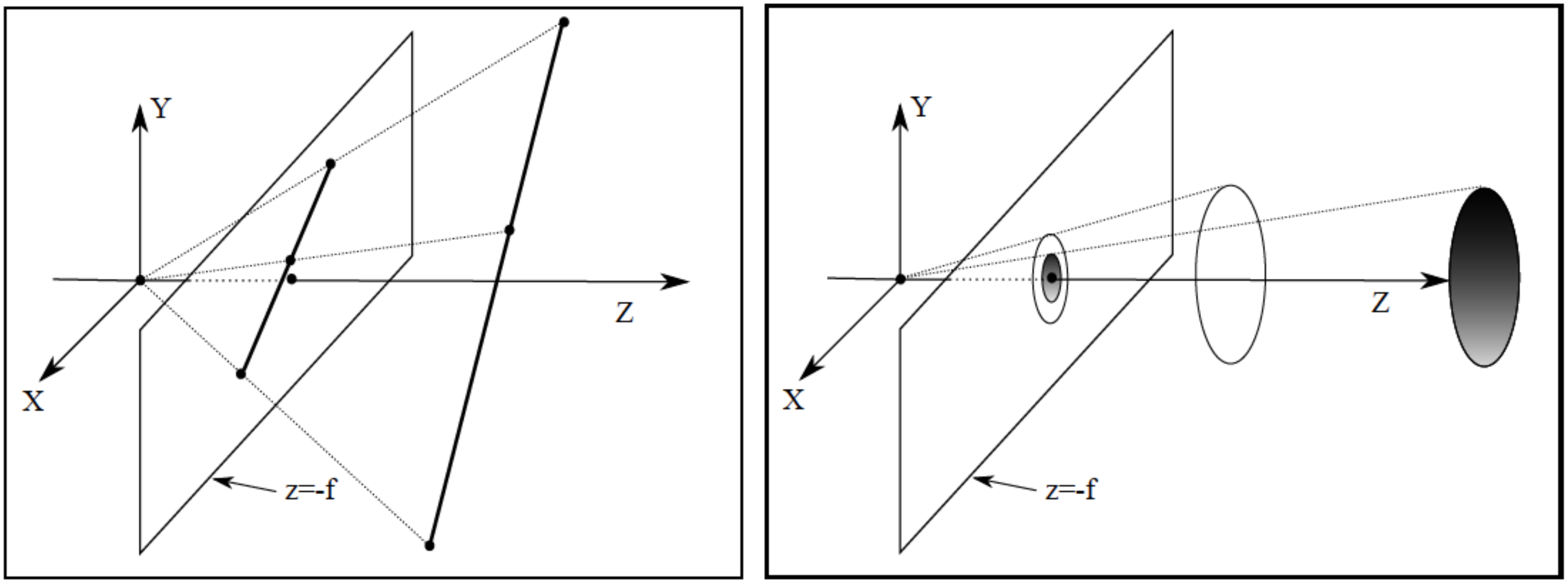
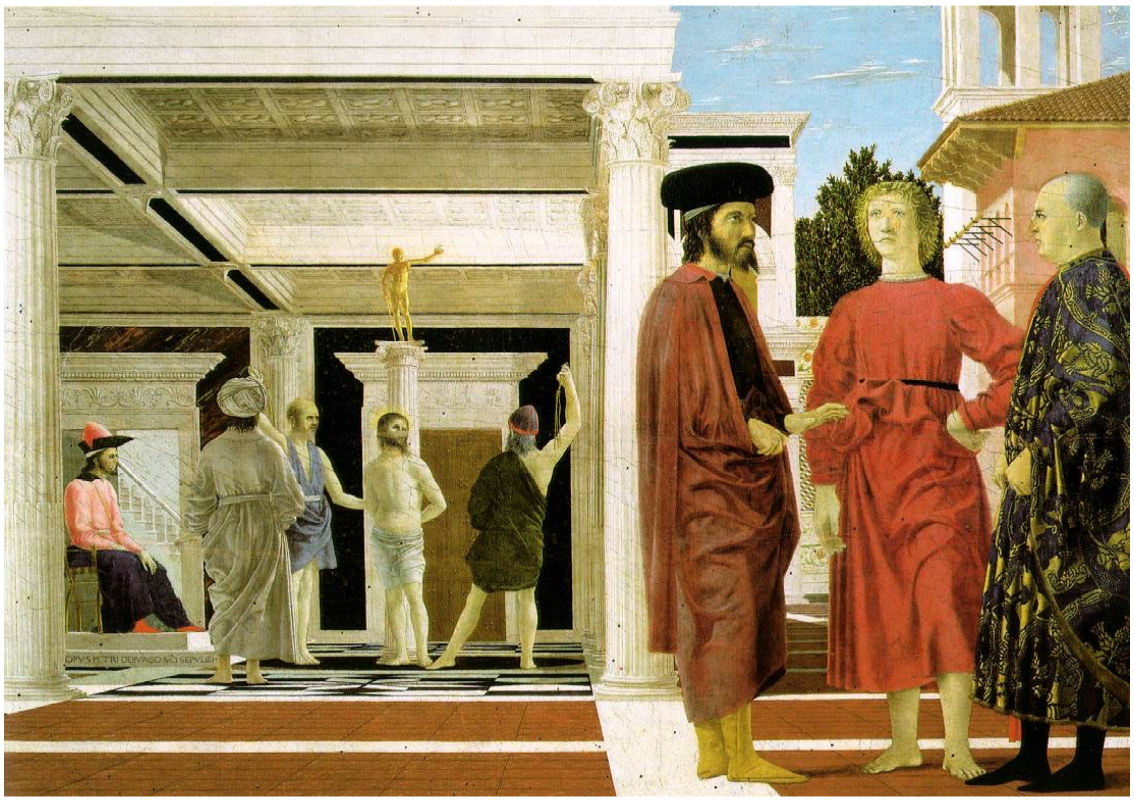


FIGURE 19.2: *Perspective projection maps almost any 3D line to a line in the image plane (***left***). Some rays from the focal point to points on the line are shown as dotted lines. The family of all such rays is a plane, and that plane must intersect the image plane in a line as long as the 3D line does not pass through the focal point. On the* **right***, two 3D objects viewed in perspective projection; the more distant object appears smaller in the image.*

Projection of lines



Piero della Francesca, *Flagellation of Christ*, 1455-1460

Parallel lines have a vanishing point

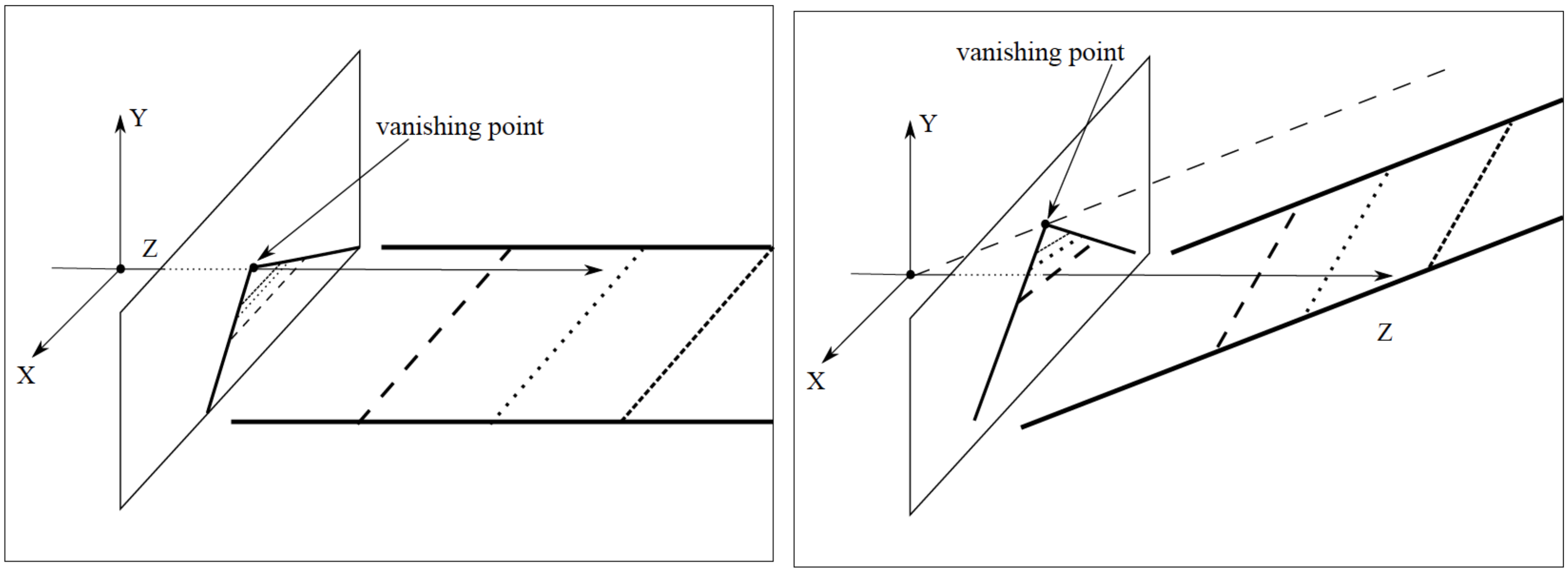
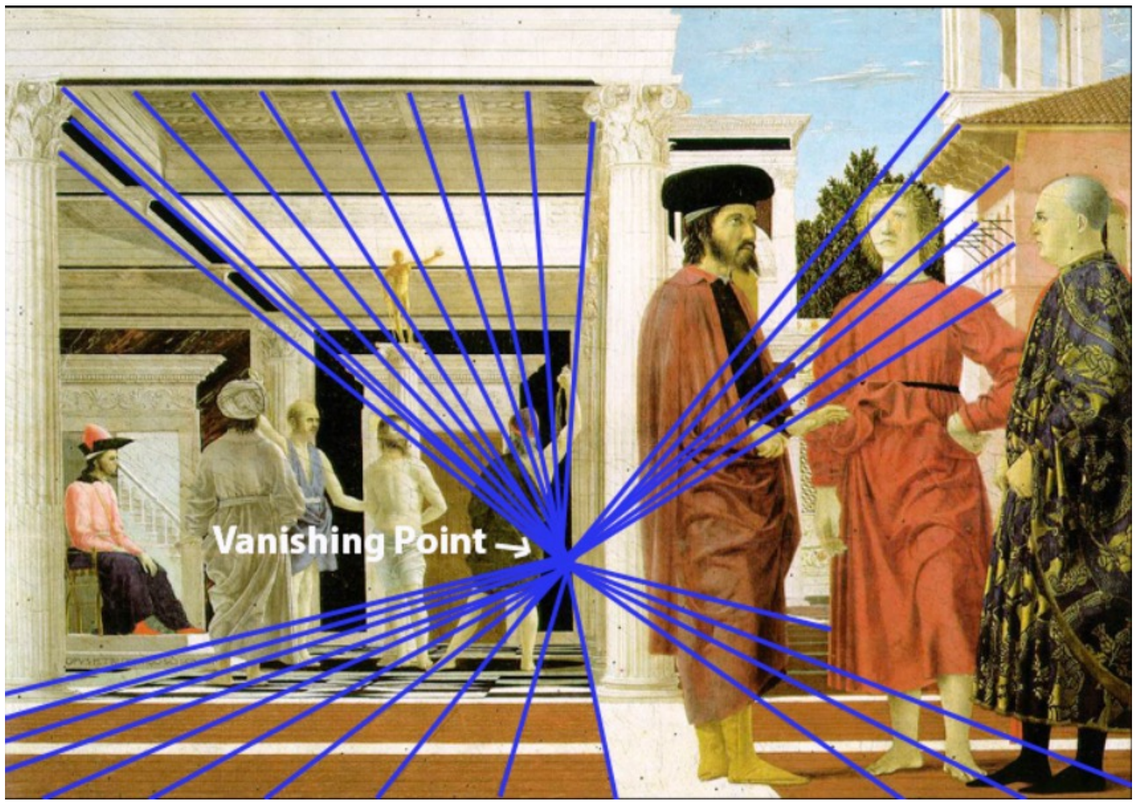


