Image Formation

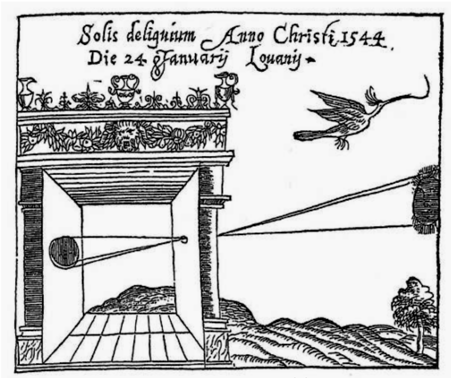
- Light
- Reflectance
- Image Capture
  - Camera
  - Lens
  - Sensor
  - Projection models
  - Camera System Parameters

Slide credits: D Forsythe, I. Kokkinos. S. Lazebnic, S. Seitz, J. Hays

Cameras - a brief history

- From the latin *Camera Obscura* - Dark Chamber
- The cameras of the 16th century were literally dark rooms with a small opening on one side
- The first cameras were used by painters who would literally sit in the box and trace the outline of the image on their canvases

Camera obscura



Gemma Frisius, 1558

- Basic principle known to Mozi (470-390 BCE), Aristotle (384-322 BCE)
- Drawing aid for artists: described by Leonardo da Vinci (1452-1519)

Source: A. Efros

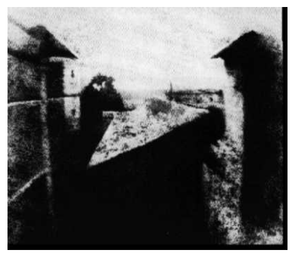
The invention of the lens

- The first major improvement occurred in the 16th century when Giovanni Battista della Porta added a lens to the design
- The lens improved the light gathering power of the device so brighter images were possible

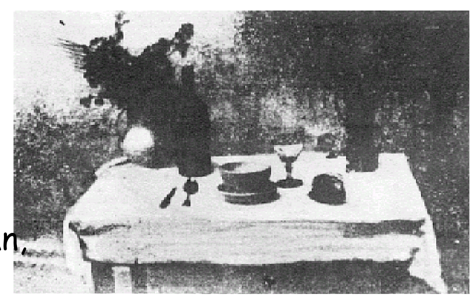
Photography

- Film
  - The first film was developed by Niepce (1822) who took the worlds first photograph of a farmyard in central France – it required an exposure time of eight hours.
  - He joined forces with Louis Jacques Mande Daguerre who improved the process. His exposures only required half an hour and were quite sharp.

The Worlds First Known Photo?

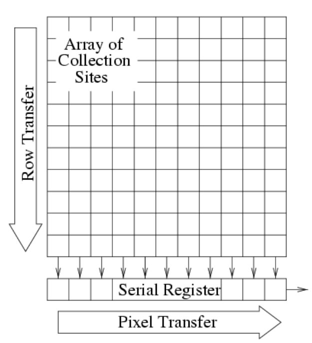


Photographs (Niepce, "La Table Servie," 1822)



Milestones:
Daguerreotypes (1839)
Photographic Film (Eastman, 1889)
Cinema (Lumière Brothers, 1895)
Color Photography (Lumière Brothers, 1908)
Television (Baird, Farnsworth, Zworykin, 1920s)

CCD Devices (1970)



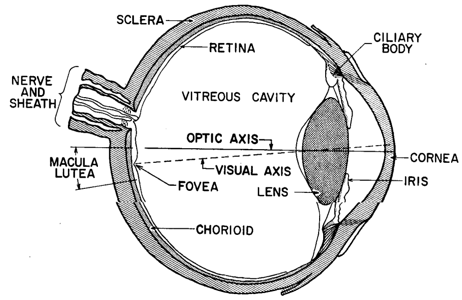
The anatomy of a modern camera

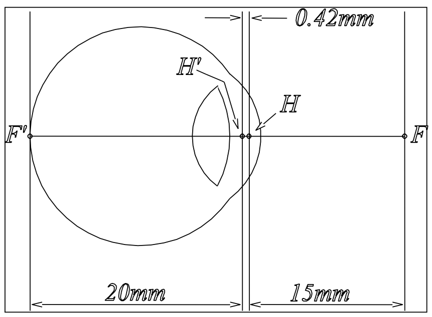
- Lens
- Shutter (exposure time)
- Diaphragm (aperture)
- Focusing Control

The Camera of the Mind

- Components of the human eye
  - Pupil
  - Lens
  - Occiliary muscle
  - retina
  - fovea
  - blind spot

The Human Eye





Helmoltz's Schematic Eye

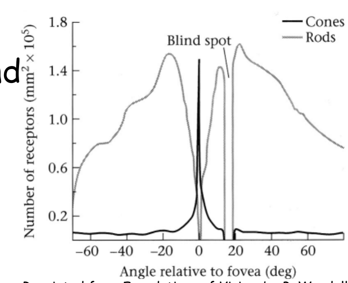
The Retina

- There are two types of photosensitive cells in the retina, rods and cones
- Cones come in three flavors which exhibit different sensitivities to different wavelengths of light, red green and blue.
- Rods are not sensitive to variations in wavelength but they are more sensitive than cones and can pick up much dimmer light
- The fovea is populated entirely by cones.

