Image Formation and Representation

CS485/685 Computer Vision

Dr. George Bebis

A Simple model of image formation

- The scene is illuminated by a single source.
- The scene reflects radiation towards the camera.
- The camera senses it via solid state cells (CCD cameras)

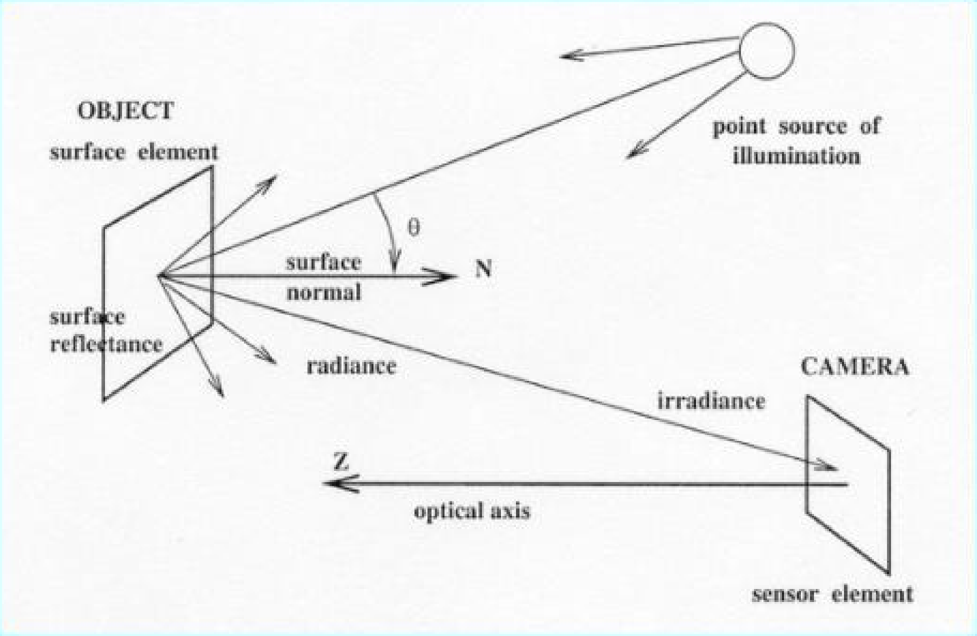


Image formation (cont'd)

- There are two parts to the image formation process:

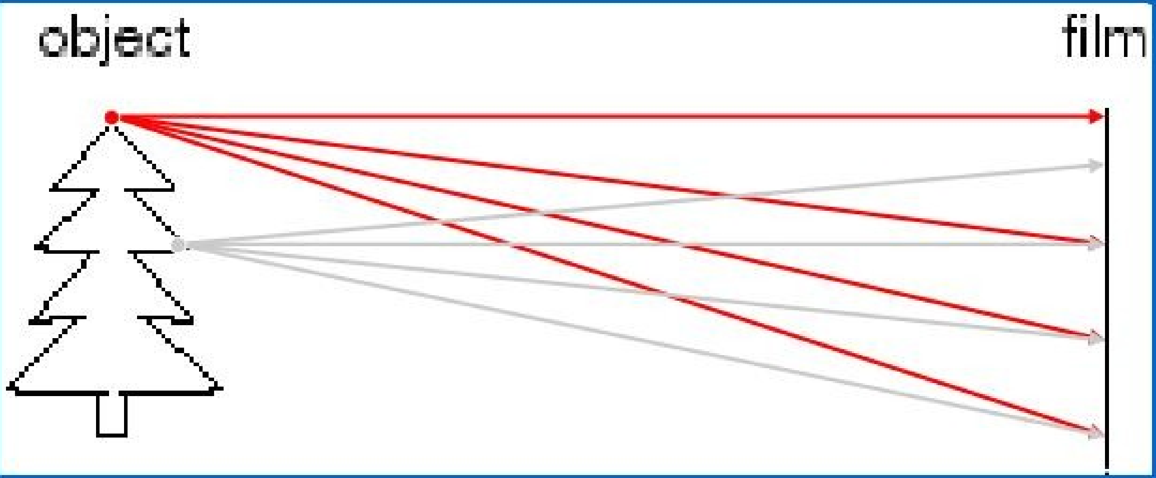
  (1) The **geometry**, which determines where in the image plane the projection of a point in the scene will be located.

  (2) The **physics of light**, which determines the brightness of a point in the image plane.

Simple model: $\mathbf{f(x,y) = i(x,y)\ r(x,y)}$

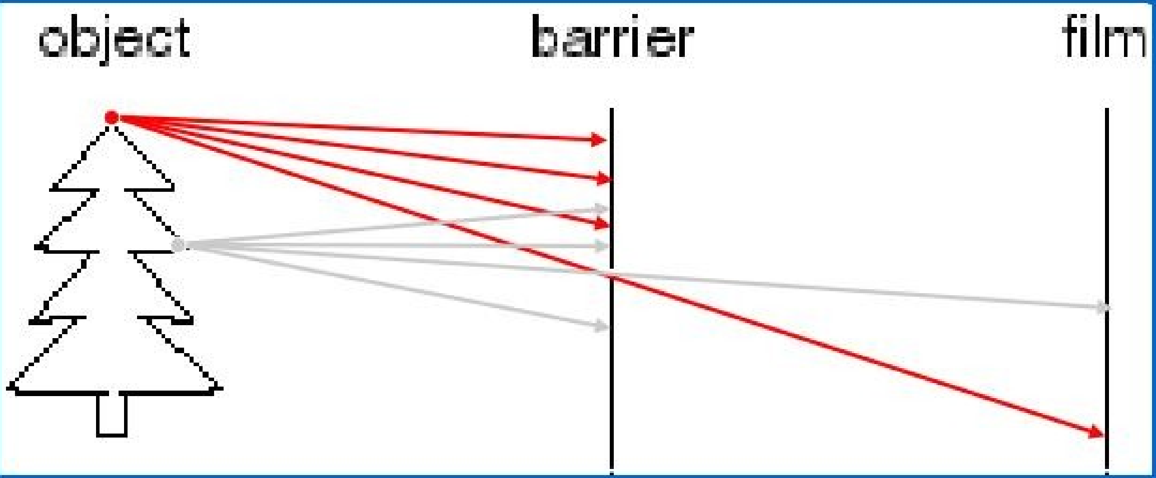
i: illumination, r: reflectance

Let’s design a camera



- Put a piece of film in front of an object - do we get a reasonable image?
  - Blurring - need to be more selective!

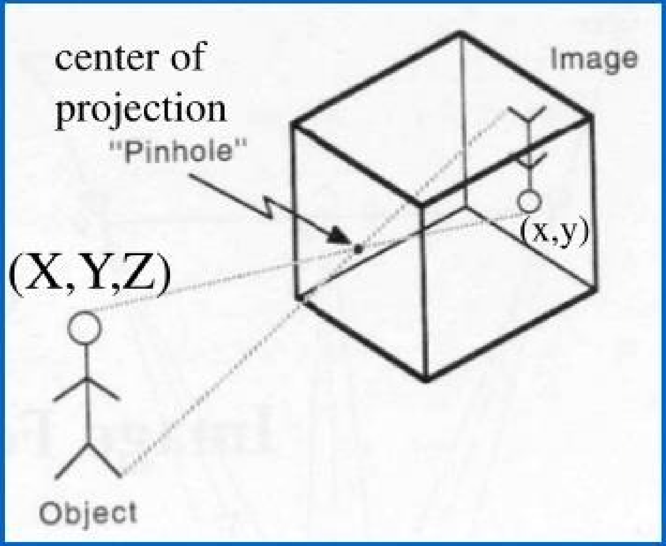
Let’s design a camera (cont’d)



- Add a barrier with a small opening (i.e. aperture) to block off most of the rays
  - Reduces blurring

“Pinhole” camera model

- The simplest device to form an image of a 3D scene on a 2D surface.
- Rays of light pass through a "pinhole" and form an inverted image of the object on the image plane.