FIGURE 19.3: *Perspective projection maps a set of parallel lines to a set of lines that meet in a point. On the* **left**, *a set of lines parallel to the z-axis, with "railway sleepers" shown. As these sleepers get further away, they get smaller in the image, meaning the projected lines must meet. The vanishing point (the point where they meet) is obtained by intersecting the ray parallel to the lines and through the focal point with the image plane. On the* **right**, *a different pair of parallel lines with a different vanishing point. The figure establishes that, if there are more than two lines in the set of parallel lines, all will meet at the vanishing point.*

Projection of lines

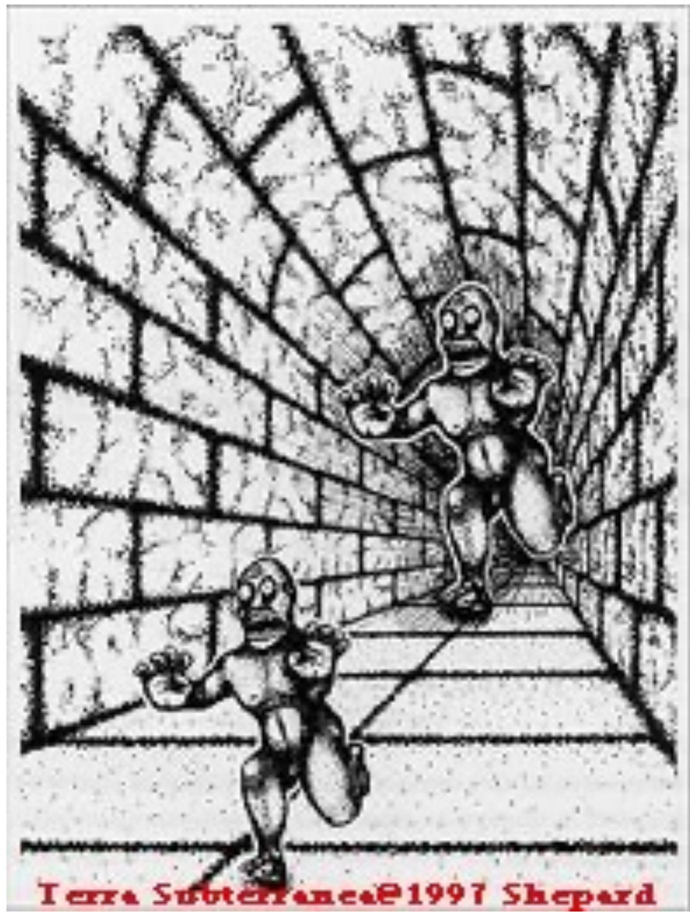
- Parallel lines meet at a *vanishing point*



Piero della Francesca, *Flagellation of Christ*, 1455-1460

Image source

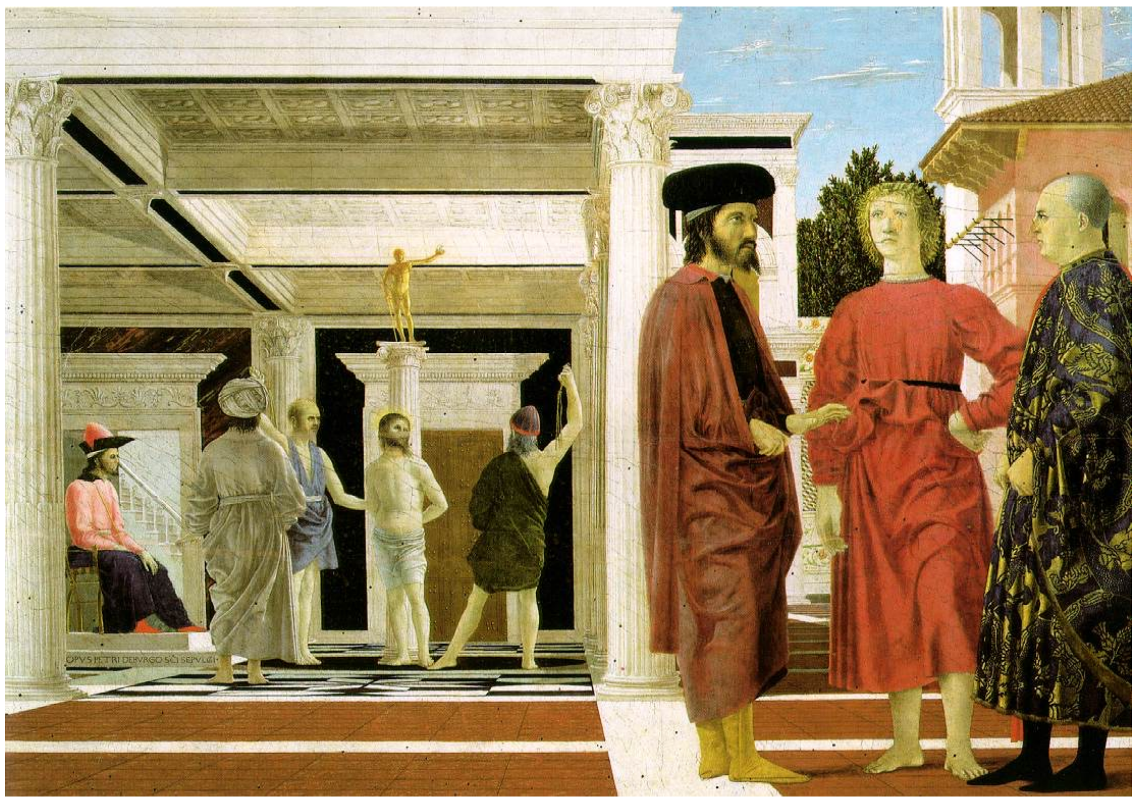
Converging lines are a powerful perspective cue



Slide by Steve Seitz

Projection of lines

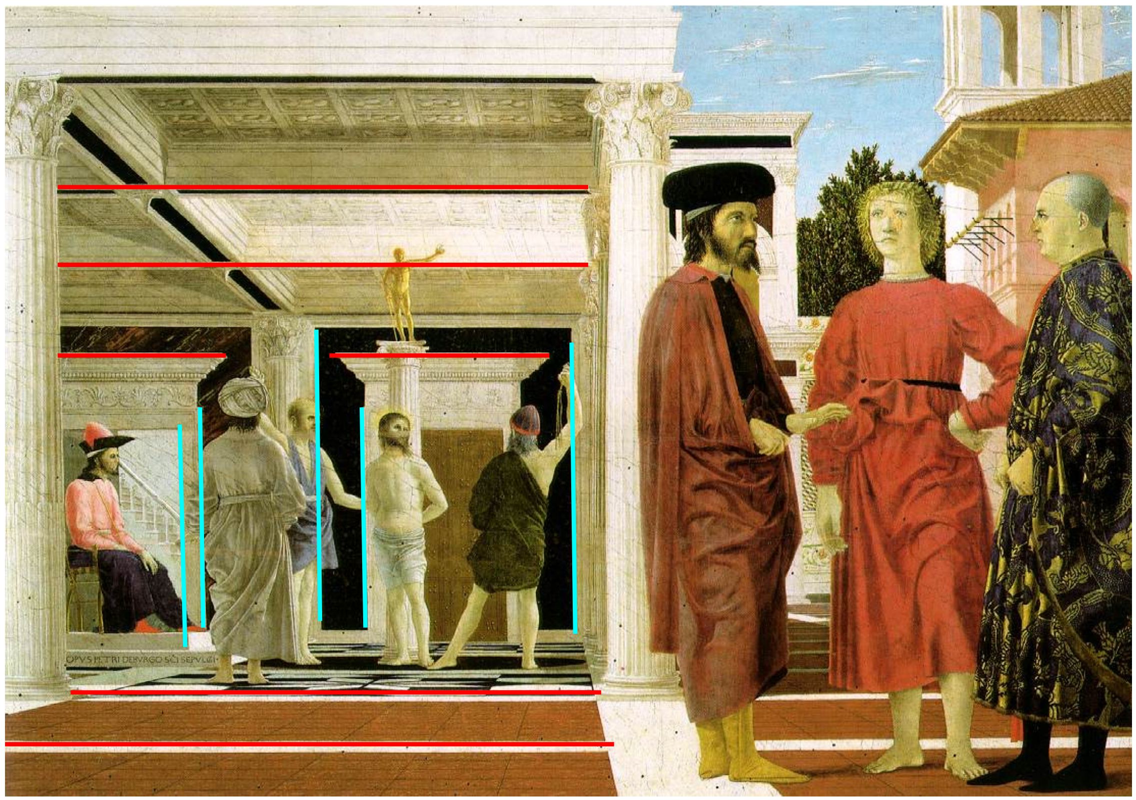
- Do *all* parallel lines converge to a vanishing point?