More Cells

- Ganglion Cells
  - The photosensitive cells transmit their information to ganglion cells which in turn transmit information to the brain via the optic nerve
- Numbers of cells
  - There are approximately 6 million cone cells, 120 million rods and 1 million optic nerve fibers.

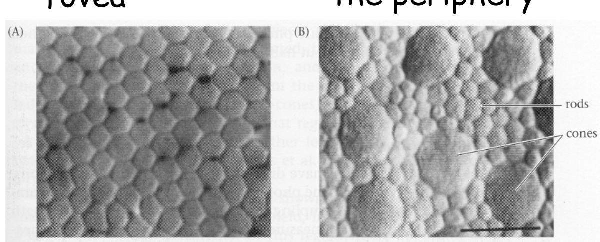
The distribution of rods and cones across the retina



Reprinted from Foundations of Vision, by B. Wandell, Sinauer Associates, Inc., (1995). © 1995 Sinauer Associates, Inc.

Cones in the fovea

Rods and cones in the periphery



Reprinted from Foundations of Vision, by B. Wandell, Sinauer Associates, Inc., (1995). © 1995 Sinauer Associates, Inc.

Image Sensors

- Images are formed by the interaction of the incident image irradiance with light sensitive elements on the image plane
- Light sensitive elements
  - Film
  - Charge Coupled Device (CCD)
  - CMOS Imaging element

Digital Images

- A digital image is an array of numbers indicating the image irradiance at various points on the image plane
- Image intensities are spatially sampled
- Intensity values are quantized (8-bits, 10-bits, 12-bits)

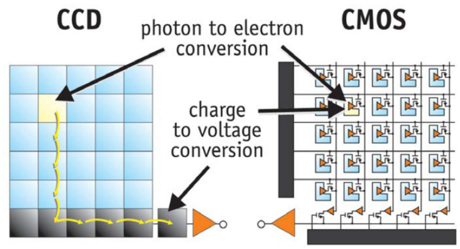
Digital Image Representations

- Acquisition: Charge-Coulped Device (CCD) arrays
- Storage: Usually in computer memory
- Display: computer hardware (video boards) and monitors

Digital B&W Cameras

- CCDs consist of a (usually) two-dimensional array of photo-sensitive cells, each corresponding to a pixel
- Light falling onto a cell's surface causes the generation of a voltage roughly proportional to intensity of incident light
- Voltage reading of each cell is converted to a digital signal
  within a CCD-specific range (usually an 8- or 10-bit number)

Digital camera



CCDs move photogenerated charge from pixel to pixel and convert it to voltage at an output node. CMOS imagers convert charge to voltage inside each pixel.

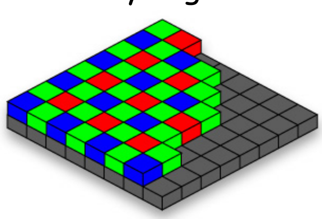
A digital camera replaces film with a sensor array

- Each cell in the array is light-sensitive diode that converts photons to electrons
- Two common types
  - **Charge Coupled Device** (CCD)
  - **Complementary metal oxide semiconductor** (CMOS)
- http://electronics.howstuffworks.com/digital-camera.htm

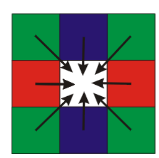
Slide by Steve Seitz

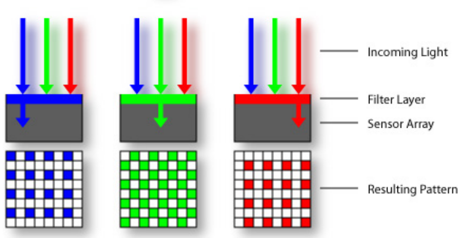
Color sensing in camera: Color filter array

Bayer grid

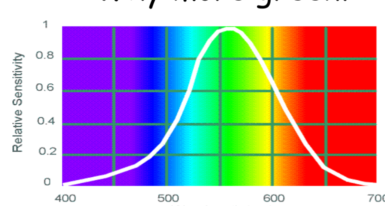


Estimate missing components from neighboring values (demosaicing)





Why more green?