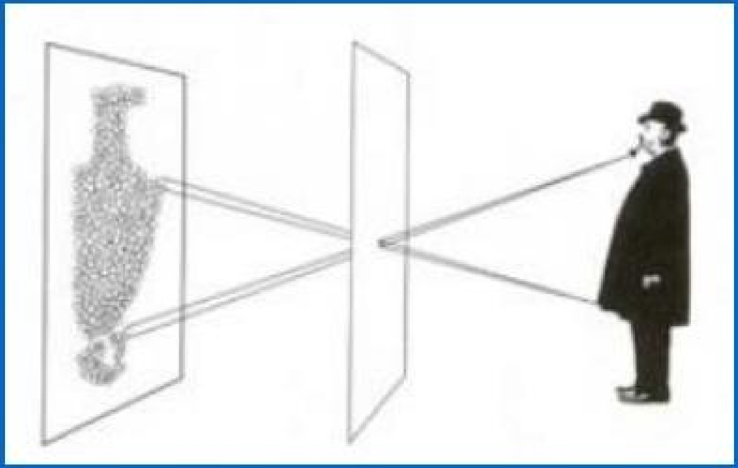


perspective projection:

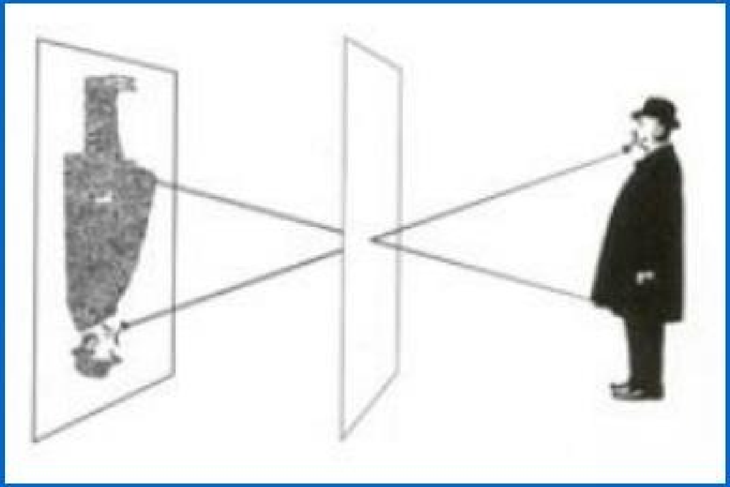
$$x = \frac{fX}{Z} \qquad y = \frac{fY}{Z}$$

f: focal length

What is the effect of aperture size?

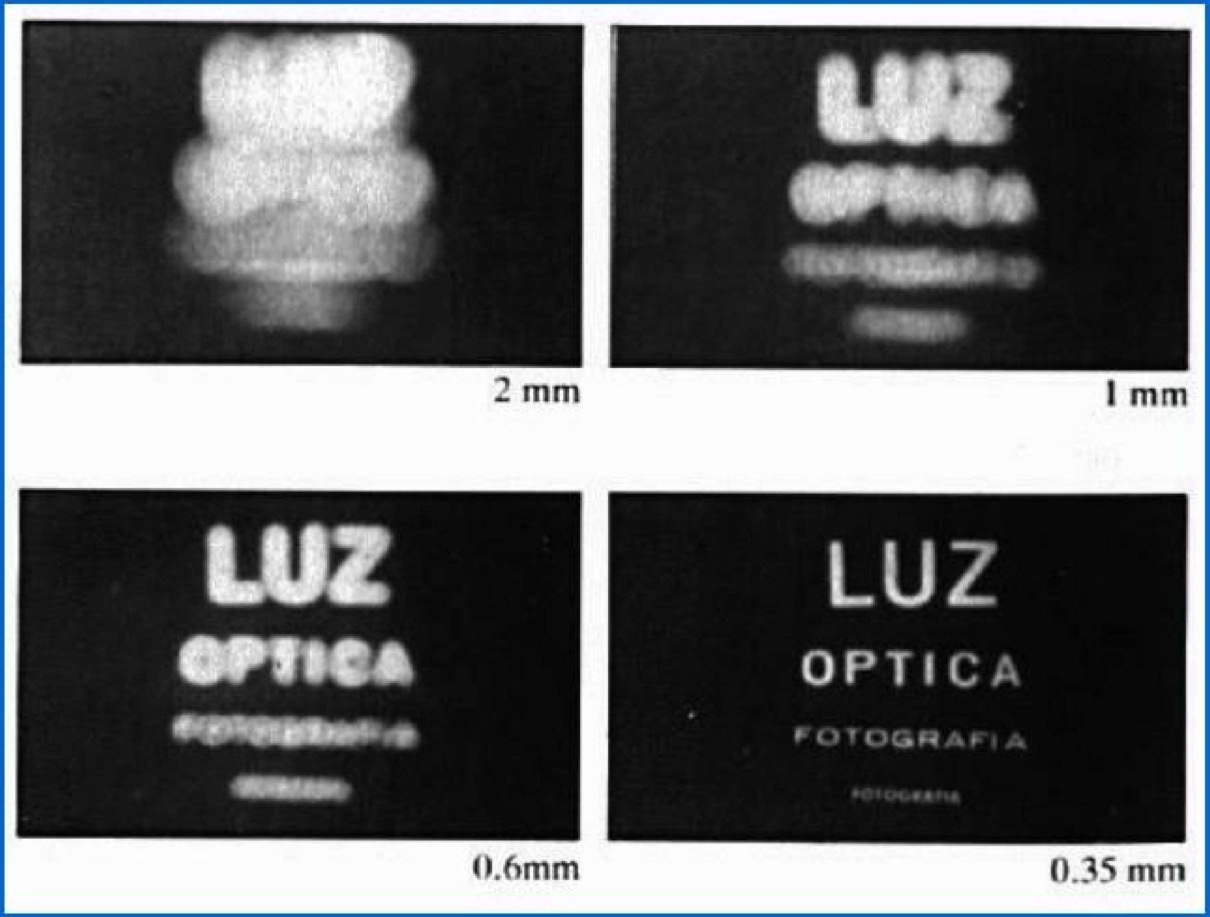


Large aperture: light from the source spreads across the image (i.e., not properly focused), making it blurry!



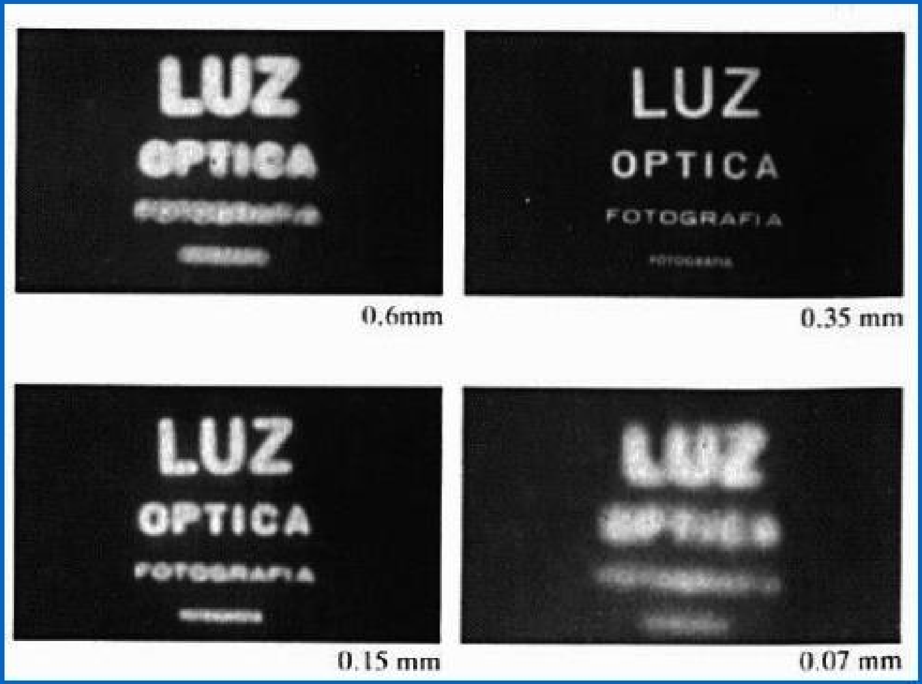
Small aperture: reduces blurring but (i) it limits the amount of light entering the camera and (ii) causes light diffraction.

Example: varying aperture size



Example: varying aperture size (cont'd)

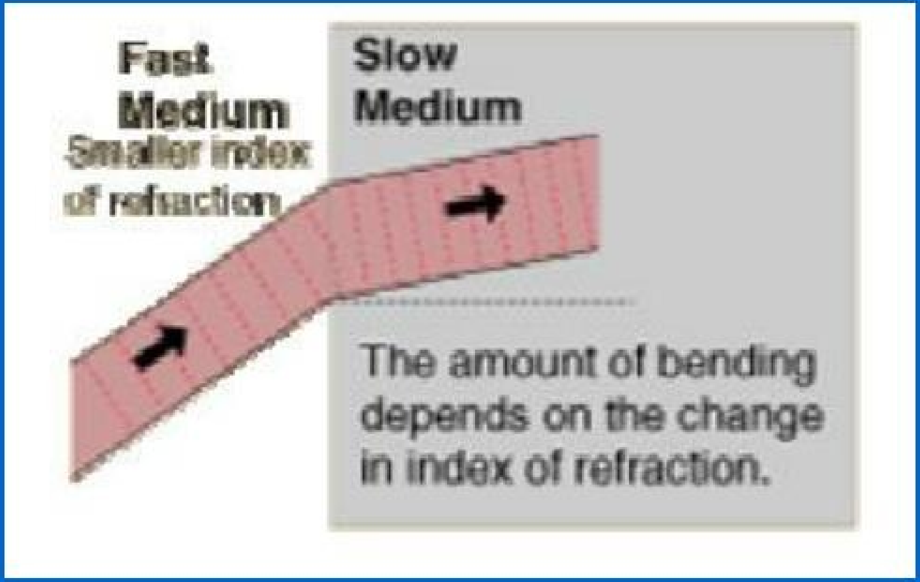
- What happens if we keep decreasing aperture size?
- When light passes through a small hole, it does not travel in a straight line and is scattered in many directions (i.e., diffraction)