Piero della Francesca, *Flagellation of Christ*, 1455-1460

Projection of lines

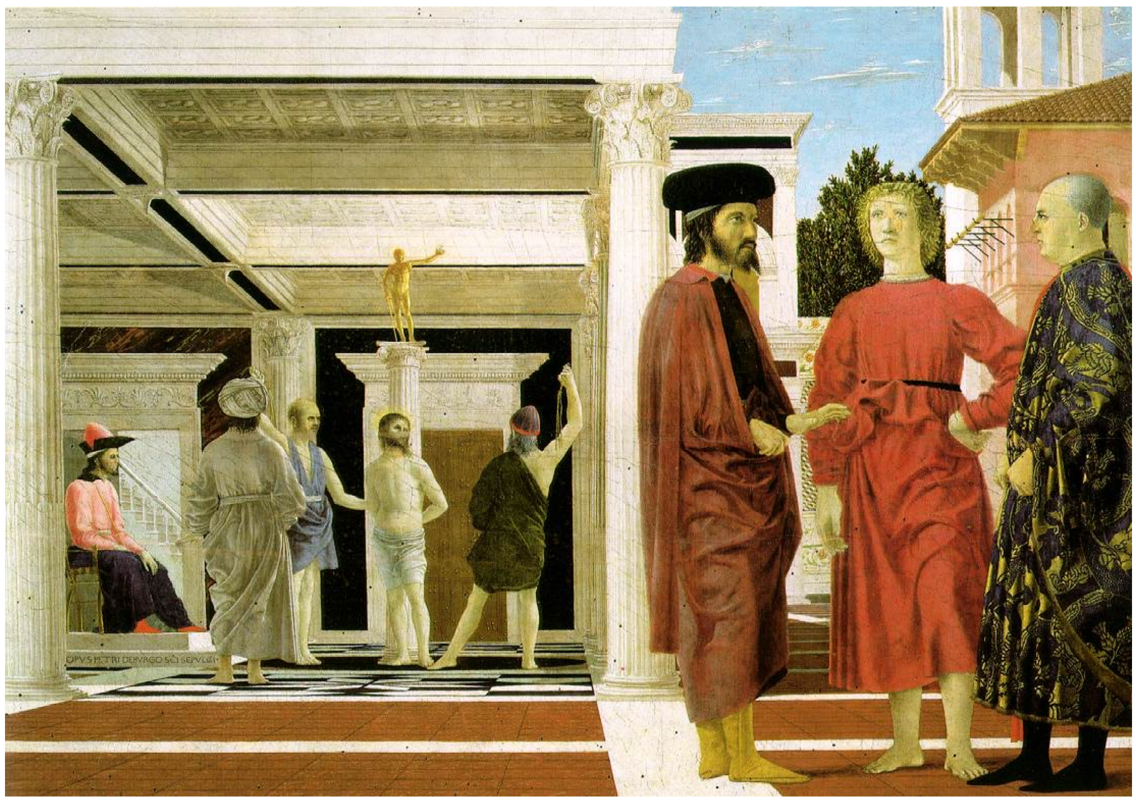
- Parallel lines that are also parallel to the image plane converge to a point "at infinity"



Piero della Francesca, *Flagellation of Christ*, 1455-1460

Projection of planes

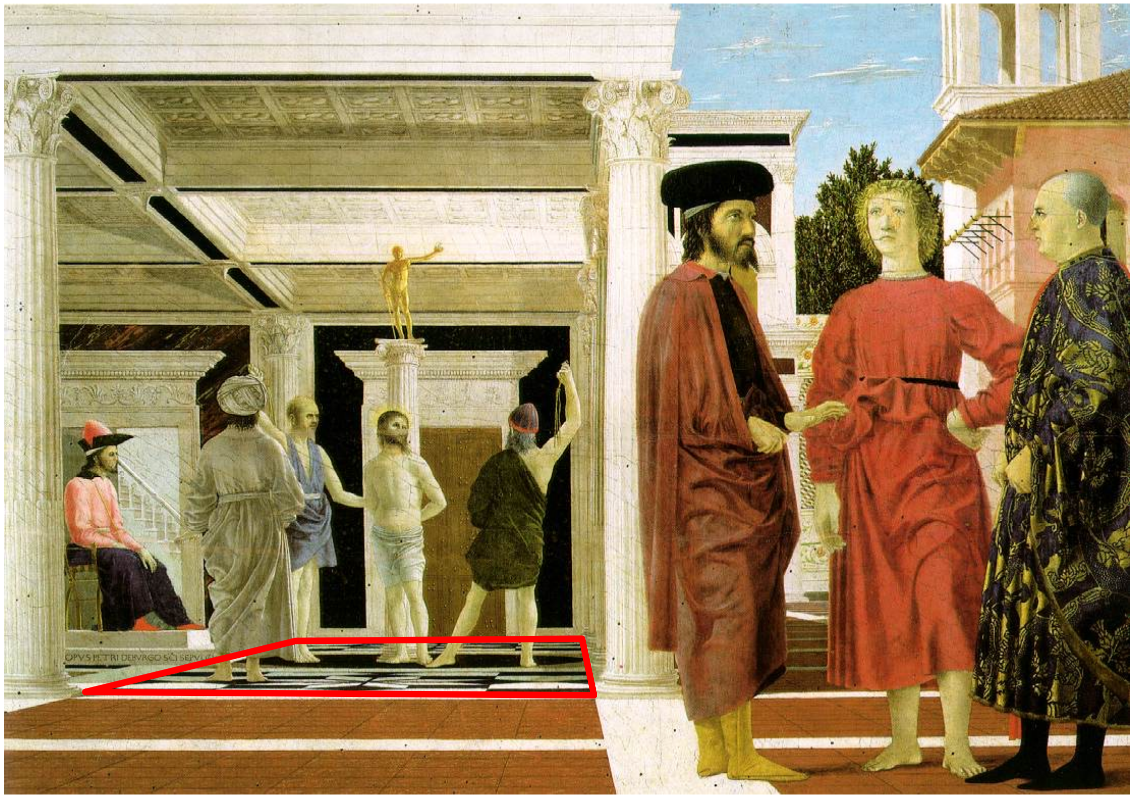
- What happens to planar patterns?



Piero della Francesca, *Flagellation of Christ*, 1455-1460

Projection of planes

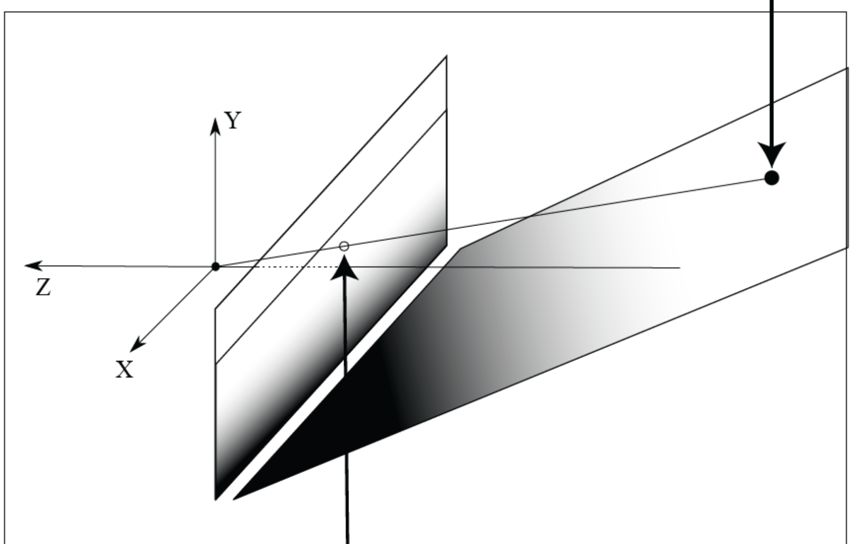
- Patterns on *non-fronto-parallel* planes are distorted by a 2D homography



Piero della Francesca, *Flagellation of Christ*, 1455-1460

A plane pattern is transformed by a homography

$$\begin{pmatrix} X(s,t) \\ Y(s,t) \\ Z(s,t) \end{pmatrix} = \begin{pmatrix} a_{11}s + a_{12}t + a_{13} \\ a_{21}s + a_{22}t + a_{23} \\ a_{31}s + a_{32}t + a_{33} \end{pmatrix}$$

