Image Formation: Pinhole Model, Perspective Projection, and Binocular Stereo

Lecture by Margrit Betke, CS 585, February 6, 2024
Pinhole camera

Image credit: public domain, Wikipedia
Pinhole camera

= Box with a hole

Image is projected upside down on side opposite to hole.
Pinhole camera

Why is a hole needed?

With an open box and no hole, light rays from everywhere go everywhere…
Pinhole camera

Why is a hole needed?

With an open box and no hole, light rays from everywhere go everywhere with the result that the image is completely washed out and the box side stays white.
Pinhole camera

Ok, so the hole is crucial.

But how big a hole?
Pinhole camera

Larger hole: Blurrier image
Pinhole camera

Larger hole:
Blurrier image

Why?
Light rays from nearby scene points also end up in the same spot on the image plane
Pinhole camera

Larger hole:
Blurrier image

Why?
Light rays from nearby scene points also end up in the same spot on the image plane

Technical term of hole: Aperture
Small Aperture                                          Large Aperture
© www.scratchapixel.com                              © www.scratchapixel.com
Pinhole camera

Smaller aperture

Fainter image

because less light travels through pinhole
Pinhole camera

Trade off:

Small aperture: Faint image but less blurry

Large aperture: Bright image but blurry
Watch Steve Seitz’ Video:

https://www.youtube.com/watch?v=F5WA26W4JaM&list=PLWfDJ5nl8UpwShx-lzLJqcp575fKpsSO&index=11

until 2:58
Pinhole camera  = Camera obscura

Camera means ‘room’ and obscura means ‘dark’ in Latin.

Historic descriptions in
- Chinese Mozi writings (~ 500 BCE),
- Aristotelian Problems (~ 300 BCE),
- Arab writings (~1000 CE).

Image credit: James Ayscough, 1755
Pinhole camera = Camera obscura

Problem: Real-time images cannot be stored!
Development of Camera Obscura to Modern Camera

The first permanent photoetching was an image produced in 1822 by the French inventor Nicéphore Niépce.

Image credit: Jonnychiwa, Wikipedia
Development of Camera Obscura to Modern Camera

Film as a storage medium:
The first flexible photographic roll film was marketed by George Eastman, founder of Kodak in 1885.

Array of linked capacitors as storage medium:
Sony unveiled the first consumer camera (Mavica) to use a charge-coupled device (CCD) for imaging, eliminating the need for film, in 1981.
Modern Film or Digital Cameras

Problem solved: Real-time images can be stored!
So then, why do we care about pinhole cameras in CS 585?
Poll: So then, why do we care about pinhole cameras in CS 585?

1. Historic reason: We need to learn how computer vision started as a research field.

2. Mathematical reason: Real cameras have complicated lens systems, not pin holes. We can simplify the geometry of image formation mathematically by ignoring the lenses.

3. Computational reason: Modern cameras post-process the projected images as if they were collected by a pinhole camera.
Ideal Pinhole Camera Model: View from Top

Coordinate System Origin = Center of Projection (CoP) = Pinhole
Ideal Pinhole Camera Model: View from Top

Optical axis = Z axis
perpendicular to image plane

Coordinate System Origin = Center of Projection (CoP) = Pinhole

image plane x
Ideal Pinhole Camera Model: View from Top

- **Optical axis** = Z axis
- **Coordinate System Origin** = Center of Projection (CoP) = Pinhole
- **focal length** f
- **image plane** X
- **scene**
Ideal Pinhole Camera Model: View from Top

Depth = Distance Z

Coordinate System Origin = Center of Projection (CoP) = Pinhole

focal length f

image plane X

scene
Pinhole Model

Coordinate System Origin = Center of Projection (CoP) = Pinhole

light ray from scene to image plane through pinhole

depth $Z$

focal length $f$

image plane
We can relate the scene and image plane coordinates $X$ and $x$ using a perspective projection equation.
Derivation of the Perspective Projection Equation

Similar triangles have same angles
Derivation of the Perspective Projection Equation

Coordinate System Origin = Center of Projection (CoP) = Pinhole

focal length f

depth Z

image plane

light ray from scene to image plane through pinhole

3D scene point

Derivation of the Perspective Projection Equation
Derivation of the Perspective Projection Equation

Perspective Projection Equation:

\[ \frac{x}{Z} = \frac{x}{f} \]
Perspective Projection Equation:
\[ \frac{x}{Z} = \frac{x}{f} \]

f is given by the camera, x is our image measurement.
Usefulness of the Perspective Projection Equation

Perspective Projection Equation:
\[ \frac{X}{Z} = \frac{x}{f} \]

- \( f \) is given by the camera, \( x \) is our image measurement.