SOLUTION: refraction

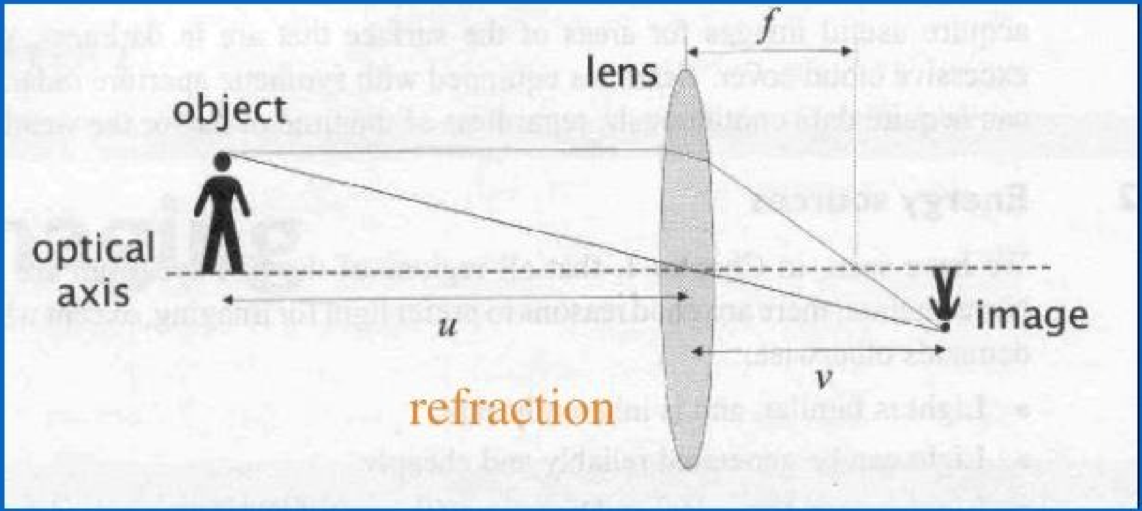
Refraction

- Bending of wave when it enters a medium where its speed is different.



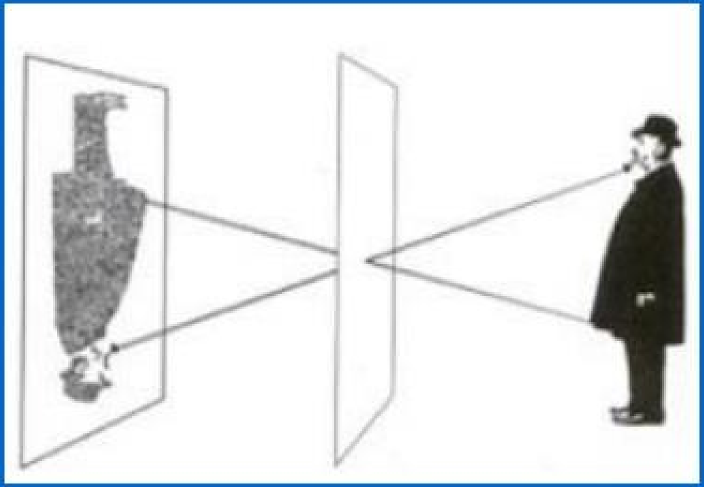
Lens

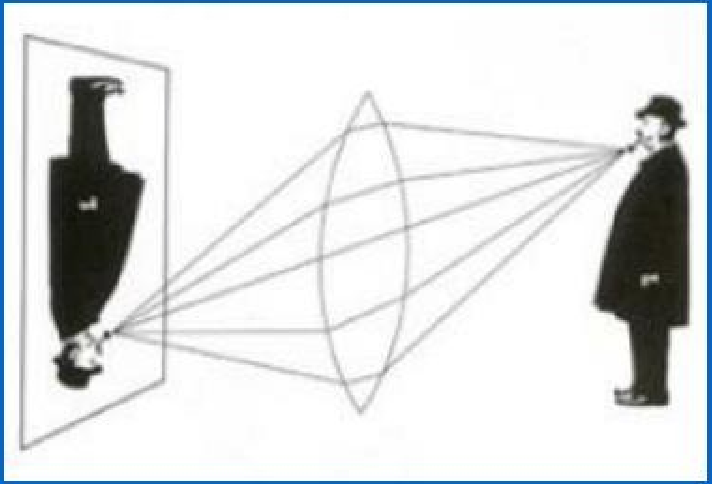
- Lens duplicate pinhole geometry without resorting to undesirably small apertures.
  - Gather all the light radiating from an object point towards the lens's finite aperture .
  - Bring light into focus at a single distinct image point.



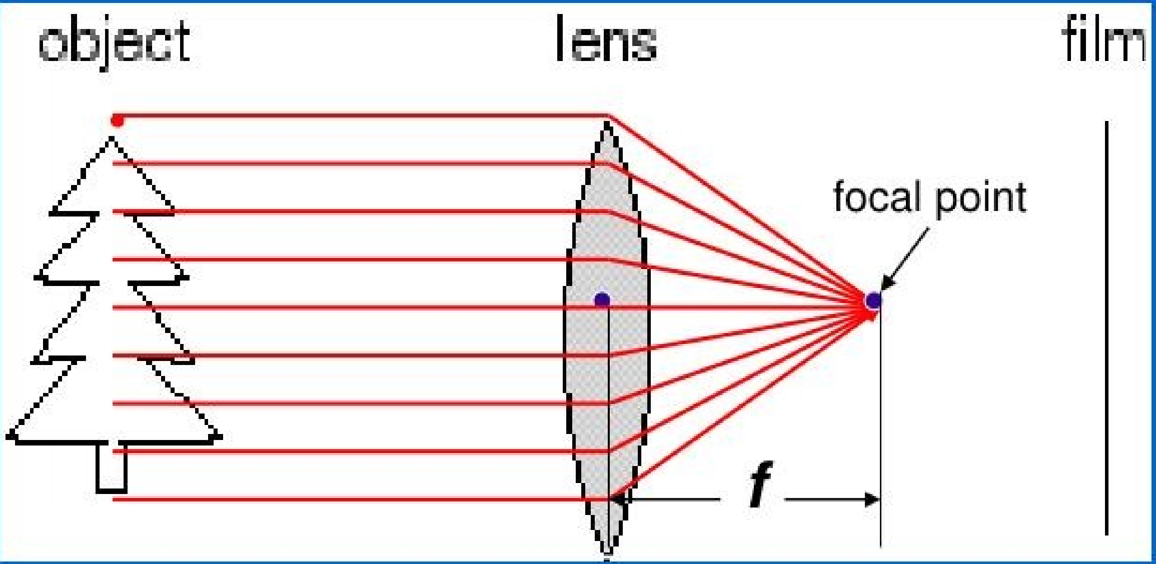
Lens (cont’d)

- Lens improve image quality, leading to sharper images.



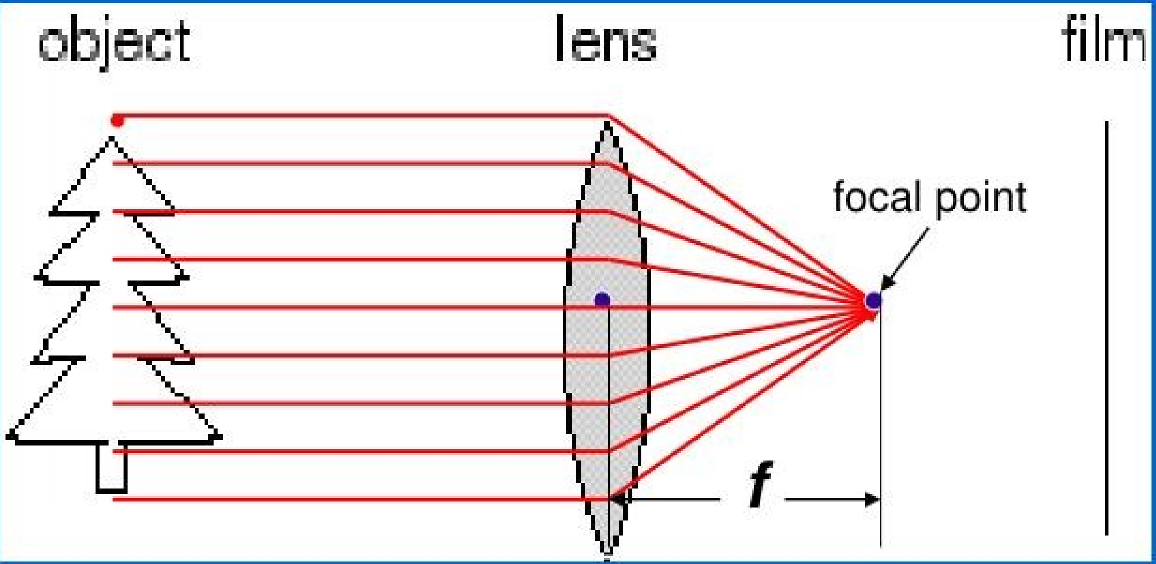


Properties of “thin” lens (i.e., ideal lens)