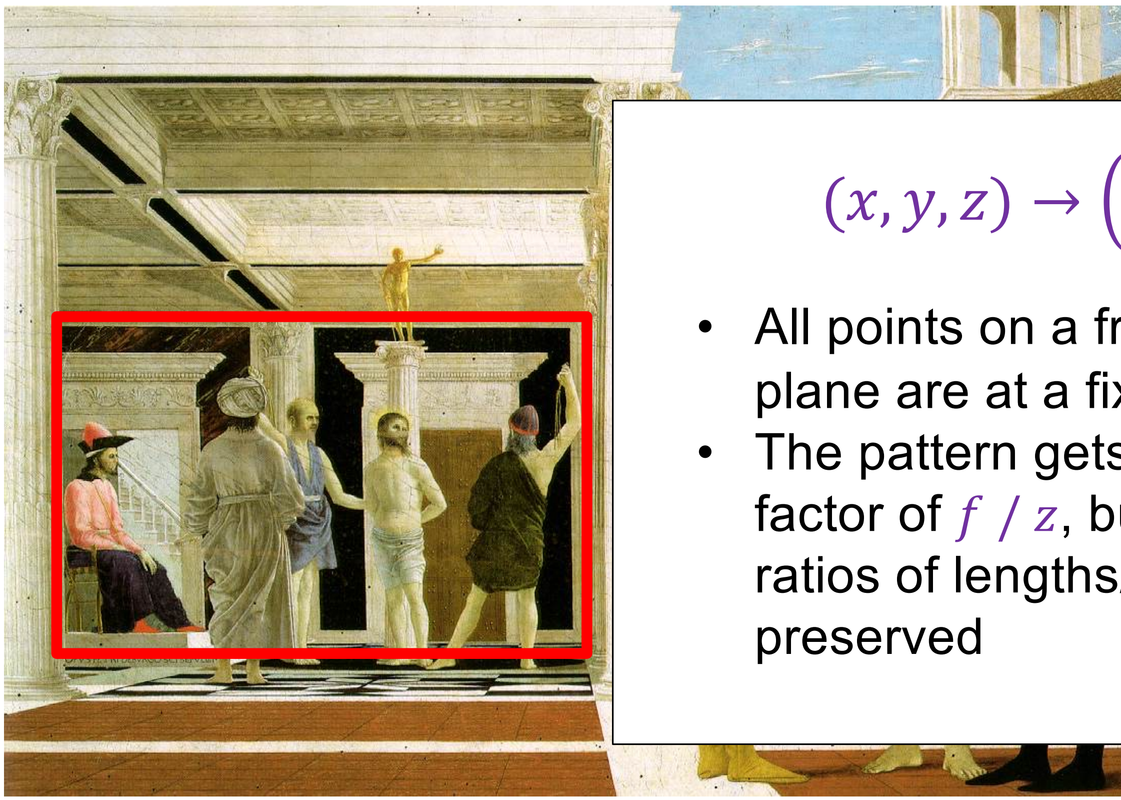


Here s, t are coordinates on the plane

$$\begin{pmatrix} \frac{X(s,t)}{Z(s,t)} \\ \\ \frac{Y(s,t)}{Z(s,t)} \end{pmatrix} = \begin{pmatrix} \frac{a_{11}s+a_{12}t+a_{13}}{a_{31}s+a_{32}t+a_{33}} \\ \\ \frac{a_{21}s+a_{22}t+a_{23}}{a_{31}s+a_{32}t+a_{33}} \end{pmatrix}$$

Projection of planes

- What about patterns on *fronto-parallel planes*?



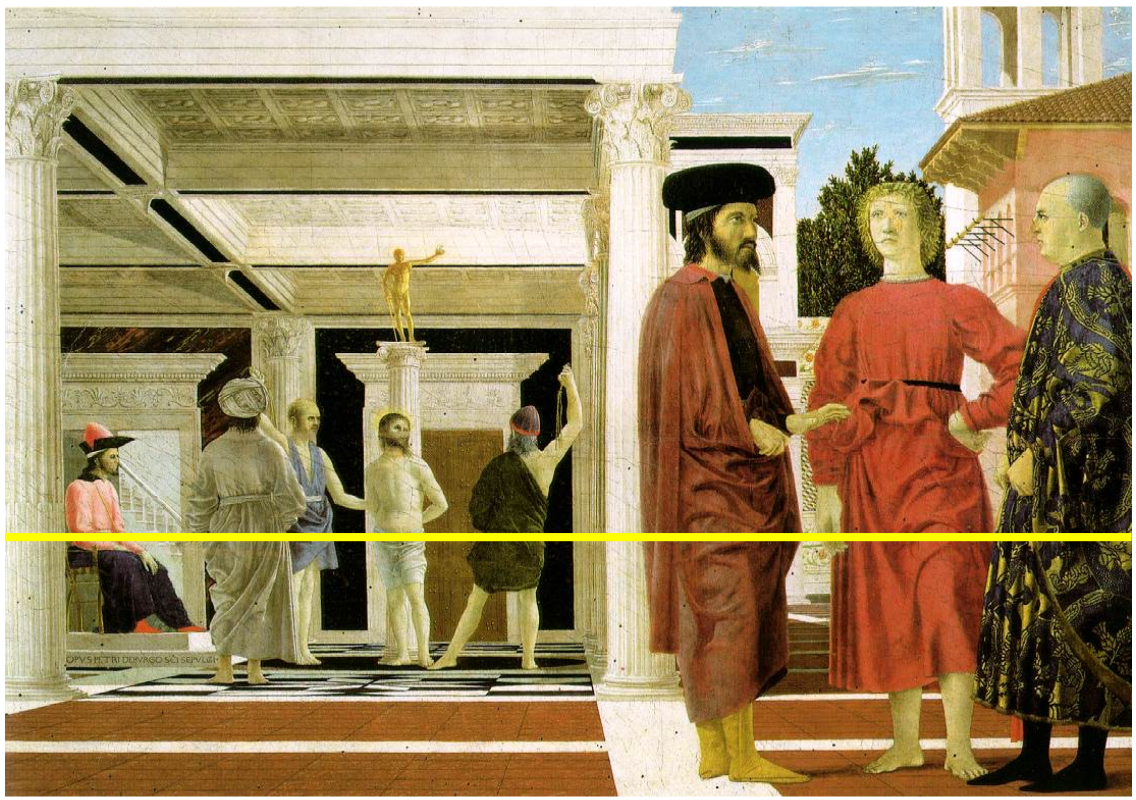
$$(x, y, z) \rightarrow \left(f\frac{x}{z}, f\frac{y}{z}\right)$$

- All points on a fronto-parallel plane are at a fixed depth z
- The pattern gets scaled by a factor of $f\ /\ z$, but angles and ratios of lengths/areas are preserved

Piero della Francesca, *Flagellation of Christ*, 1455-1460

Projection of planes

- Non-fronto-parallel planes have *vanishing lines*

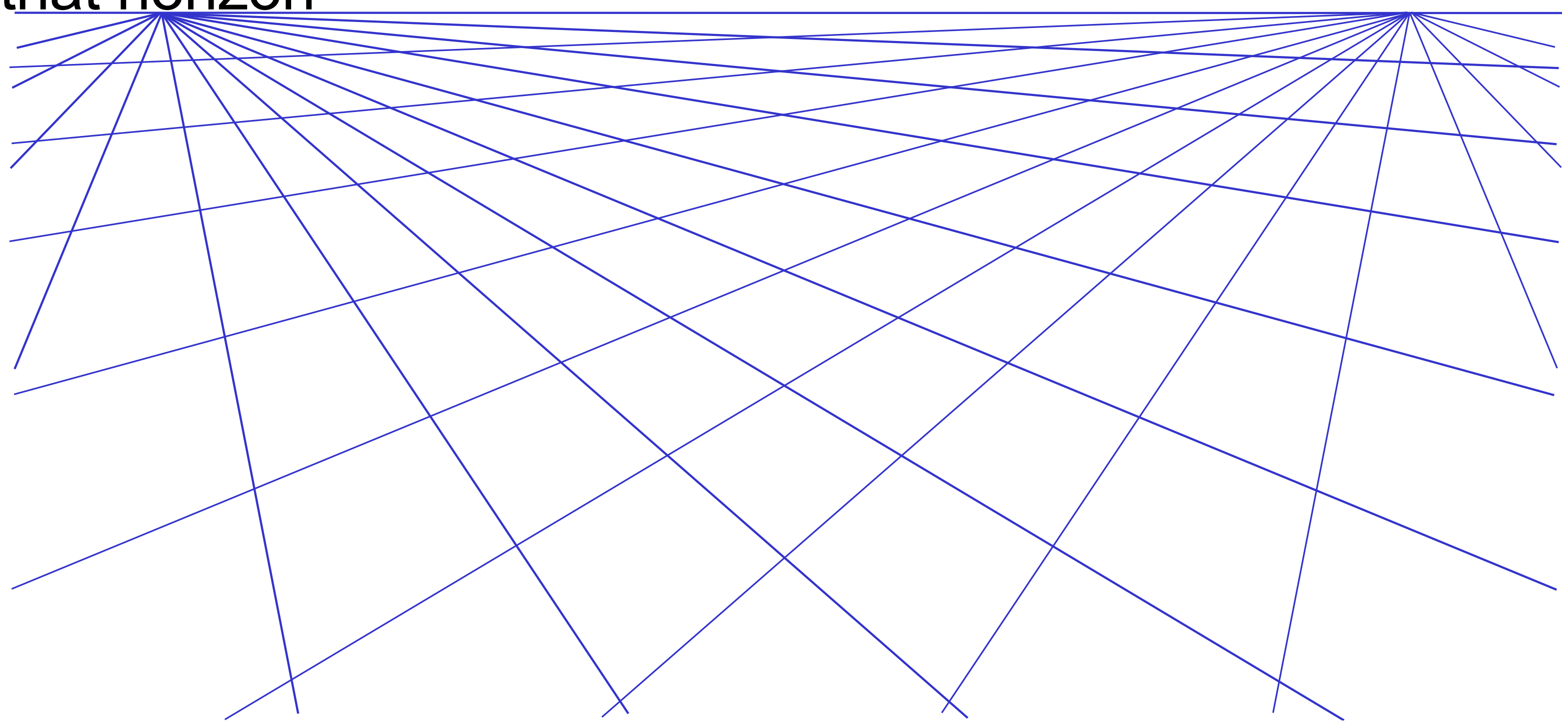


Vanishing line of the ground plane

Piero della Francesca, *Flagellation of Christ*, 1455-1460

Vanishing lines of planes

- Each family of parallel planes is associated with a *vanishing line (=horizon)* in the image
- Sets of lines parallel to that plane have vanishing points on that horizon



Horizons (= vanishing lines)