If we know the object depth \( Z \), we can compute \( X \):
\[ X = \frac{xZ}{f} \]

If we know the object width \( X \), we can compute \( Z \):
\[ Z = \frac{xf}{x} \]
Perspective Projection Equation:

\[ \frac{x}{Z} = \frac{x}{f} \]

- \(f\) is given by the camera, \(x\) is our image measurement

If we do not know the object depth \(Z\) and object width \(X\), we need another equation!

=> Use a Binocular Stereo System
Binocular Stereo

Special Considerations:

1) It is convenient to use the same type of camera on the left and the right
   ⇒ Same focal length $f$

2) Place the optical axes parallel to each other

3) Align the cameras so that the left and right images are in the same plane
   (here viewed in cross section)
Binocular Stereo

In a Monocular System: Coordinate System Origin = Center of Projection (CoP) = Pinhole
Binocular Stereo

In a Monocular System: Coordinate System Origin = Center of Projection (CoP) = Pinhole
In a Binocular System: Coordinate System Origin in the middle between CoPs
Binocular Stereo

In a Monocular System: Coordinate System Origin = Center of Projection (CoP) = Pinhole
In a Binocular System: Coordinate System Origin in the middle between CoPs
Binocular Stereo

- Depth $Z$
- CoP$_{left}$
- CoP$_{right}$
- Light rays from scene to image planes

3D scene point

- CoP left
- CoP right
- Light rays from scene to image planes
- Depth $Z$
Binocular Stereo

Baseline $b = \text{Distance between Center of Projections (CoPs)}$

Depth $Z$

Coordinate System Origin

Left image plane $x_{\text{left}}$

Right image plane $x_{\text{right}}$

Light rays from scene to image planes

$\mathbf{x}$

$\mathbf{f}$
Binocular Stereo

Coordinate System Origin

3D scene point

light rays from scene to image planes

depth Z

Baseline \( b = \) Distance between Center of Projections (CoPs)

CoP \(_{\text{left}}\)

CoP \(_{\text{right}}\)

\( f \)

left image plane

right image plane

left image plane

right image plane

\( x_{\text{left}} \)

\( x_{\text{right}} \)

© Betke
Binocular Stereo

Projection Equation:
\[ \frac{x_{\text{left}}}{f} = (\text{something})/Z \]

Baseline \( b = \text{Distance between Center of Projections (CoPs)} \)
Binocular Stereo

Coordinate System
Origin

Baseline $b =$ Distance between Center of Projections (CoPs)

$$x_{right}/f = (\text{something else})/Z$$
Binocular Stereo

Baselines $b$

Distance between Center of Projections (CoPs)

Coordinate System Origin

3D scene point

Depth $Z$

Left image plane

Right image plane

$X_{\text{left}}$

$X_{\text{right}}$

$X$

$f$

$b/2$

$X_{\text{CoP left}}$

$X_{\text{CoP right}}$

© Betke
Binocular Stereo

Projection Equations:
\[ \frac{x_{\text{left}}}{f} = \frac{X + b/2}{Z} \]
\[ \frac{x_{\text{right}}}{f} = \frac{X - b/2}{Z} \]

Baseline \( b \) = Distance between Center of Projections (CoPs)
Binocular Stereo

Projection Equations:

\[
\frac{x_{\text{left}}}{f} = \frac{X + b/2}{Z}
\]

\[
\frac{x_{\text{right}}}{f} = \frac{X - b/2}{Z}
\]
Binocular Stereo

Projection Equations:

\[
\frac{x_{\text{left}}}{f} = \frac{X + b/2}{Z} \\
\frac{x_{\text{right}}}{f} = \frac{X - b/2}{Z}
\]

Subtract the 2nd equation from the 1st equation:

\[
\frac{x_{\text{left}} - x_{\text{right}}}{f} = \frac{(X + b/2 - X + b/2)}{Z}
\]
Binocular Stereo

Coordinate System

Origin

Baseline $b = \text{Distance between Center of Projections (CoPs)}$

Depth $Z$

CoP left

CoP right

Projection Equations:

\[
x_{\text{left}}/f = (X+b/2)/Z \\
x_{\text{right}}/f = (X-b/2)/Z
\]