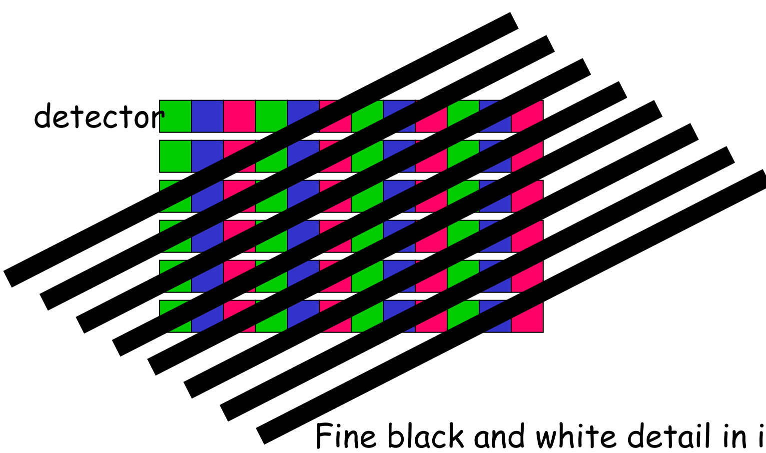
Human Luminance Sensitivity Function

Source: Steve Seitz

Problem with demosaicing: color moire

Slide by F. Durand

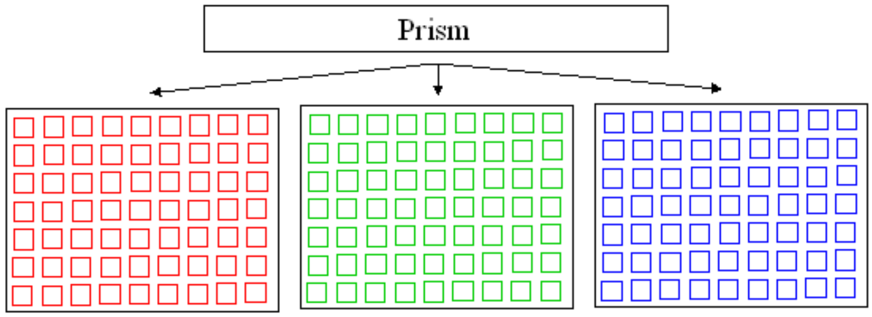
The cause of color moire



Fine black and white detail in image
misinterpreted as color information

Slide by F. Durand

Expensive Digital Color Cameras



- Three-CCD color cameras precisely aligned

Spatial Sampling

- The Image irradiance function across the image plane is sampled to obtain the digital image
- The spacing of the image elements limits the resolution of the image
- Due to the frequency content of the irradiance function is limited by the effective aperture of the camera
- Spatial aliasing

Video Images

- For a video camera, Images are taken sequentially by opening and closing the shutter 30 times per second
- Temporal aliasing
- Motion blur

Sensing Color

- In a 3 CCD video camera the light path is split into three components which are passed through colored light filters and then imaged
- As a result - a color image contains three channels of information; red, green and blue image intensities

Other Input Devices

- Mouse
- Function Buttons
- Tablet
- Joystick & Trackball
- Data Glove
- Magnetic Sensors
- Range Scanner

Image Storage

- As continuous string of bytes
- As a two-dimensional array of pixels
- Way of storage can affect performance of algorithms
- Color images can be stored using either 24 or 8 bits
- When we have more colors, we can use a look up table

Gamma Correction

- The light output of a monitor's phosphors is generally not proportional to the voltage applied
- Brightness function $I = aV^g$
- Image colors look "distorted"
- gamma correction color lookup table

Pixel Representations

- Resolution
- Quantization
- Efficiency in access

Image Resolution

Image Quantization

- **Number of bits per pixel (or color) determines number of possible values**

Image Storage Issues

- 4-bytes per pixel is more efficient than 3, because it aligns better with integer boundaries in memory.
- 4^{th} pixel is often the alpha component between 0 and 1
- Final color is [aR aG aB]
- Used to merge two pixel values

Bandwidth Requirements

- Binary
  - 1 bit * 640x480 * 30 = 9.2 Mbits/second
- Grey
  - 1 byte * 640x480 * 30 = 9.2 Mbytes/second
- Color
  - 3 bytes * 640x480 * 30 = 27.6 Mbytes/second (actually about 37 mbytes/sec)
- Typical operation:
  - 3x3 convolution: 9 multiplies + 9 adds ➔ 180 Mflops