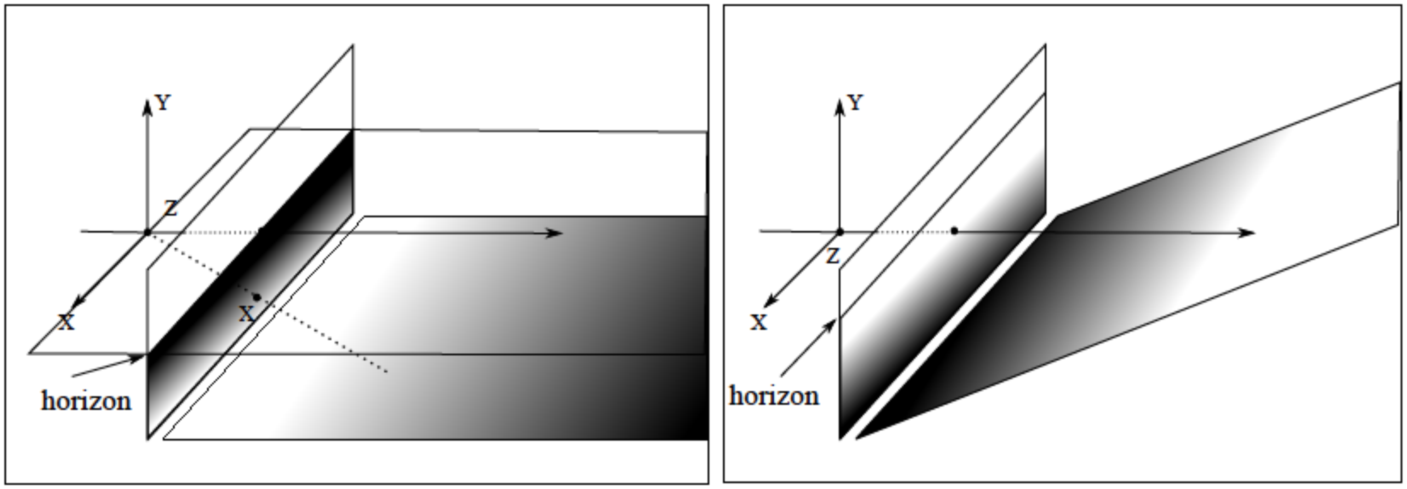
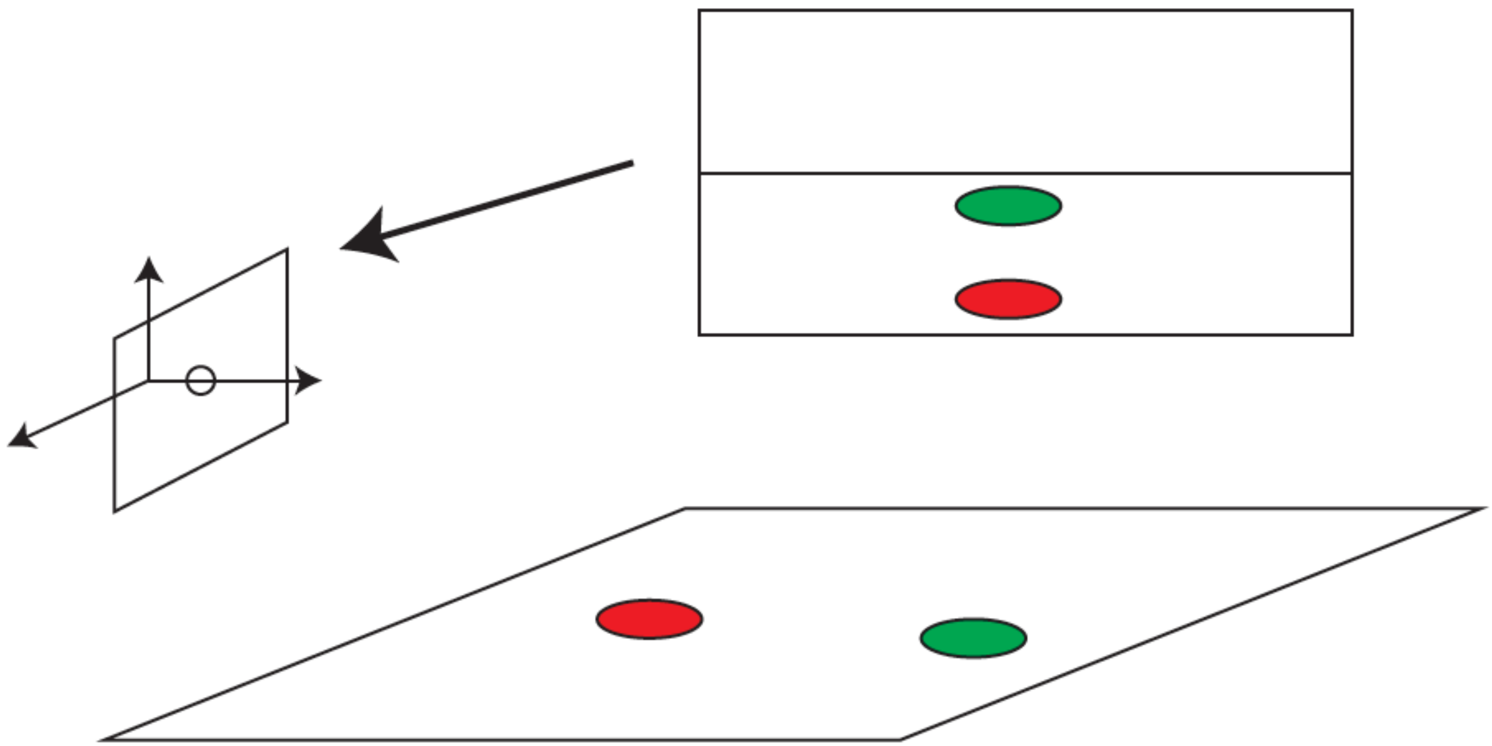


FIGURE 19.4: **Left** *shows a plane in 3D (in this case, $y = -1$). The intersection of the plane through the focal point parallel to the 3D plane (in this case, $y = 0$) and the image plane, forms an image line called the horizon. This line cuts the image plane into two parts. Construct the ray through the focal point and a point* $\mathbf{x}$ *in the image plane. For* $\mathbf{x}$ *on one side of the horizon, this ray will intersect the 3D plane in the half space $z > 0$ (and so in front of the camera, shown here). If* $\mathbf{x}$ *is on the other side of the horizon, the intersection will be in the half space $z < 0$ (and so behind the camera, where it cannot be seen).* **Right** *shows a different 3D plane with a different horizon. The gradients on the planes indicate roughly where points on the 3D plane appear in the image plane (light points map to light, dark to dark).*

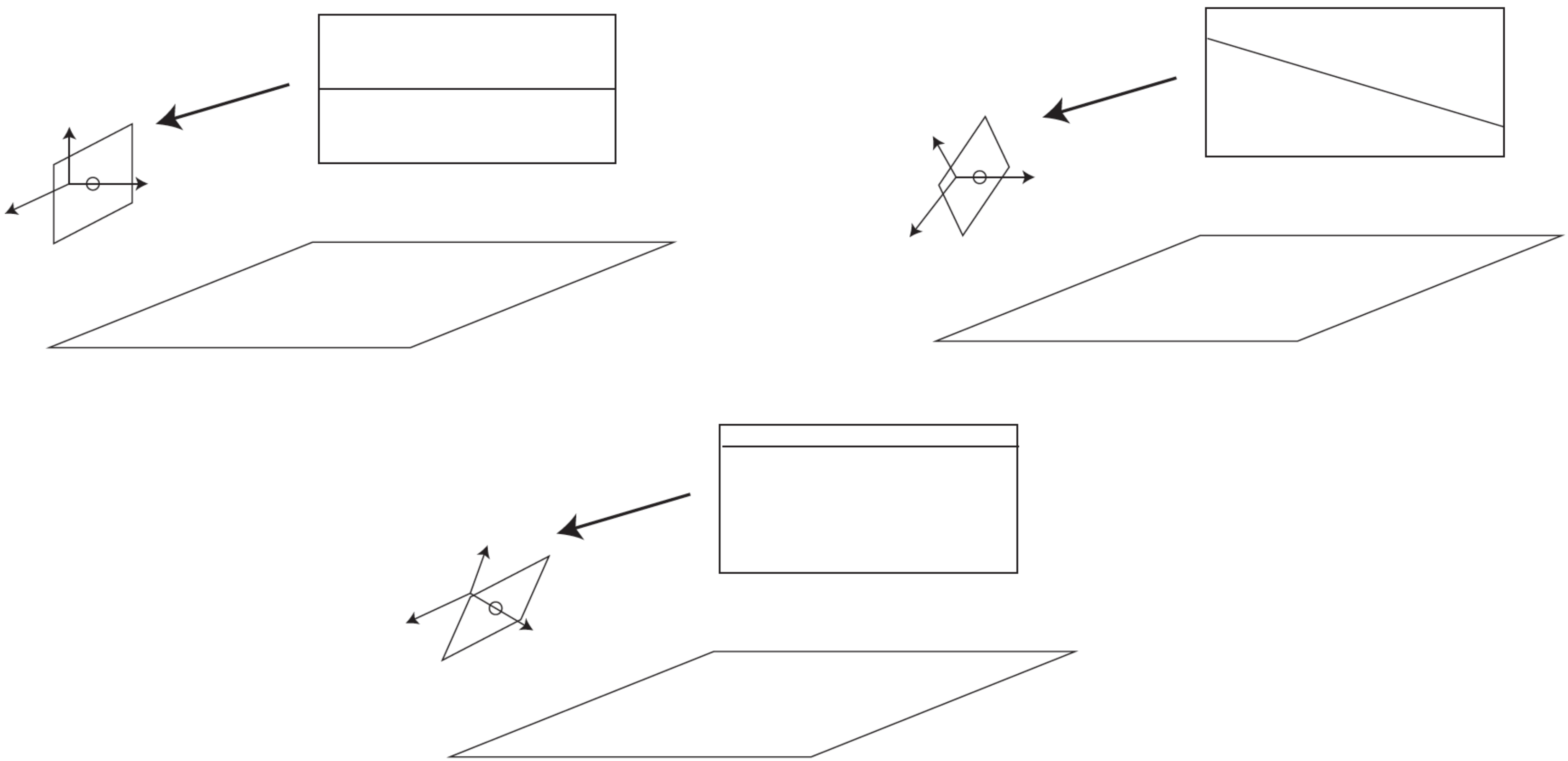
Horizons are informative - I



Horizons are informative - II

Image source: S. Seitz

Horizons are informative, III



Projection of 3D shapes

- What is are the relationships between the geometric properties of general 3D surfaces and their 2D projections?

Barbara Hepworth sculpture

Projection of 3D shapes

- What is the shape of the projection of a sphere?