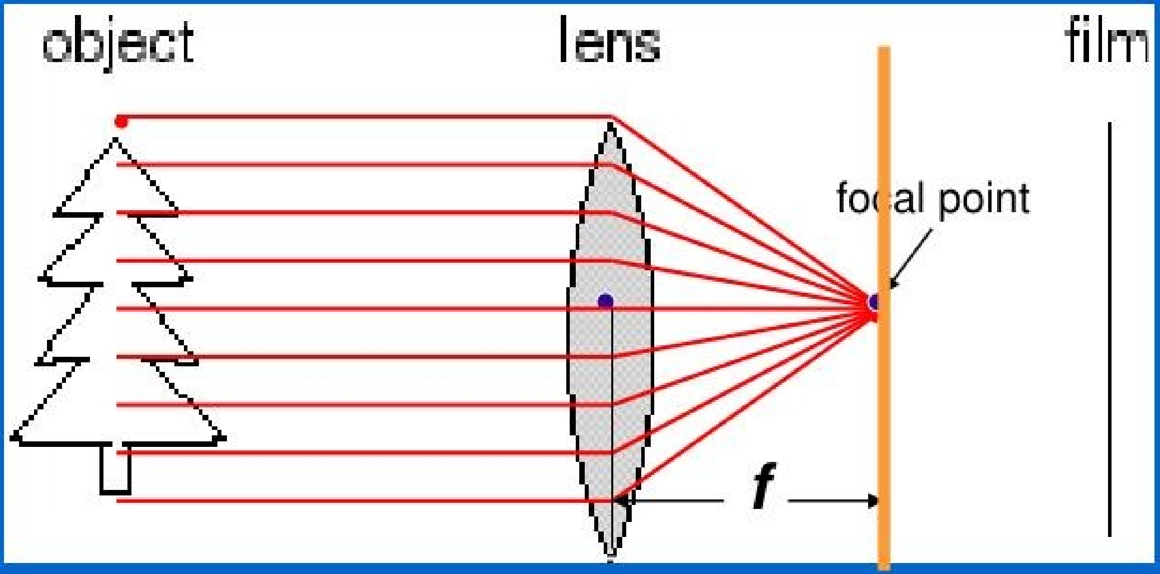
- Light rays passing through the center are not deviated.
- Light rays passing through a point far away from the center are deviated more.

Properties of “thin” lens (i.e., ideal lens)



- All parallel rays converge to a single point.
- When rays are perpendicular to the lens, it is called focal point.

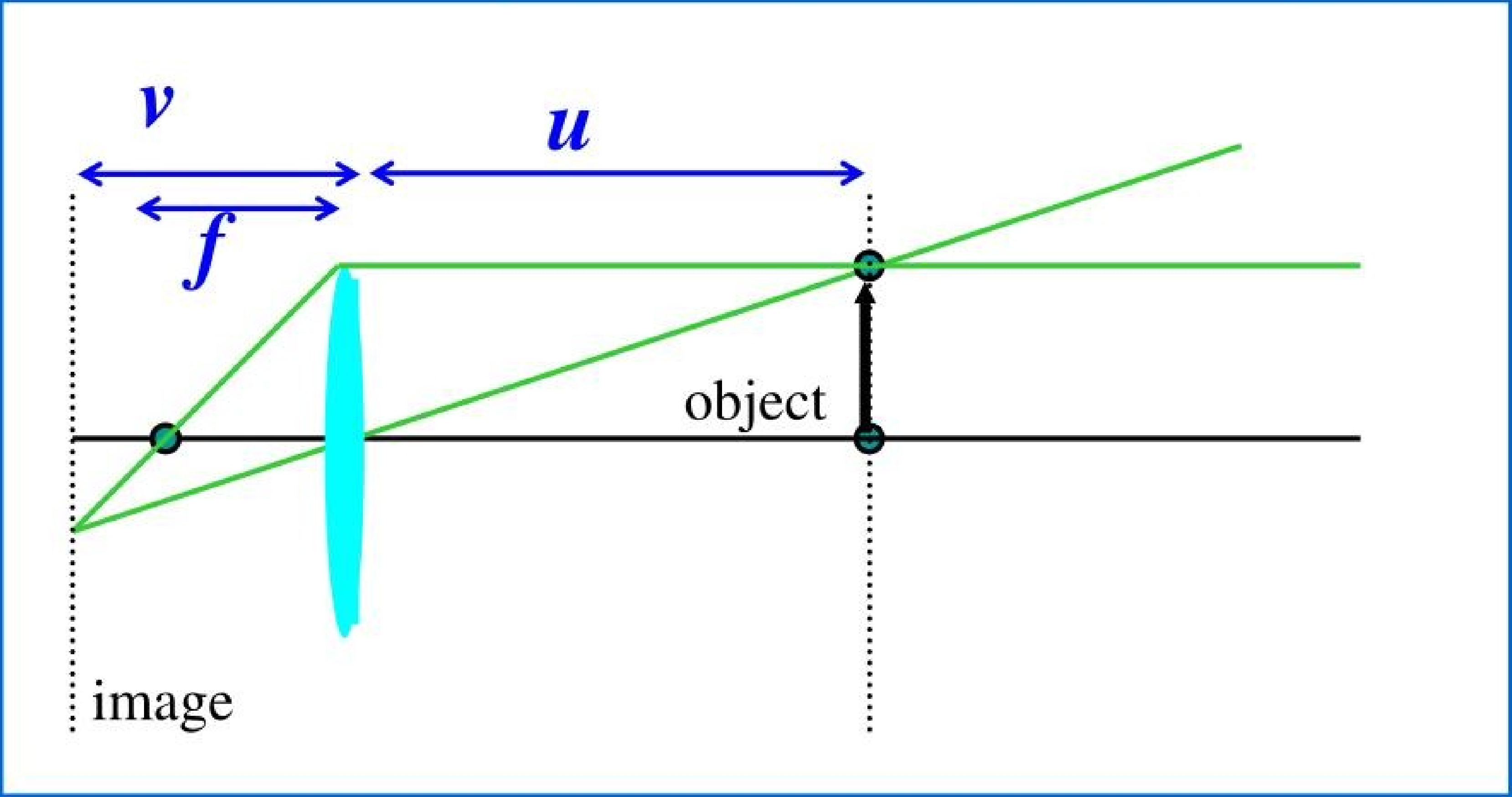
Properties of “thin” lens



- The plane parallel to the lens at the focal point is called the focal plane.
- The distance between the lens and the focal plane is called the focal length (i.e., f) of the lens.