Subtract the 2\text{nd} equation from the 1\text{st} equation:

\[
(x_{\text{left}} - x_{\text{right}})/f = (X+b/2-X+b/2)/Z
\]

which results in:

\[
(x_{\text{left}} - x_{\text{right}})/f = b/Z
\]
Binocular Stereo

Projection Equations:

\[
\frac{x_{\text{left}}}{f} = \frac{X+b/2}{Z} \quad \frac{x_{\text{right}}}{f} = \frac{X-b/2}{Z}
\]

Subtract the 2nd equation from the 1st equation:

\[
\frac{x_{\text{left}} - x_{\text{right}}}{f} = \frac{(X+b/2-X+b/2)}{Z}
\]

which results in:

\[
Z = \frac{bf}{x_{\text{left}} - x_{\text{right}}}
\]
Binocular Stereo

Projection Equations:
\[
x_{\text{left}}/f = (X+b/2)/Z \\
x_{\text{right}}/f = (X-b/2)/Z
\]

Subtract the 2\textsuperscript{nd} equation from the 1\textsuperscript{st} equation:
\[
(x_{\text{left}} - x_{\text{right}})/f = (X+b/2-X+b/2))/Z
\]
which results in:
\[
(x_{\text{left}} - x_{\text{right}})/f = b/Z \quad \text{or:}
\]
\[
Z = bf/ (x_{\text{left}} - x_{\text{right}})
\]
Binocular Stereo

Projection Equations:
\[
\frac{x_{\text{left}}}{f} = \frac{X+b/2}{Z} \quad \frac{x_{\text{right}}}{f} = \frac{X-b/2}{Z}
\]
which results in:
\[
\frac{(x_{\text{left}} - x_{\text{right}})}{f} = \frac{b}{Z} \quad \text{or:} \quad Z = \frac{bf}{(x_{\text{left}} - x_{\text{right}})}
\]

Disparity \( \delta \)
Binocular Stereo

Key Equation:

\[ Z = \frac{b f}{x_{\text{left}} - x_{\text{right}}} \]

or

\[ Z = \frac{b f}{\delta} \]

- **Z** = Unknown depth
- **b** = selected baseline
- **f** = focal length of camera
- **\( \delta \)** = disparity of imaged scene point

Baseline **b** = Distance between Center of Projections (CoPs)
What happens with the images when we make the baseline smaller?

Original baseline: ________________________________

Smaller baseline: ________________________________
Poll: What happens with the images when we make the baseline **smaller**?

The disparity

1. increases
2. stays the same
3. decreases
Making the Baseline Smaller Reduces the Disparity
What happens when we increase the depth = distance between scene and image plane?
What happens with the images when we increase the depth?
What happens with the images when we increase the depth?

The disparity

1. increases
2. stays the same
3. decreases
Increasing the depth makes the disparity smaller.

\[ \delta_{\text{near } Z} > \delta_{\text{far } Z} \]
Summary of Concepts: Binocular Stereo

• Today we considered a special case:
  • parallel optical axes
  • image planes aligned
  • same focal length

• Combining perspective projection equations for both cameras yields formula $Z = \frac{bf}{\delta}$

• We discussed how disparity $\delta$ changes with changes in $b$ or $Z$
Back to the Single Camera Pinhole Model

Coordinate System Origin = Center of Projection (CoP) = Pinhole

Projection Equation:
\[
x/Z = x/f
\]
Placing Image Plane in Front of Pinhole

Projection Equation:

\[
\frac{x}{Z} = \frac{x}{f}
\]
Placing Image Plane in Front of Pinhole

Projection Equation: \( \frac{x}{Z} = \frac{x}{f} \)

This is done for mathematical convenience: Image \( x \) and scene \( X \) are measured in the same direction on the x-axis (a positive \( x \) means a positive \( X \)). The focal length is often set to \( f=1 \).
Back to the Original Single Camera Pinhole Model

Projection Equation:
\[
x/Z = x/f
\]
Let’s rotate the camera counterclockwise:
Vertical Dimension Included:

3D scene point

Optical Axis = $Z$ axis

CoP

$(0,0)$(x,y)

image plane

© Betke
Vertical Dimension Included:

3D scene point \((X,Y,Z)\)

Optical Axis = \(Z\) axis

depth

Optical Axis = \(Z\) axis

CoP \((0,0,0)\)

Principal Point = intersection of optical axis with image plane

image plane

© Betke
Vertical Dimension Included:

3D scene point

Optical Axis = Z axis

CoP

(0,0,0)

f

image plane

Perspective Projection Equations:

\[ \frac{X}{Z} = \frac{x}{f} \]

and

\[ \frac{Y}{Z} = \frac{y}{f} \]

relate scene \((X,Y,Z)\) and image \((x,y)\) coordinates
Why are perspective projection equations important for computer vision and image understanding?