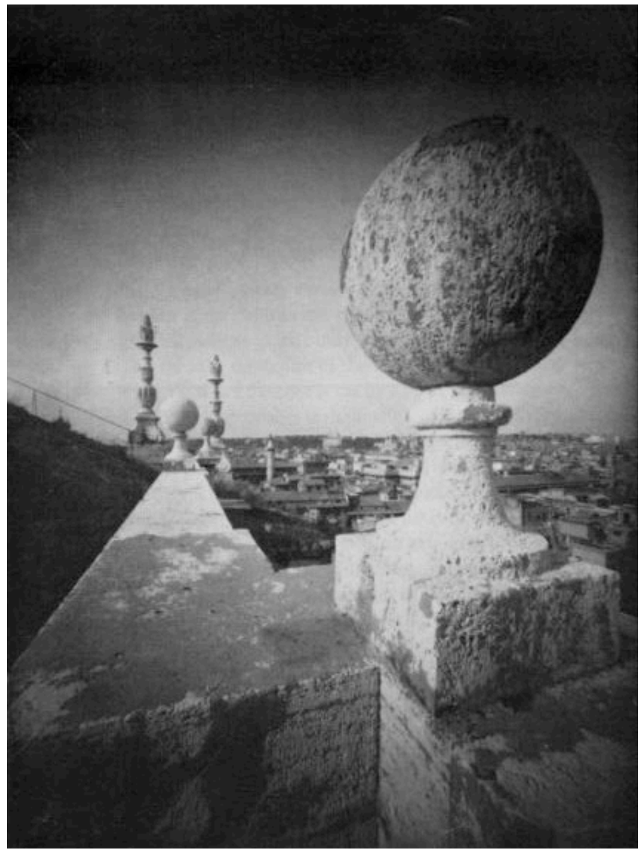
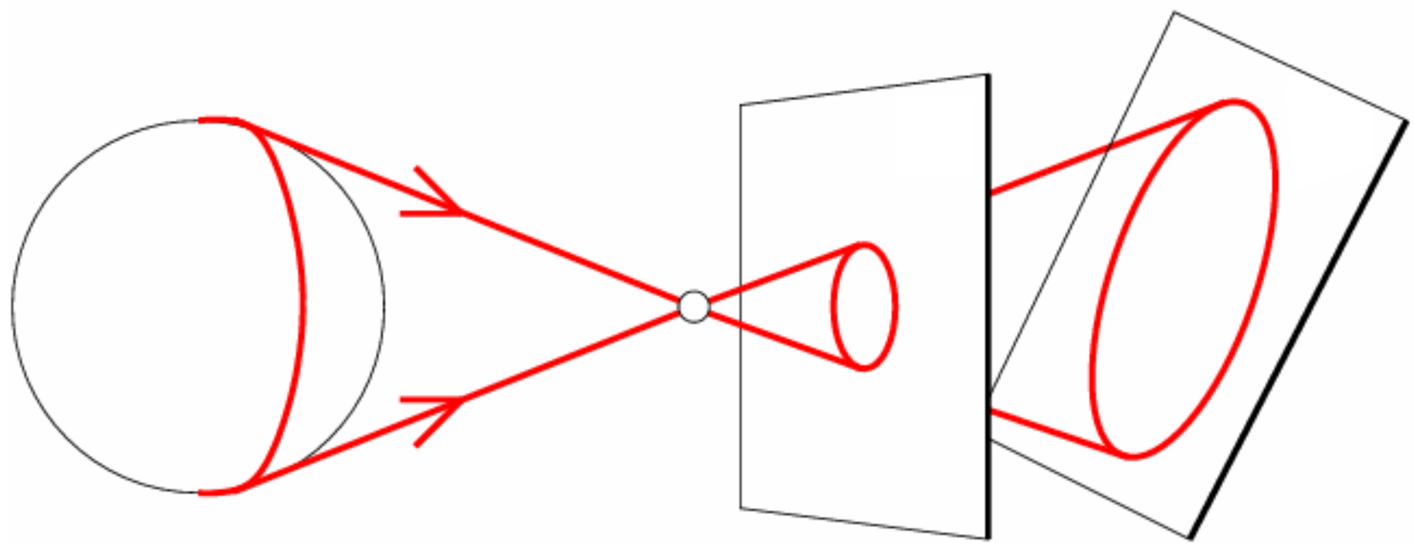


Image source: F. Durand

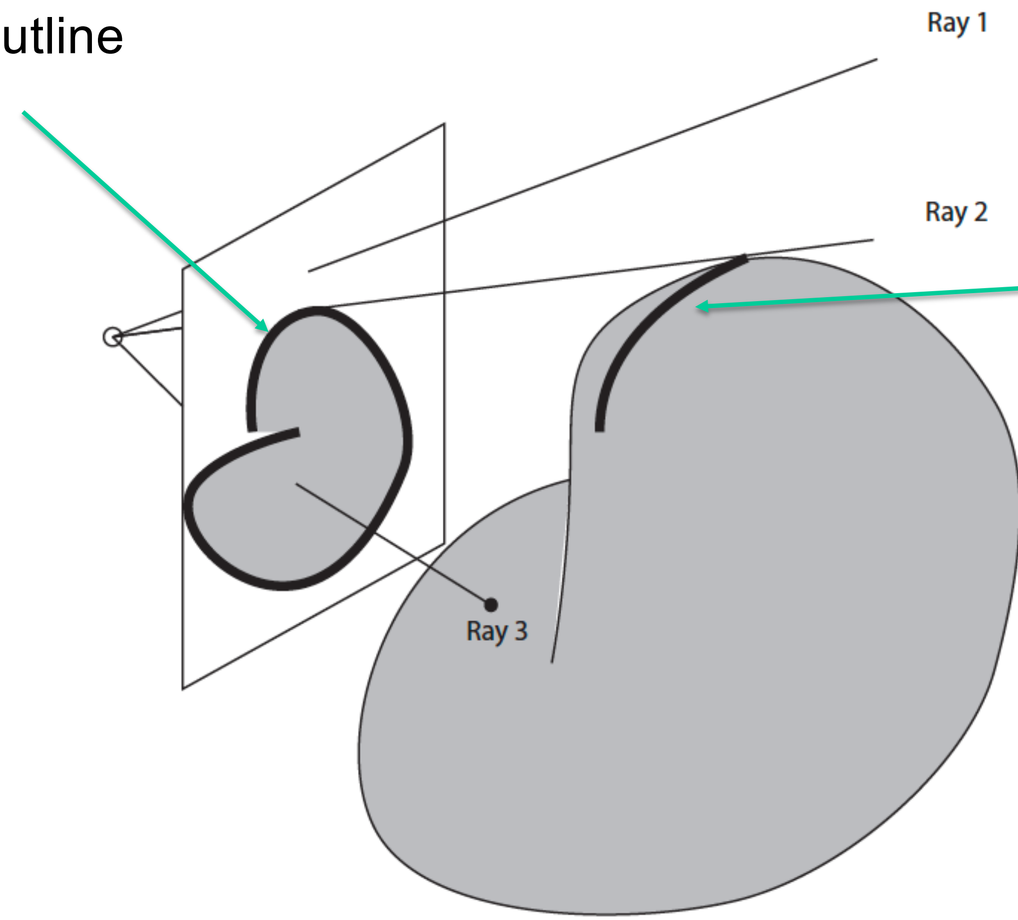
Projection of 3D shapes

- What is the shape of the projection of a sphere?



Projection of 3D shapes

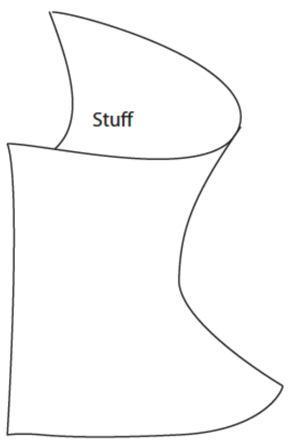
Outline

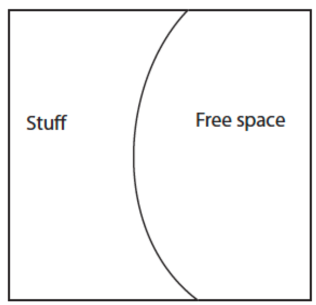


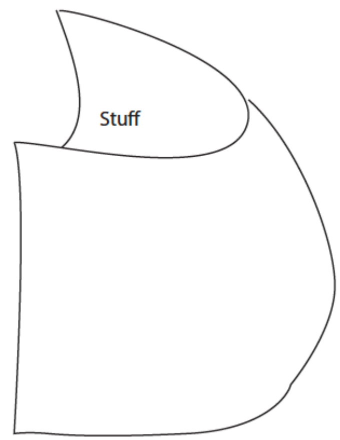
Contour generator

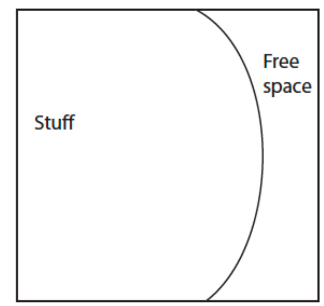
Relation between contour generator
And outline is geometrically very
Complicated. Surprisingly hard to say
Much about surface from outline.