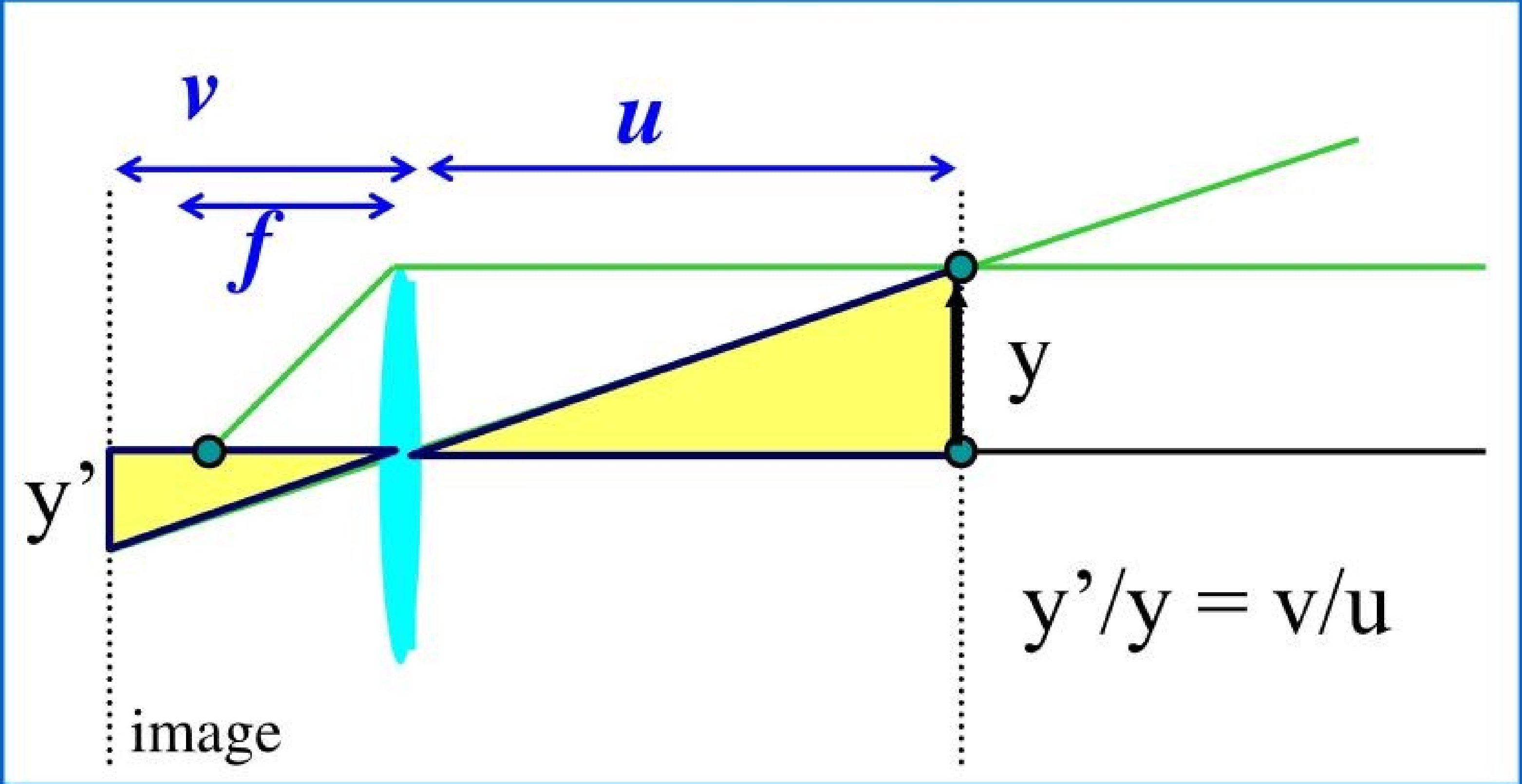
Thin lens equation

Assume an object at distance u from the lens plane:



Thin lens equation (cont'd)

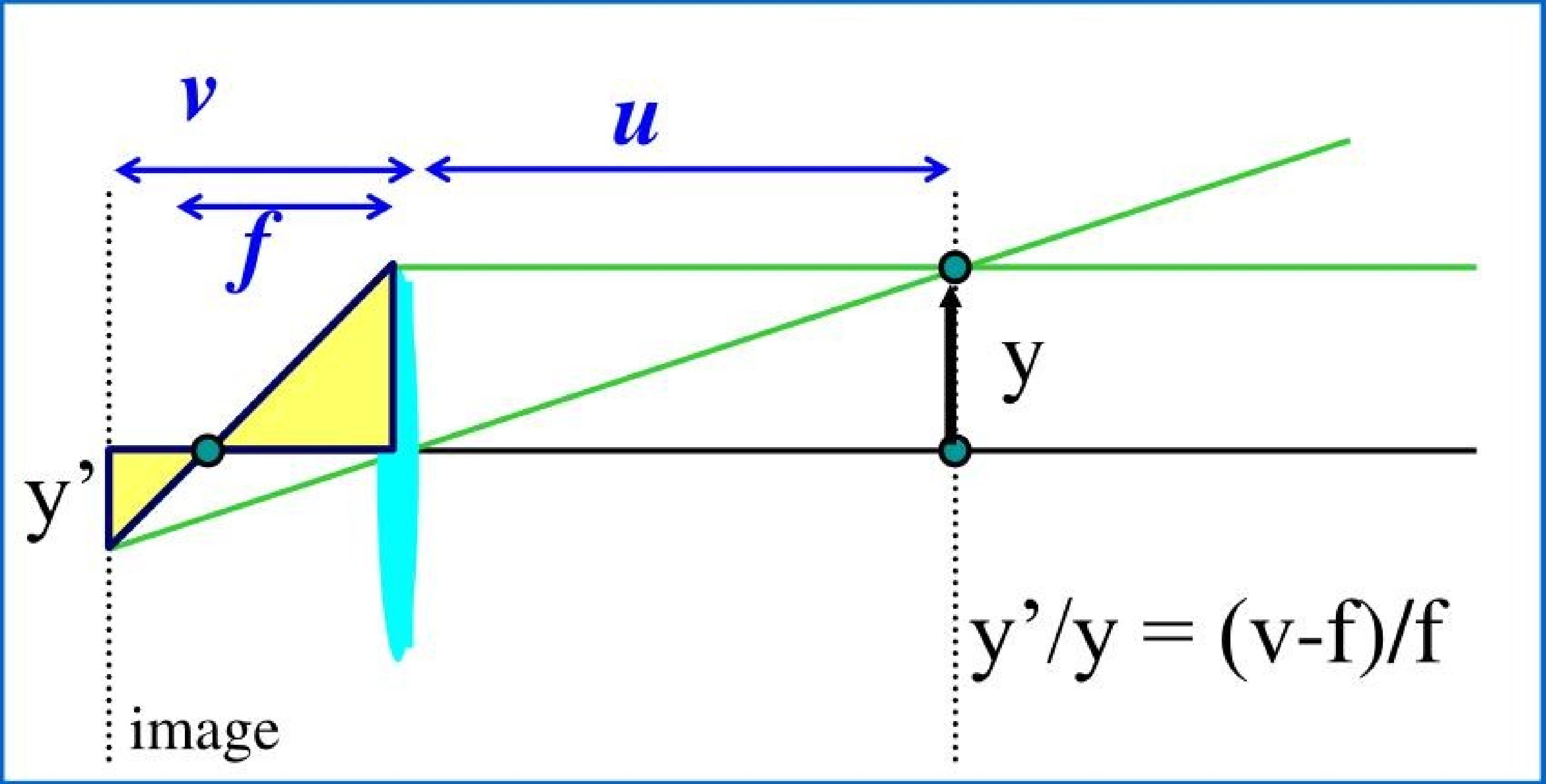
Using similar triangles:



$y'/y = v/u$

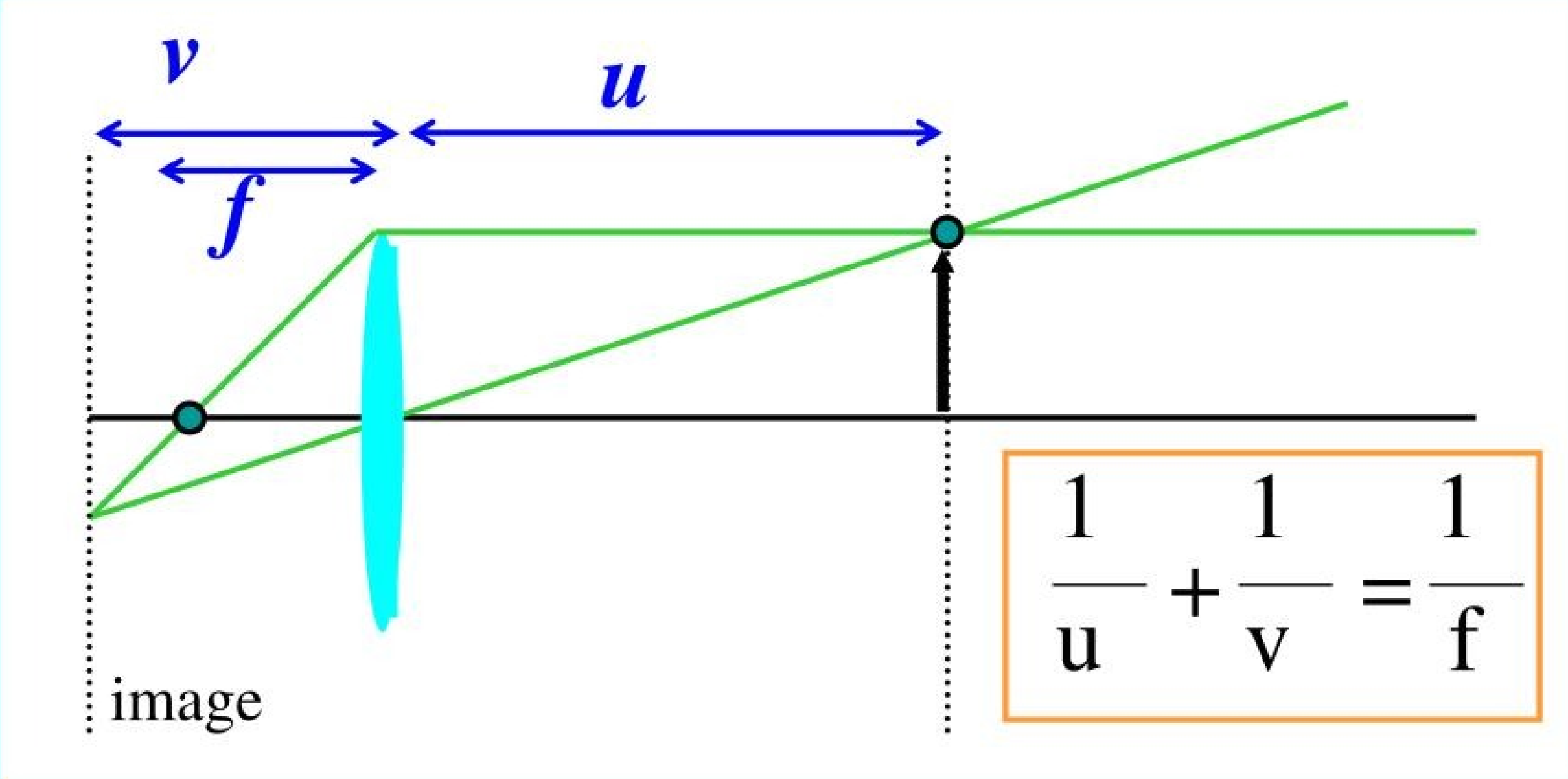
Thin lens equation (cont'd)