Coordinate Systems

- Camera frame
- Image plane coordinates
- Pixel coordinates

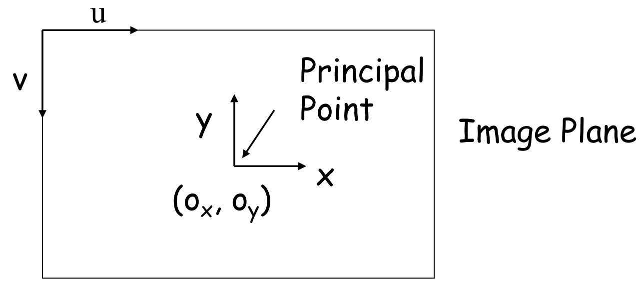


Image Plane to Pixels

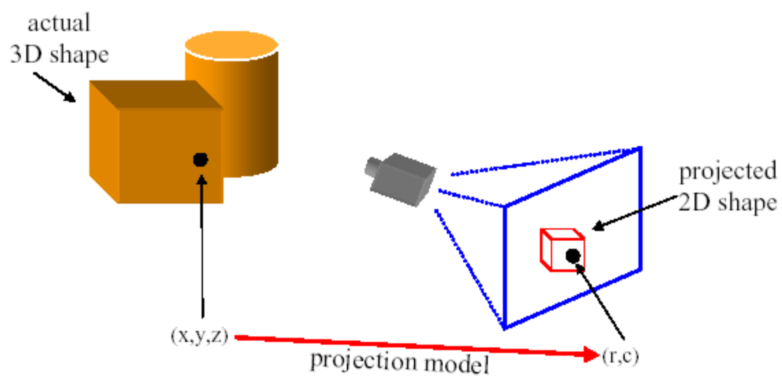
- There is an affine relationship between image plane coordinates (x,y) and pixel coordinates (u, v)

$$x = (u - o_x)s_x$$

$$y = -(v - o_y)s_y$$

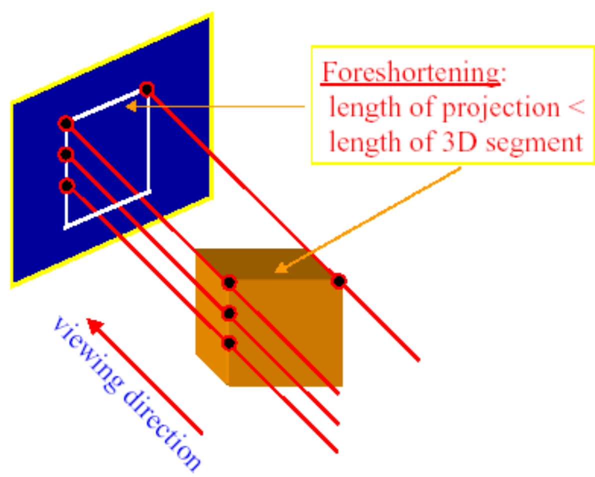
Projection Models

- Projection models are geometrical models that allow us to determine the projection of a 3D point onto a 2D photograph



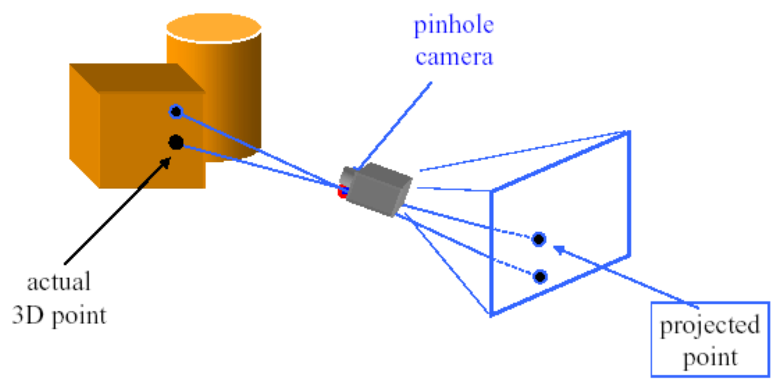
Orthographic Projection

- 3D points are projected along parallel rays onto the image plane

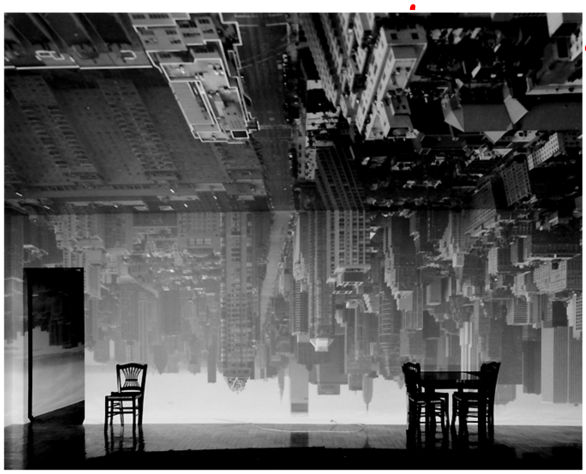


Perspective Projection

- Points are projected onto the image plane along rays that pass through a "pinhole" camera



Turning a room into a camera



After scouting rooms and reserving one for at least a day, Morell masks the windows except for the aperture. He controls three elements: the size of the hole, with a smaller one yielding a sharper but dimmer image; the length of the exposure, usually eight hours; and the distance from the hole to the surface on which the outside image falls and which he will photograph. He used 4 x 5 and 8 x 10 view cameras and lenses ranging from 75 to 150 mm.