Projection of 3D shapes









J. Koenderink. What does the occluding contour tell us about solid shape? Perception 13 (321-330), 1984

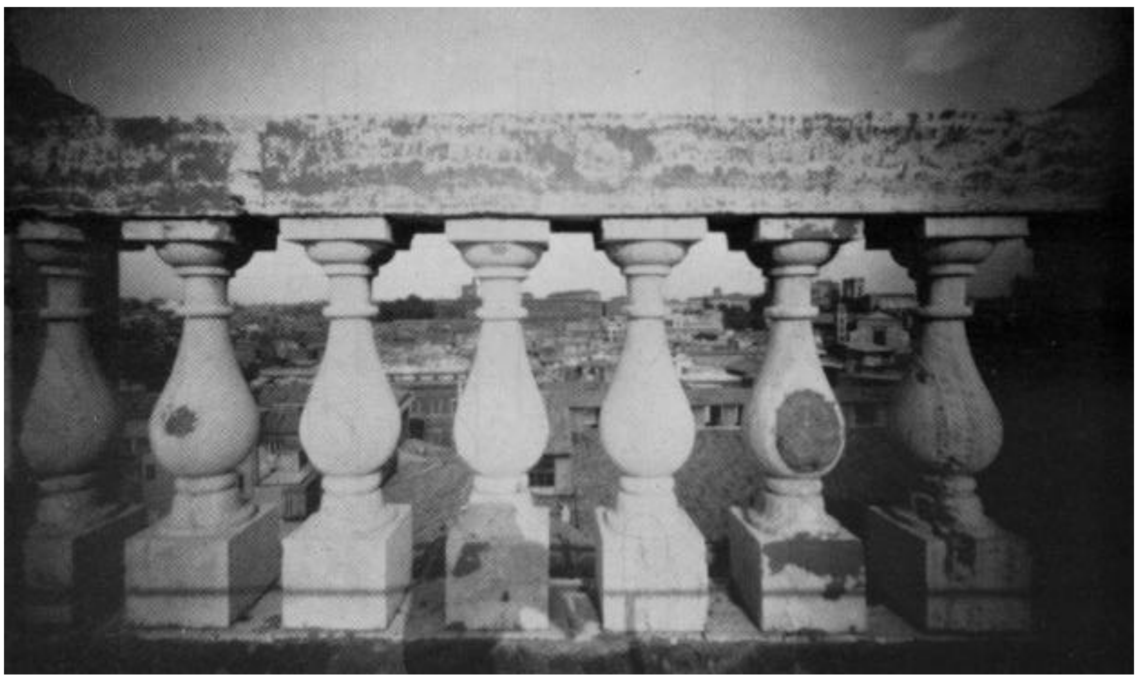
Projection of 3D shapes

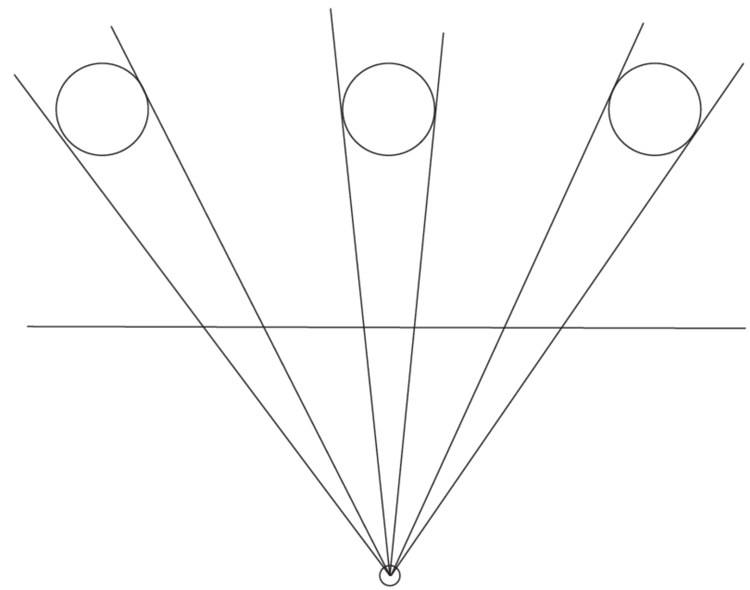
Figure 4. Details from Dürer's "Samson killing the lion". (Bartsch #2; the print dates from 1498.)

J. Koenderink. What does the occluding contour tell us about solid shape? Perception 13 (321-330), 1984

Projection of 3D shapes

- Are the widths of the projected columns equal?
  - The exterior columns are wider
  - This is not an optical illusion, and is not due to lens flaws
  - Phenomenon pointed out by Leonardo Da Vinci





Source: F. Durand

Perspective distortion: People

Overview

- Motivation: recovery of 3D structure
- Pinhole projection model
- Properties of projection
- Perspective projection matrix
- Orthographic projection

Homogeneous coordinates

- Add an extra coordinate and use an equivalence relation
- for 2D
  - equivalence relation
    k*(X,Y,Z) is the same as (X,Y,Z)
- for 3D
  - equivalence relation
    k*(X,Y,Z,T) is the same as (X,Y,Z,T)
- "Ordinary" or "non-homogeneous" coordinates
  - properly called affine coordinates
  - in 3D, affine -> homogeneous $(x, y, z) \to k * (x, y, z, 1)$
  - in 3D, homogeneous to affine $(X, Y, Z, T) \to \left(\frac{X}{T}, \frac{Y}{T}, \frac{Z}{T}\right)$

Homogenous coordinates

- Notice (0, 0, 0, 0) is meaningless (HC's for 3D)
  - also (0, 0, 0) in 2D
- Basic notion
  - Possible to represent points "at infinity" by careful use of zero
  - Where parallel lines intersect
    - eg

$$(tX, tY, tZ, 1) \quad \text{and} \quad (tX + a, tY + b, tZ + c, 1)$$

$$\text{intersect at} \quad (X, Y, Z, 0)$$

  - Where parallel planes intersect (etc)
- Can write the action of a perspective camera as a matrix

Homogeneous coordinates

Example: 23.1 *Lines on the affine plane*

Lines on the affine plane form one important example of homogeneous coordinates. A line is the set of points (x, y) where $ax+by+c=0$. We can use the coordinates (a, b, c) to represent a line. If $(d, e, f) = \lambda(a, b, c)$ for $\lambda \neq 0$ (which is the same as $(d, e, f) \equiv (a, b, c)$), then (d, e, f) and (a, b, c) represent the same line. This means the coordinates we are using for lines are homogeneous coordinates, and the family of lines in the affine plane is a projective plane. Notice that encoding lines using affine coordinates must leave out some lines. For example, if we insist on using $(u, v, 1) = (a/c, b/c, 1)$ to represent lines, the corresponding equation of the line would be $ux + vy + 1 = 0$. But no such line can pass through the origin – our representation has left out every line through the origin.

Homogeneous coordinates for a line

23.1.2 The projective line

In homogenous coordinates, we represent a point on a 1D space with two coordinates, so (X_1, X_2) (by convention, homogeneous coordinates are written with capital letters). Two sets of homogeneous coordinates (U_1, U_2) and (V_1, V_2) represent different points if there is no $\lambda \neq 0$ such that $\lambda(U_1, U_2) = (V_1, V_2)$. The set of all distinct points is known as a *projective line*. You should think of the projective line as an ordinary line (an *affine line*) with an "extra point". Every point on an affine line has a corresponding point on a projective line. A point on an affine line is given by a single coordinate x. This point can be identified with the point on a projective line given by $(X_1, X_2) = \lambda(x, 1)$ (for $\lambda \neq 0$) in homogeneous coordinates. The extra point has coordinates $(X_1, 0)$. These are the homogeneous coordinates of a single point (check this), but this point would be "at infinity" on the affine line.

There isn't anything special about the point on the projective line given by $(X_1, 0)$. You can see this by identifying the point x on the affine line with $(X_1, X_2) = \lambda(1, x)$ (for $\lambda \neq 0$). Now $(X_1, 0)$ is a point like any other, and $(0, X_2)$ is "at infinity". A little work establishes that there is a 1-1 mapping between the projective line and a circle (exercises).

You can see the point at infinity

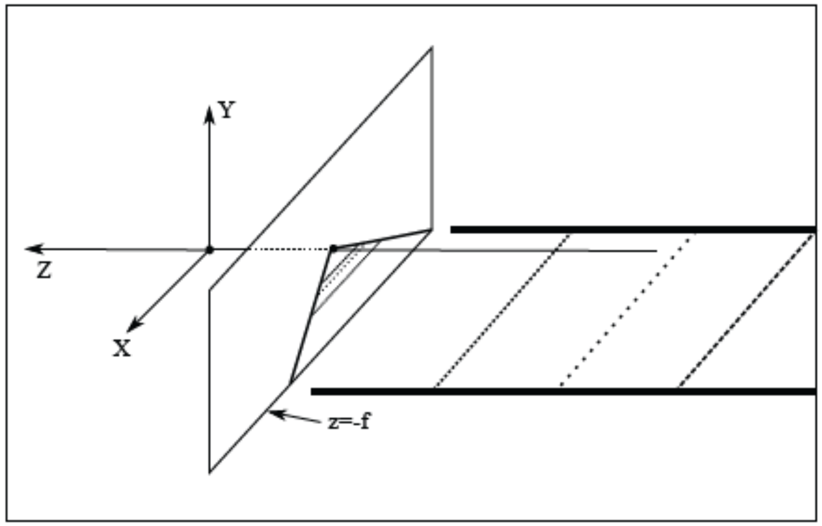


FIGURE 23.1: *The point at infinity is not just an abstraction: you can see it. Recall that lines that are parallel in the world can intersect in the image at a vanishing point. This vanishing point is the image of the point "at infinity" on the parallel lines. For example, on the plane* $y = -1$ *in the camera coordinate system, draw two lines* $(1, -1, t)$ *and* $(-1, -1, t)$ *(the lines shown in the figure). Now these lines project to* $(f1/t, f(-1/t), f)$ *and* $(f(-1/t), f(-1/t), f)$ *on the image plane, and their vanishing point is* $(0, 0, f)$*. This vanishing point occurs when the parameter* t *reaches infinity — it is the image of the point at infinity.*

Homogeneous coordinates for the plane

23.1.3 The projective plane

The space represented by three homogeneous coordinates is known as a *projective plane*. You can map an *affine plane* (the usual plane, with coordinates x, y) to a projective plane by writing $(X_1, X_2, X_3) = (x, y, 1)$. Notice that there are points on the projective plane — the points where $X_3 = 0$ — that are missing. These points form a projective line (check this!). This line is often referred to as the *line at infinity*.

You can see the line at infinity, too!