Using similar triangles:



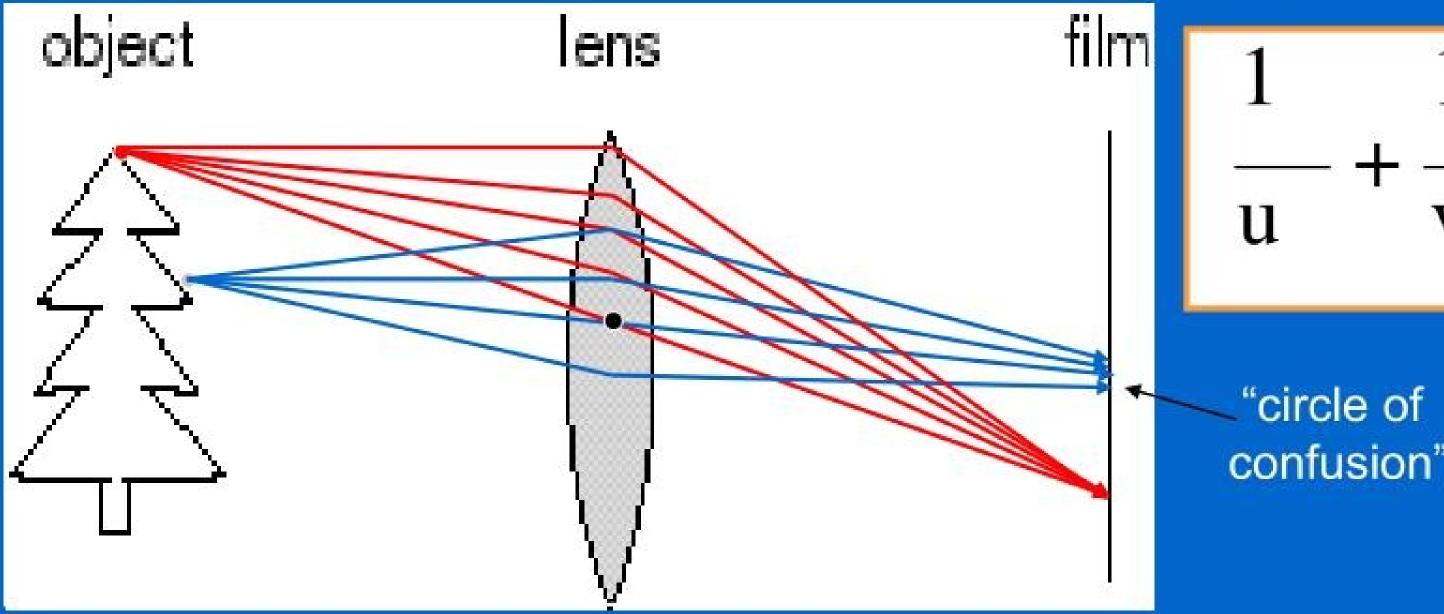
Thin lens equation (cont'd)

Combining the equations:



$$\boxed{\frac{1}{u}+\frac{1}{v}=\frac{1}{f}}$$

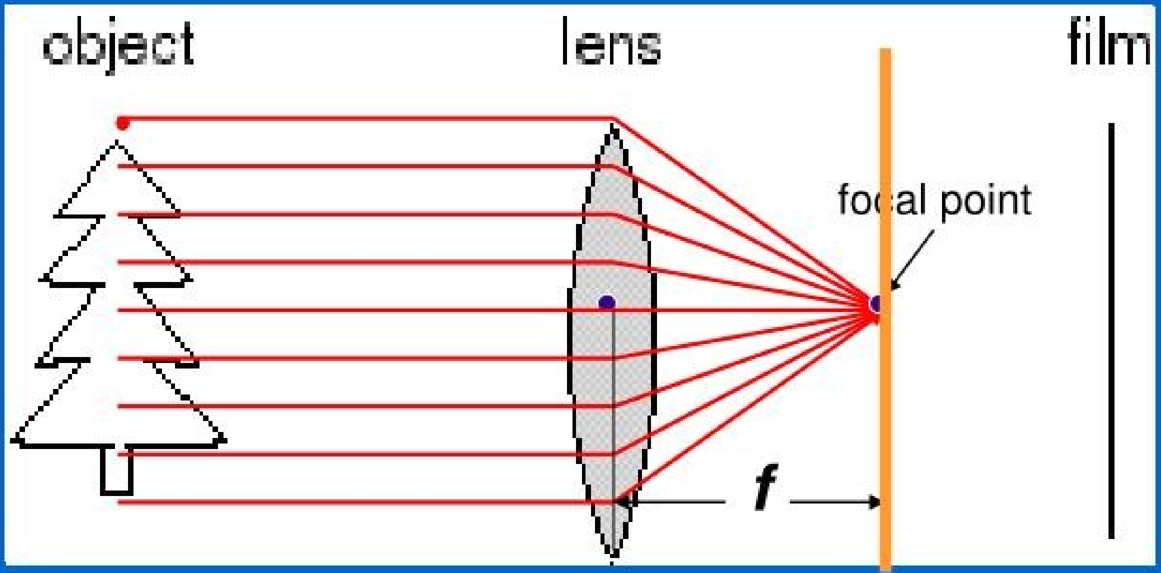
Thin lens equation (cont'd)



$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$$

- The thin lens equation implies that only points at distance u from the lens are "in focus" (i.e., focal point lies on image plane).
- Other points project to a "blur circle" or "circle of confusion" in the image (i.e., blurring occurs).

Thin lens equation (cont'd)

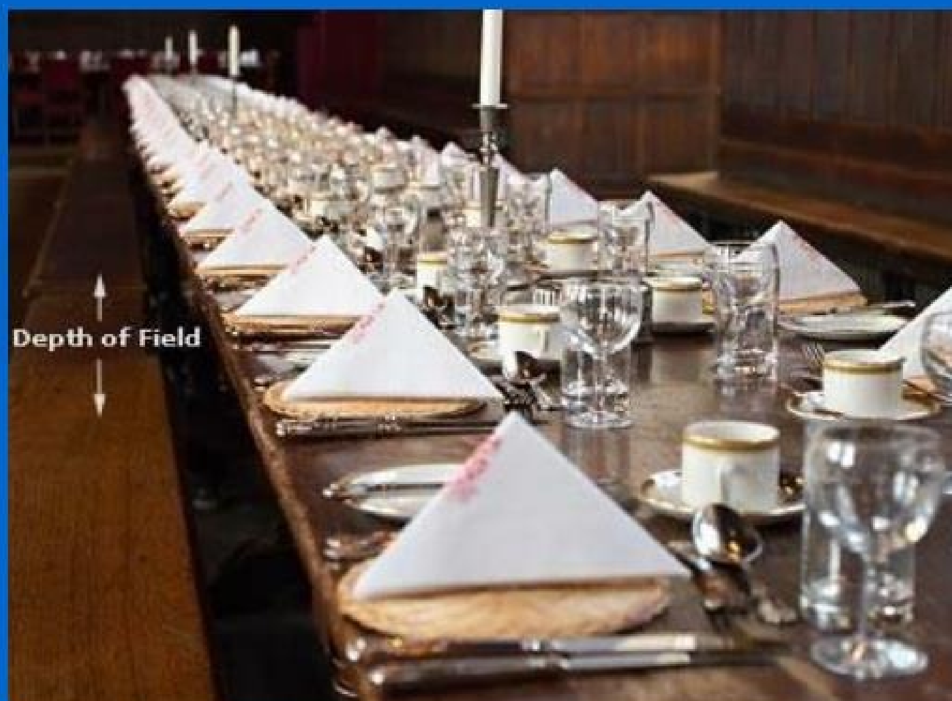


$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$$

- When objects move far away from the camera, then the focal plane approaches the image plane.