After he's done inside, it gets harder. "I leave the room and I am constantly checking the weather, I'm hoping the maid reads my note not to come in, I'm worrying that the sun will hit the plastic masking and it will fall down, or that I didn't trigger the lens."

From *Grand Images Through a Tiny Opening*, **Photo District News**, February 2005

- Abelardo Morell, Camera Obscura Image of Manhattan View Looking South in Large Room, 1996

http://www.abelardomorell.net/camera_obscura1.html

Pinhole cameras everywhere

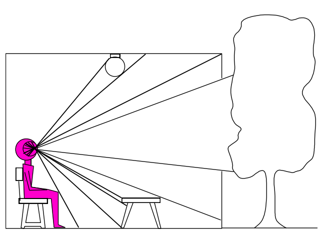
- Tree shadow during a solar eclipse
  - photo credit: Nils van der Burg

http://www.physicstogo.org/index.cfm

Slide by Steve Seitz

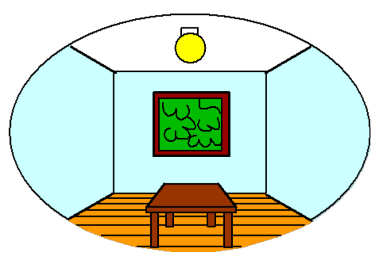
Dimensionality reduction: from 3D to 2D

3D world



Point of observation

2D image



What properties of the world are preserved?

- Straight lines, incidence

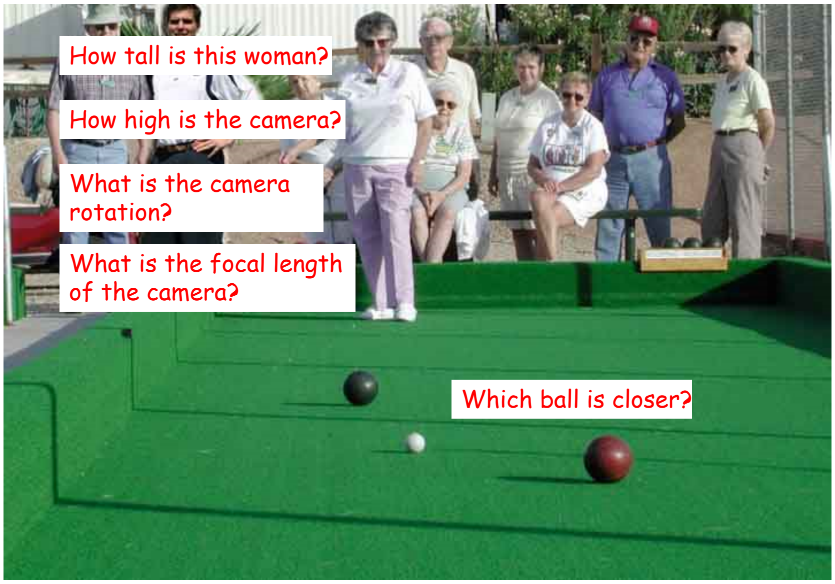
What properties are not preserved?