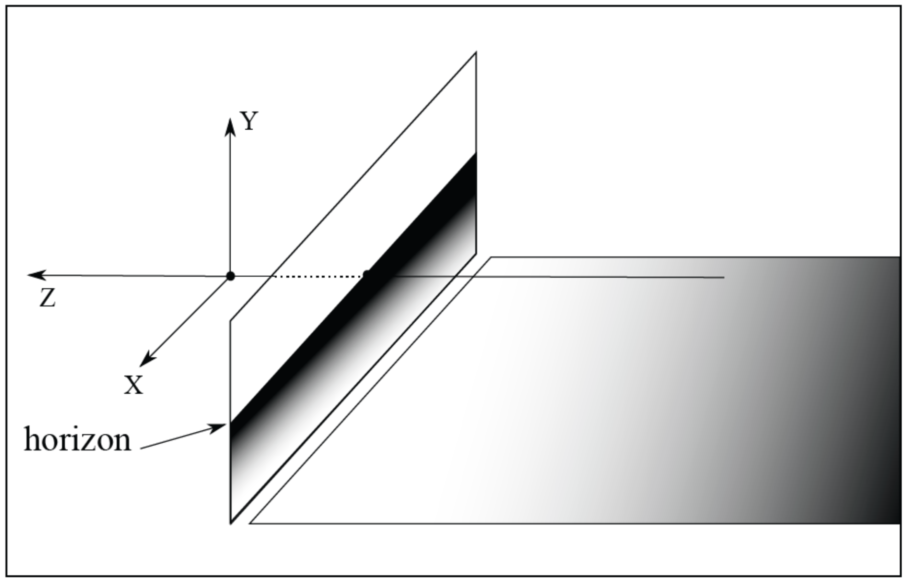
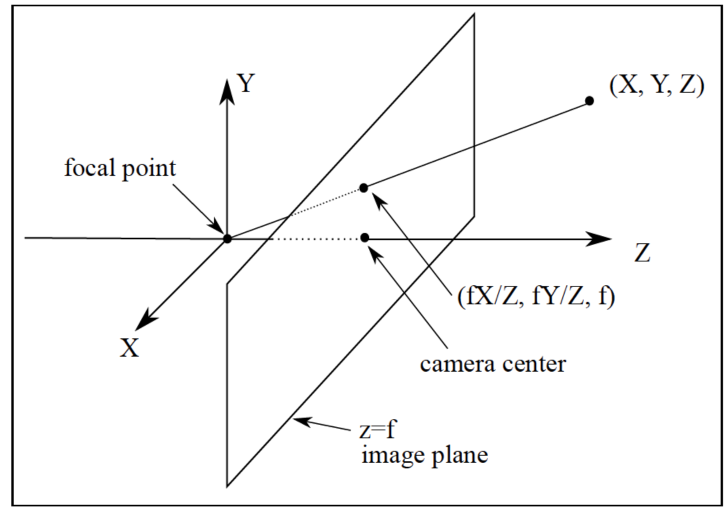


FIGURE 23.2: *The line at infinity is not just an abstraction: you can see it. Recall that, viewed in a perspective camera, planes have a horizon. This is the image of the line at infinity. For example, the plane $y = -1$ in the camera coordinate system has a horizon in the image as shown in the figure. The points on this plane can be written $(x, -1, z)$, and project to $(fx/z, -f/z, f)$. When z reaches infinity, we see the line $y = 0$ in the image plane. This is the image of the line at infinity.*

The camera matrix



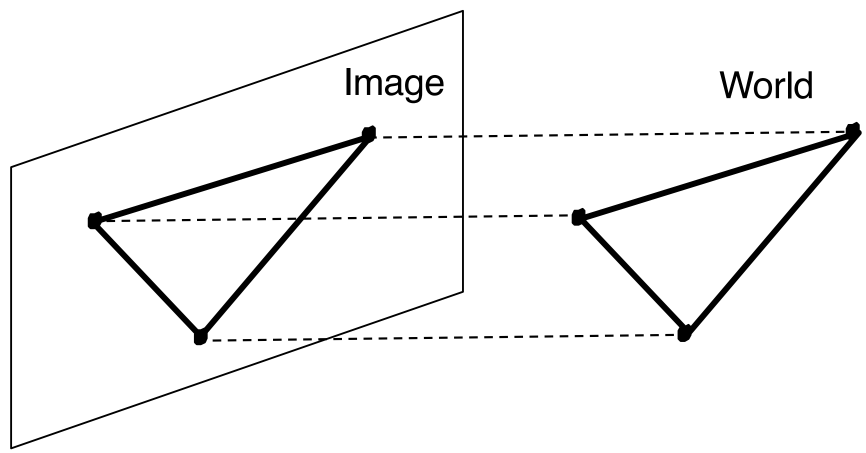
- Turn previous expression into HC's
  - HC's for 3D point are (X,Y,Z,T)
  - HC's for point in image are (U,V,W)

In Camera coordinate system:

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix}$$

Orthographic projection

- Special case of perspective projection
  - Distance from center of projection to image plane is infinite
  - Also called “parallel projection”



Slide by Steve Seitz

Orthographic projection

- Special case of perspective projection
  - Distance from center of projection to image plane is infinite
  - Also called “parallel projection”

Orthographic projection

19.1.3 Scaled Orthographic Projection and Orthographic Projection

Under some circumstances, perspective projection can be simplified. Assume the camera views a set of points which are close to one another compared with the distance to the camera. Write $\mathbf{X}_i = (X_i, Y_i, Z_i)$ for the i'th point, and assume that $Z_i = Z(1+\epsilon_i)$, where ϵ_i is quite small. In this case, the distance to the set of points is much larger than the *relief* of the points, which is the distance from nearest to furthest point. The i'th point projects to $(fX_i/Z_i, fY_i/Z_i)$, which is approximately $(f(X_i/Z)(1-\epsilon_i), f(Y_i/Z)(1-\epsilon_i))$. Ignoring ϵ_i because it is small, we have the projection model

$$(X, Y, Z) \rightarrow (f/Z)(X, Y) = s(X, Y).$$

Orthographic projection and people

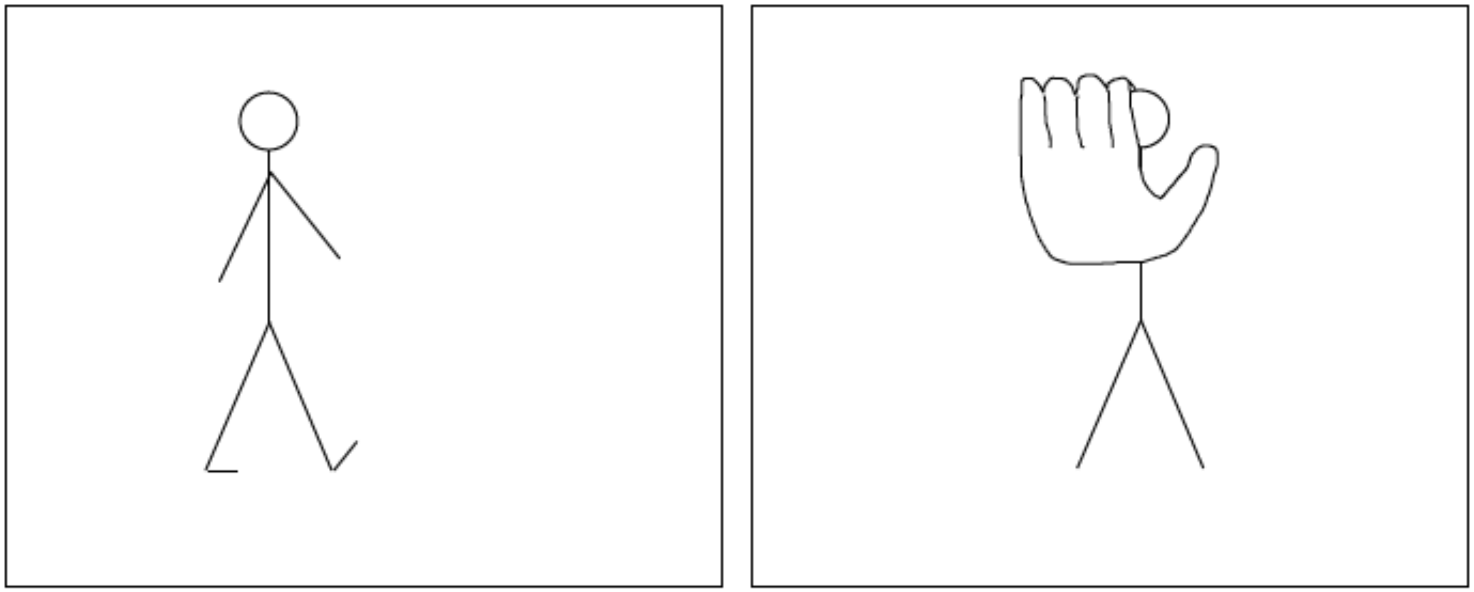


FIGURE 19.5: *The pedestrian on the* **left** *is viewed from some way away, so the distance to the pedestrian is much larger than the change in depth over the pedestrian. In this case, which is quite common for views of people, scaled orthography will apply. The celebrity on the* **right** *is holding a hand up to prevent the camera viewing their face; the hand is quite close to the camera, and the body is an armslength away. In this case, perspective effects are strong. The hand looks big because it is close, and the head looks small because it is far.*

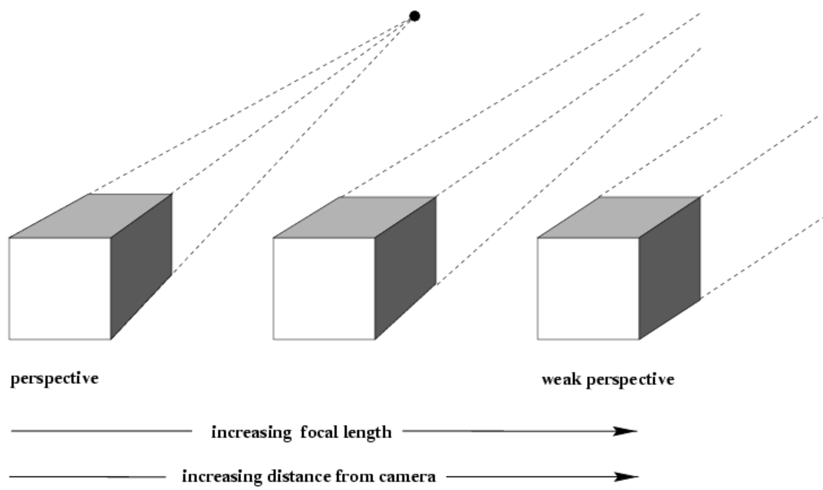
Camera matrix for orthographic projection

- Almost never encounter orthographic projection

> **Remember this:** *Scaled orthographic projection maps*
>
> $$(X, Y, Z) \rightarrow s(X, Y)$$
>
> *where s is some scale. The model applies when the distance to the points being viewed is much greater than their relief. Many views of people have this property.*

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \mathcal{C} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \mathcal{W} \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix}$$

Approximating an orthographic camera



center at infinity

Source: Hartley & Zisserman