Depth of Field

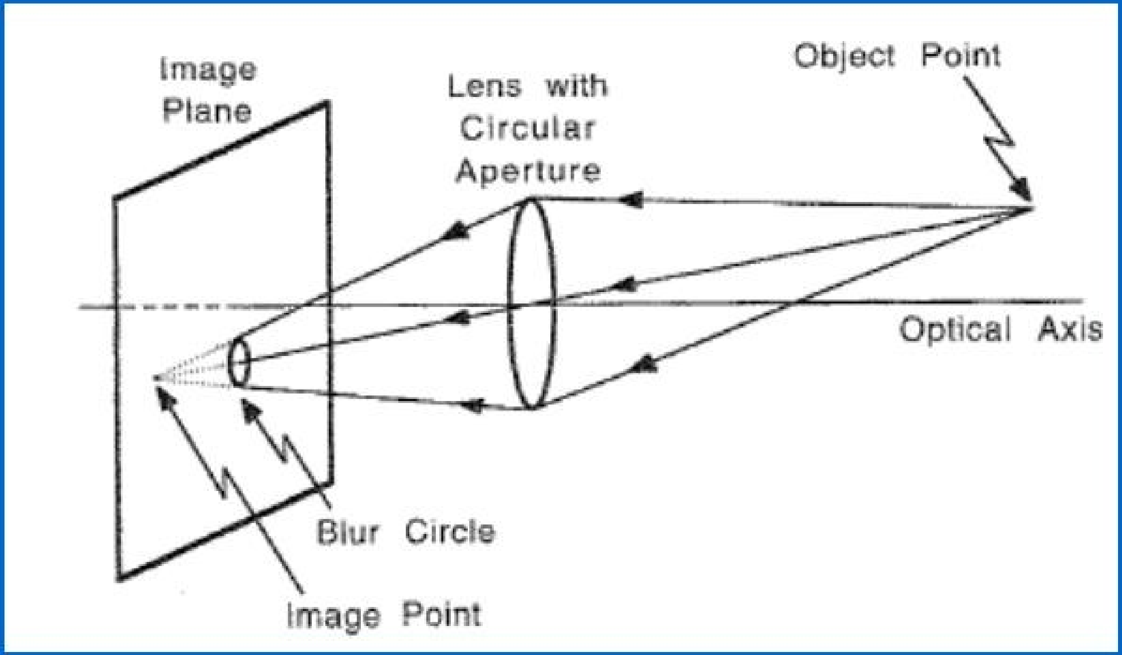
The range of depths over which the world is approximately sharp (i.e., in focus).



http://www.cambridgeincolour.com/tutorials/depth-of-field.htm

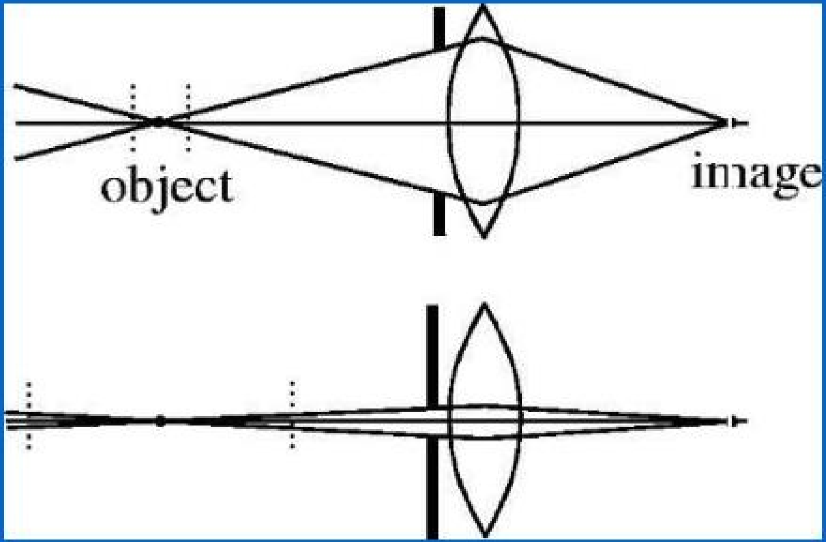
How can we control depth of field?

- The size of blur circle is proportional to aperture size.



How can we control depth of field? (cont'd)

- Changing aperture size (controlled by diaphragm) affects depth of field.
  - A smaller aperture increases the range in which an object is approximately in focus (but need to increase exposure time).
  - A larger aperture decreases the depth of field (but need to decrease exposure time).

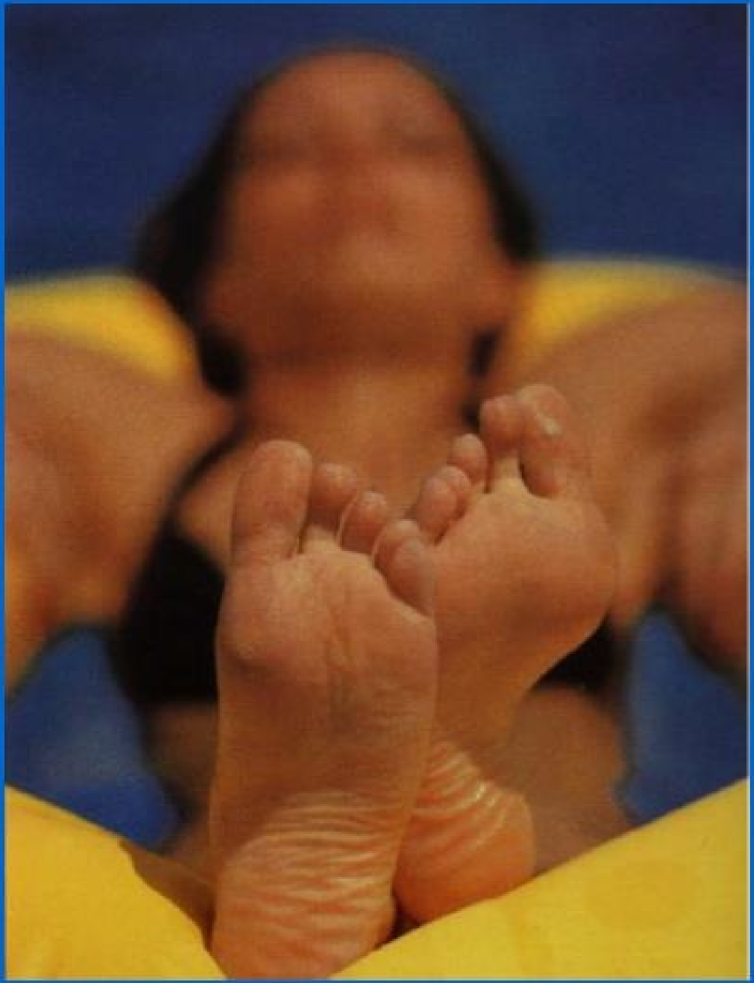


Varying aperture size

Large aperture = small DOF

Small aperture = large DOF

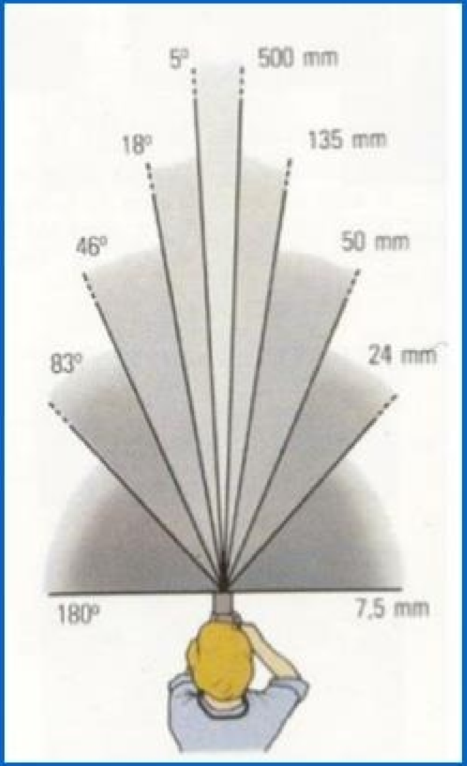
Another Example

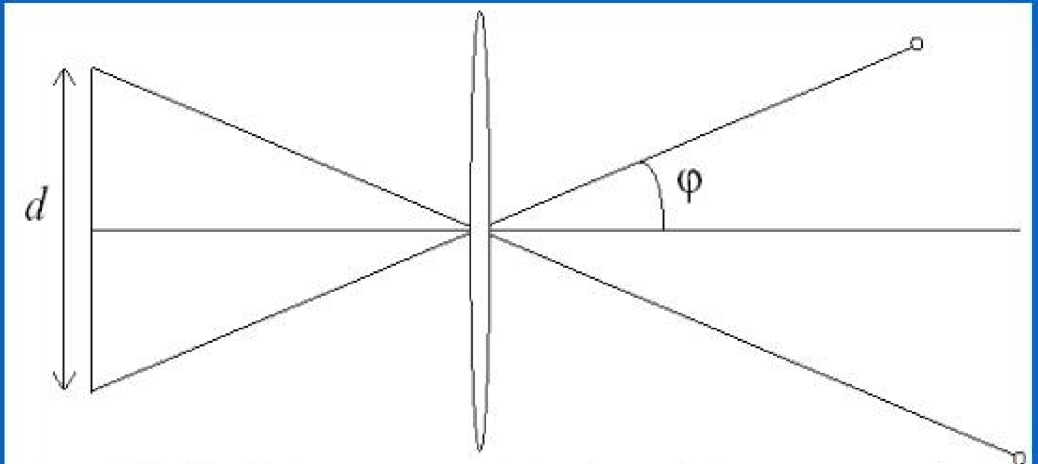


Large aperture = small DOF

Field of View (Zoom)

- The cone of viewing directions of the camera.
- Inversely proportional to focal length.