- Angles, lengths

Slide by A. Efros

Figures © Stephen E. Palmer, 2002

Single-view Geometry

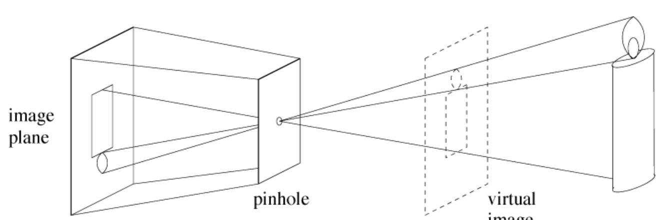


The Pinhole Camera Model

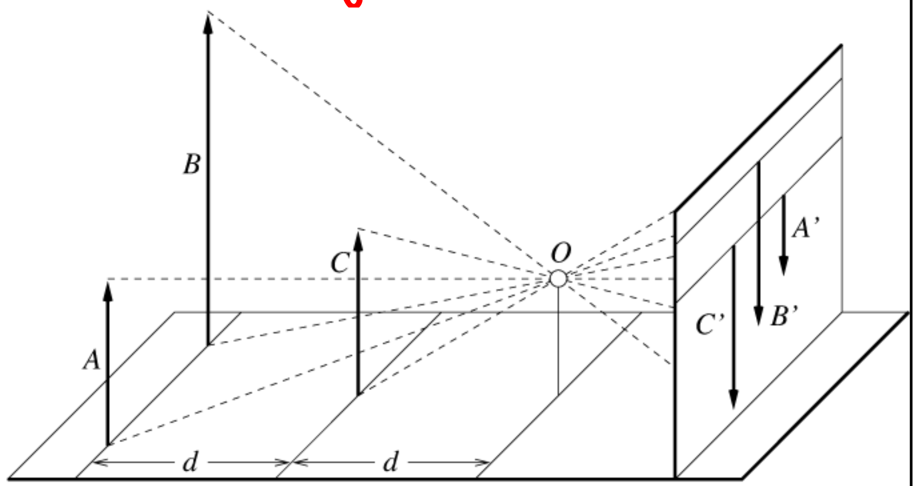
- In this model the camera is represented by a focal point, O, and an imaging plane, p

$$x = f\,\frac{X}{Z}$$

$$y = f\,\frac{Y}{Z}$$

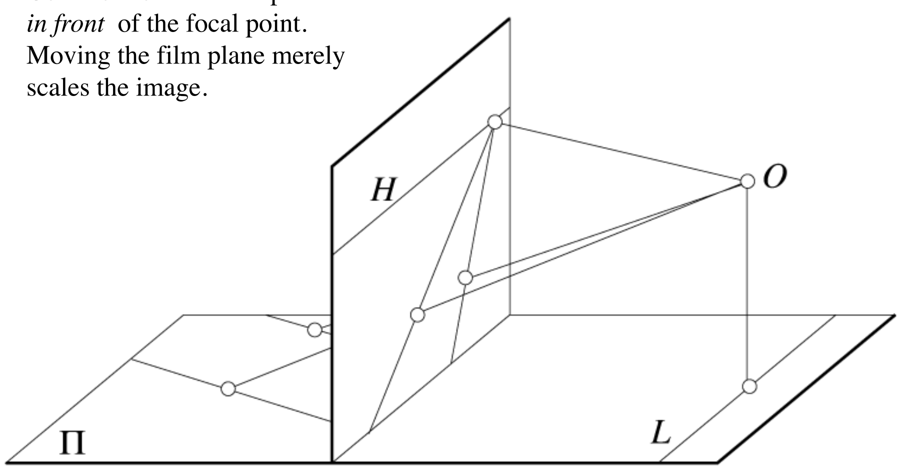


Distant objects are smaller



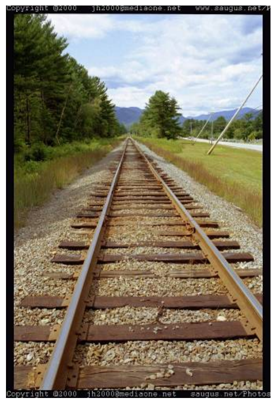
Parallel lines meet

Common to draw film plane
in front of the focal point.
Moving the film plane merely
scales the image.



Vanishing points

- *All parallel lines converge to a vanishing point*
  - Each direction in space is associated with its own vanishing point
  - Exception: directions parallel to the image plane

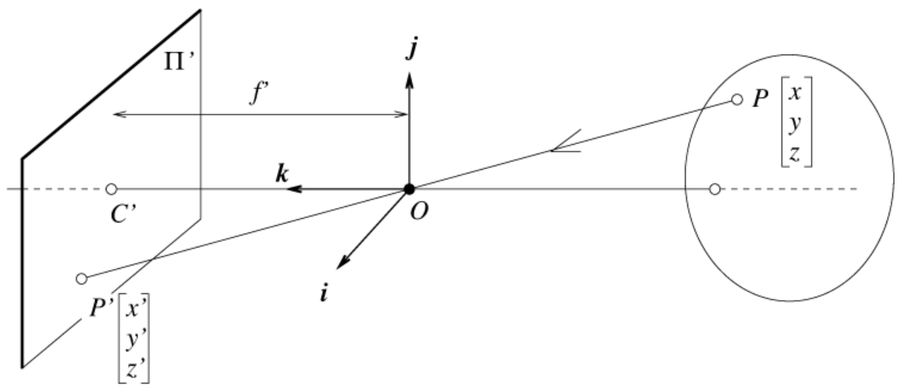


Vanishing points

- each set of parallel lines (=direction) meets at a different point
  - The *vanishing point* for this direction
- Sets of parallel lines on the same plane lead to *collinear* vanishing points.
  - The line is called the *horizon* for that plane

- Good ways to spot faked images
  - scale and perspective don't work
  - vanishing points behave badly
  - supermarket tabloids are a great source.