Perspective Projection Equations:

\[ \frac{X}{Z} = \frac{x}{f} \]

and

\[ \frac{Y}{Z} = \frac{y}{f} \]

relate scene \((X,Y,Z)\) and image coordinates \((x,y)\)
Poll: Why are perspective projection equations important for computer vision and image understanding?

1. Mathematical reason: We need to learn about the mathematical underpinnings of computer vision.
2. Historical reason: All computer vision students need to learn about these fundamental equations.
3. Computational reason: We can process measurements in the image to infer information about the scene.
Why are perspective projection equations important for computer vision and image understanding?

Perspective Projection Equations:

\[
\frac{X}{Z} = \frac{x}{f} \quad \text{and} \quad \frac{Y}{Z} = \frac{y}{f}
\]

Use measured image coordinates (x,y) to interpret the scene (X,Y,Z)
Example: Self-driving Cars:

Estimate distance to car in front

Distance in meters: \( Z = c_{\text{horiz}} \frac{f W}{w} = (22 \text{ pix/mm})(50 \text{ mm})(1.77 \text{ m})/(100 \text{ pix}) \)
\[ = 19.47 \text{ m} \]

- Typical width of a car: \( W = 1.77 \text{ m} \)
- Car width measured in image: \( w = 100 \text{ pixel} \)
- focal length \( f = 50 \text{ mm} \)
- 35-mm camera: pixel-to-mm conversion \( c_{\text{horiz}} = 22 \text{ pixel/mm} \)
Why are perspective projection equations important for computer vision and image understanding?

Perspective Projection Equations:

\[
\frac{X}{Z} = \frac{x}{f} \\
\frac{Y}{Z} = \frac{y}{f}
\]

In the self-driving car example:

Use measured image coordinates \((w, y)\) of car in front to interpret its distance: \((W, Y, Z)\)
Orthographic Projection

Alternative to perspective projection when imaged objects are far away

- **Perspective:**
  - Light rays assumed parallel to optical axis
  - Large focal length $f$
  - Image plane in front of CoP

- **Orthographic:**
  - Light rays assumed parallel to optical axis
Orthographic Projection

Alternative to perspective projection when imaged objects are far away

used when distances of points in the scene do not differ much, i.e., far away building
Orthographic Projection

Assume all Z’s are approximately at fixed distance $Z_0$ and \[ \frac{|\Delta Z|}{Z_0} \ll 1 \]
i.e., distances of points in the scene do not differ much

Then $x = (f/Z_0)X$ and $y = (f/Z_0)Y$.

Or assume $x = X$ and $y = Y$. $\Rightarrow$ Simplifies image analysis to

2D problem.
Vanishing Point

Definition:

Point at which receding parallel lines viewed in perspective appear to converge

Photo by Wikipedia user: MikKBDFJGEmalak
Vanishing Point

Definition:

Point at which receding parallel lines viewed in perspective appear to converge.

Photo by Wikipedia user: MikKBDFJKGeMalak
Vanishing Point

Definition:
Point at which receding parallel lines viewed in perspective appear to converge

Photo by Wikipedia user: MikKBDFJKGeMalak
Why do parallel lines intersect when projected?
Why do parallel lines intersect when projected?

Illustration by Wikipedia user: Mgunyho
One Vanishing Point
Learning Objectives

Be able to explain:

• Pinhole model, camera obscura, center of projection, aperture, principal point, optical axis, focal length, depth, perspective projection, disparity
• What happens to an image if the aperture is increased or decreased?
• What is the impact of changing the baseline or the distance to the scene on the images in a binocular camera system?
• What is the difference between the perspective and orthographic projection models and when should you use them?
• How could you use a perspective projection equation to estimate the distance of a car in front?
• What is a vanishing point? Why could it be useful for highway scene analysis?