The equation of projection



The equation of projection

- Cartesian coordinates:
  - We have, by similar triangles, that (x, y, z) -> (f x/z, f y/z, -f)
  - Ignore the third coordinate, and get

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

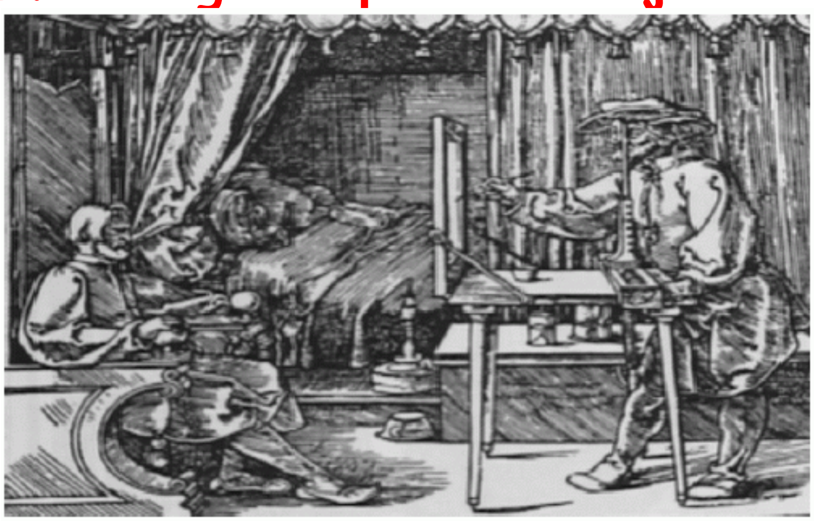
Weak Perspective Model

- Appropriate when the depth of field of the relevant objects is small compared to the distance of those objects from the center of projection
- Adv: easy
- Disadv: wrong

$$x = f\frac{X}{\bar{Z}}$$

$$y = f\frac{Y}{\bar{Z}}$$

Emulating Perspective Projection

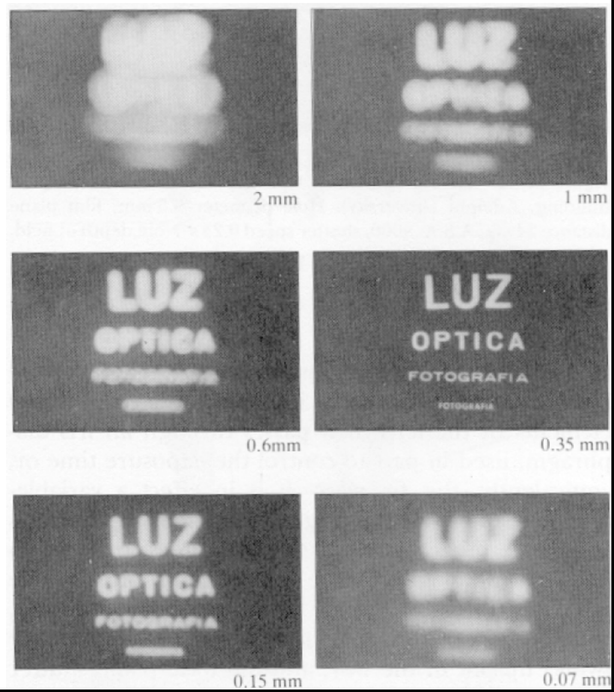


- Woodcut from Durer's "A Treatise on Measurement" (late 15th century)

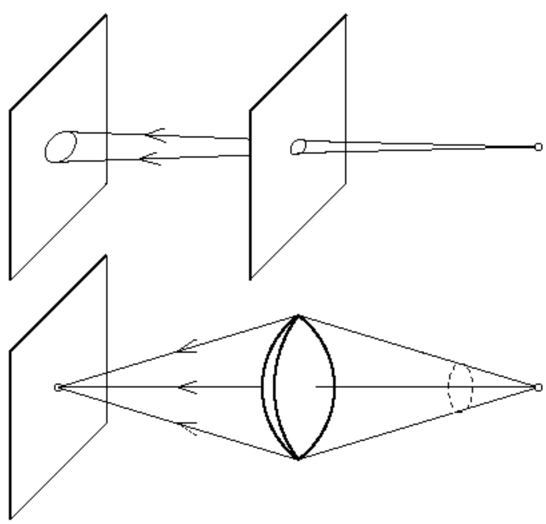
Pinhole too big -
many directions are averaged, blurring the image

Pinhole too small-
diffraction effects blur the image

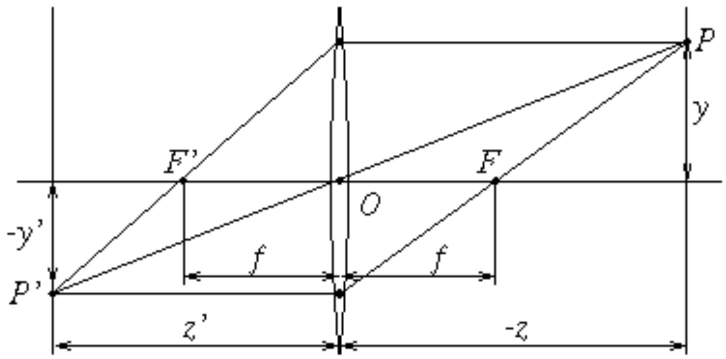
Generally, pinhole cameras are *dark*, because a very small set of rays from a particular point hits the screen.



The reason for lenses



The thin lens

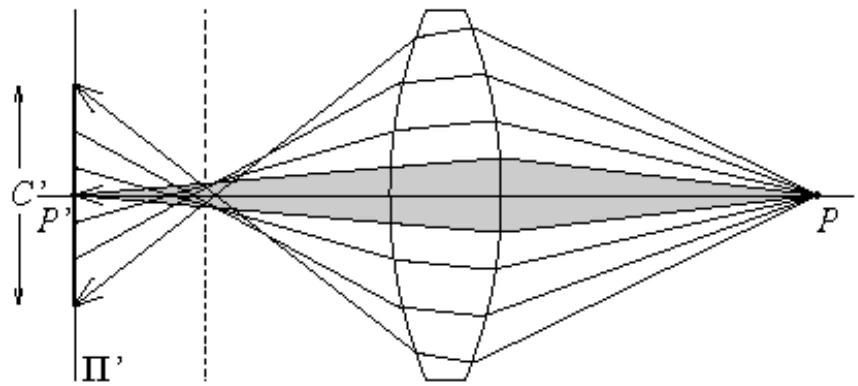


$$\frac{1}{z'} - \frac{1}{z} = \frac{1}{f}$$

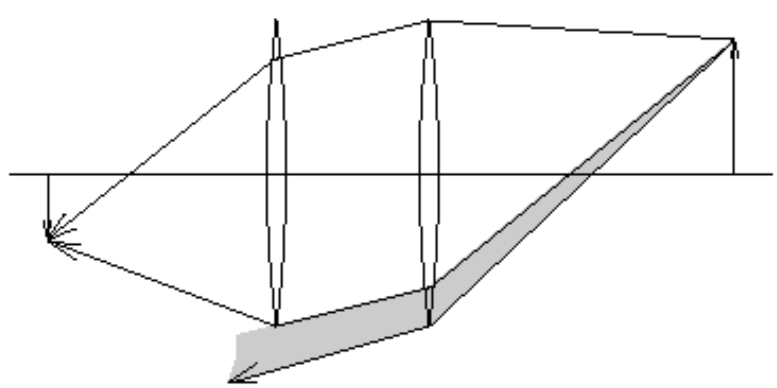
Issues with real lenses

- Depth of field
- Focus
  - Point Spread Function
- Lens aberrations
  - Chromatic aberration
  - spherical aberration

Spherical aberration



Vignetting



Other (possibly annoying) phenomena

- Chromatic aberration
  - Light at different wavelengths follows different paths; hence, some wavelengths are defocussed
  - Machines: coat the lens
  - Humans: live with it
- Scattering at the lens surface
  - Some light entering the lens system is reflected off each surface it encounters (Fresnel's law gives details)
  - Machines: coat the lens, interior
  - Humans: live with it (various scattering phenomena are visible in the human eye)
- Geometric phenomena (Barrel distortion, etc.)

Radial Distortion

- Note how straight lines are distorted into curves by radial distortion in this image

Radial Distortion cont'd

- Corresponds to a dilation of the image
- It is most pronounced in optical systems with a wide field of view

$$x = x_d(1 + k_1 r^2 + k_2 r^4)$$

$$y = y_d(1 + k_1 r^2 + k_2 r^4)$$

$$r^2 = x_d^{\ 2} + y_d^{\ 2}$$

Summary of camera parameters

- Extrinsic
  - Translation vector T
  - Rotation Matrix R
- Intrinsic
  - Focal length
  - Image center
  - Pixel size
  - Radial distortion coefficients

Camera calibration

- Issues:
  - what are intrinsic parameters of the camera?
  - what is the camera matrix? (intrinsic+extrinsic)
- General strategy:
  - view calibration object
  - identify image points
  - obtain camera matrix by minimizing error
  - obtain intrinsic parameters from camera matrix
- Error minimization:
  - Linear least squares
    - easy problem numerically
    - solution can be rather bad
  - Minimize image distance
    - more difficult numerical problem
    - solution usually rather good,
    - start with linear least squares
  - Numerical scaling is an issue