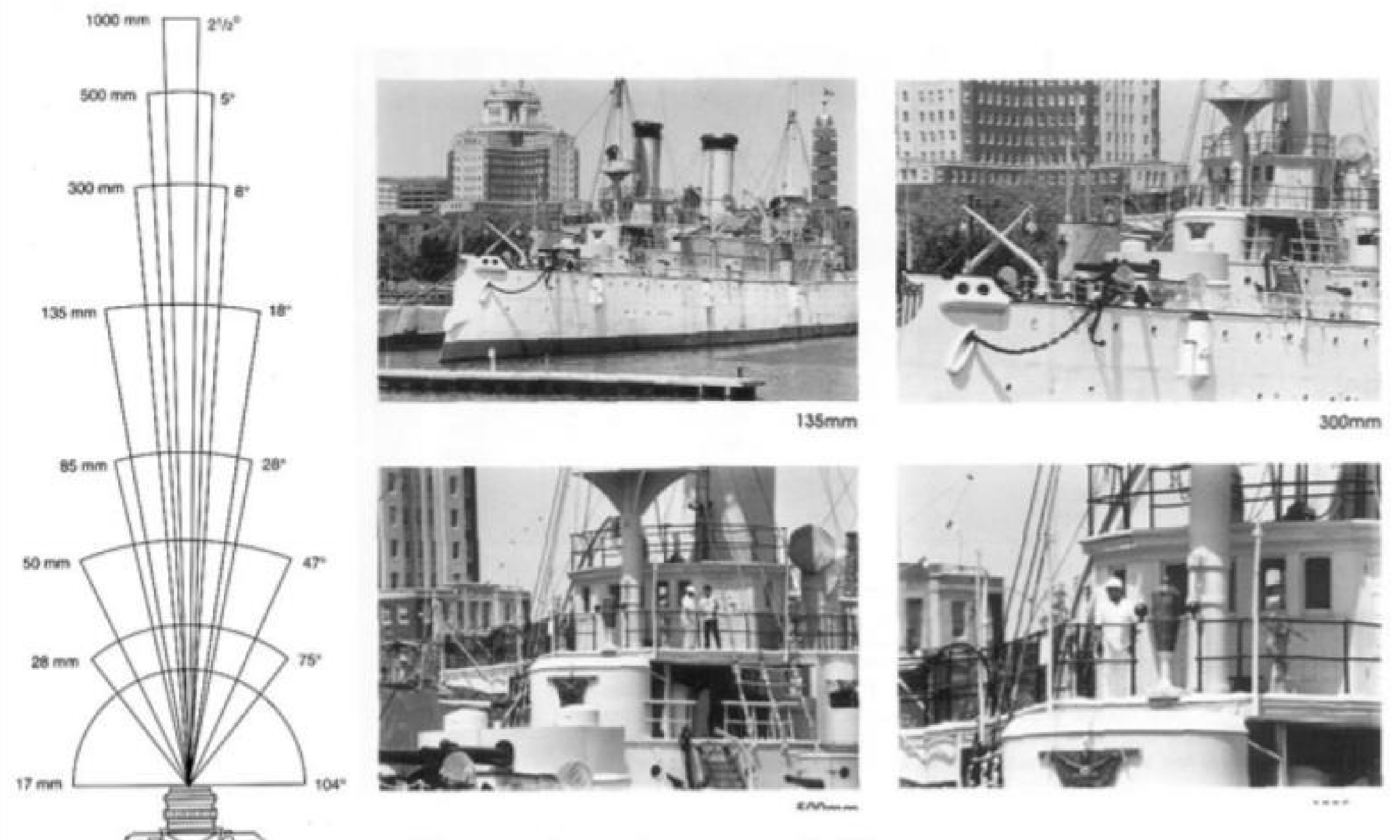




Size of field of view governed by size of the camera retina:

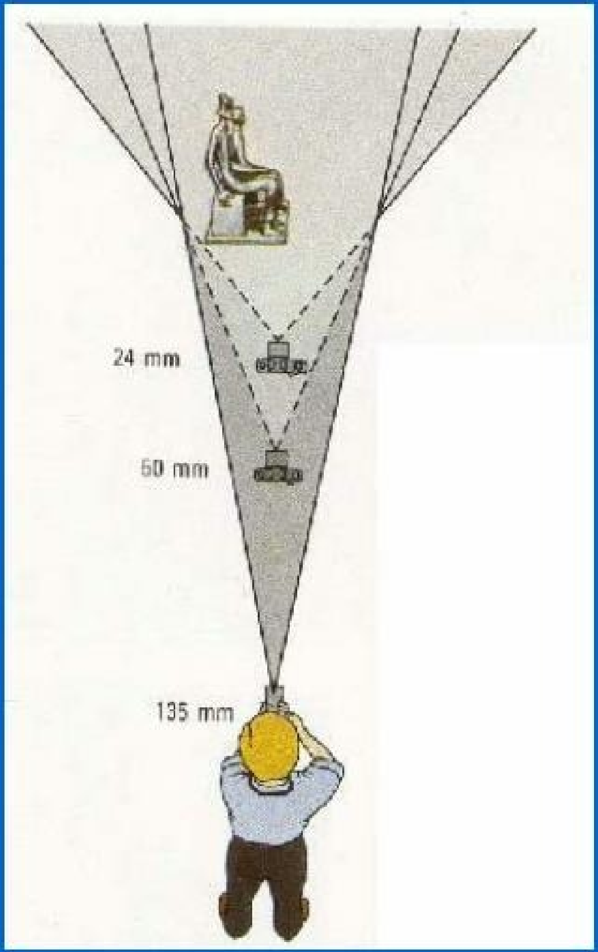
$$\varphi = \tan^{-1}\left(\frac{d}{2f}\right)$$

Field of View (Zoom)



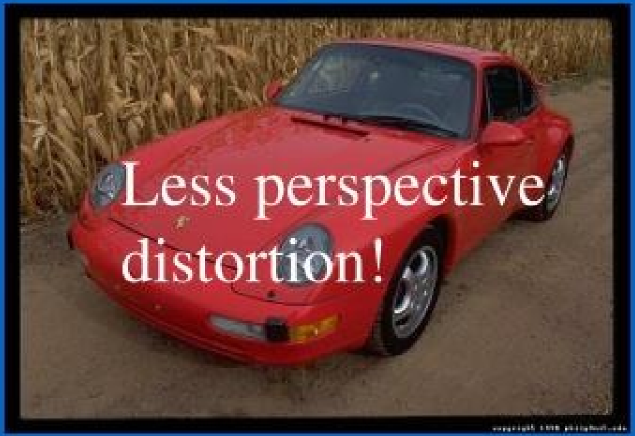
From London and Upton

Reduce Perspective Distortions by varying Distance / Focal Length



Small f (i.e., large FOV), camera close to car

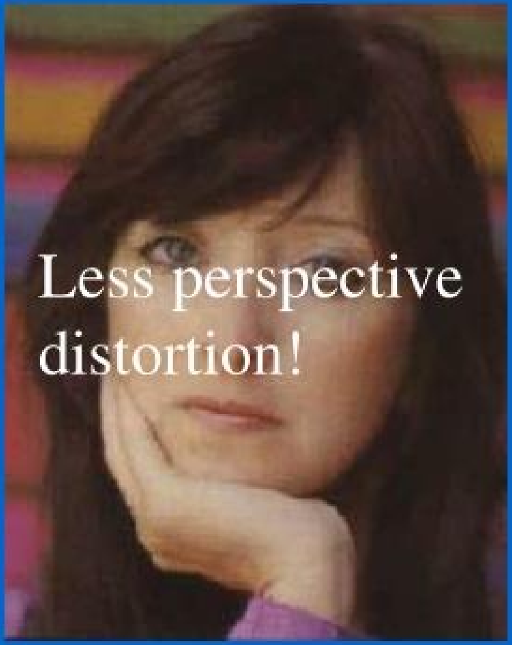
Large f (i.e., small FOV), camera far from car



Same effect for faces

wide-angle

standard

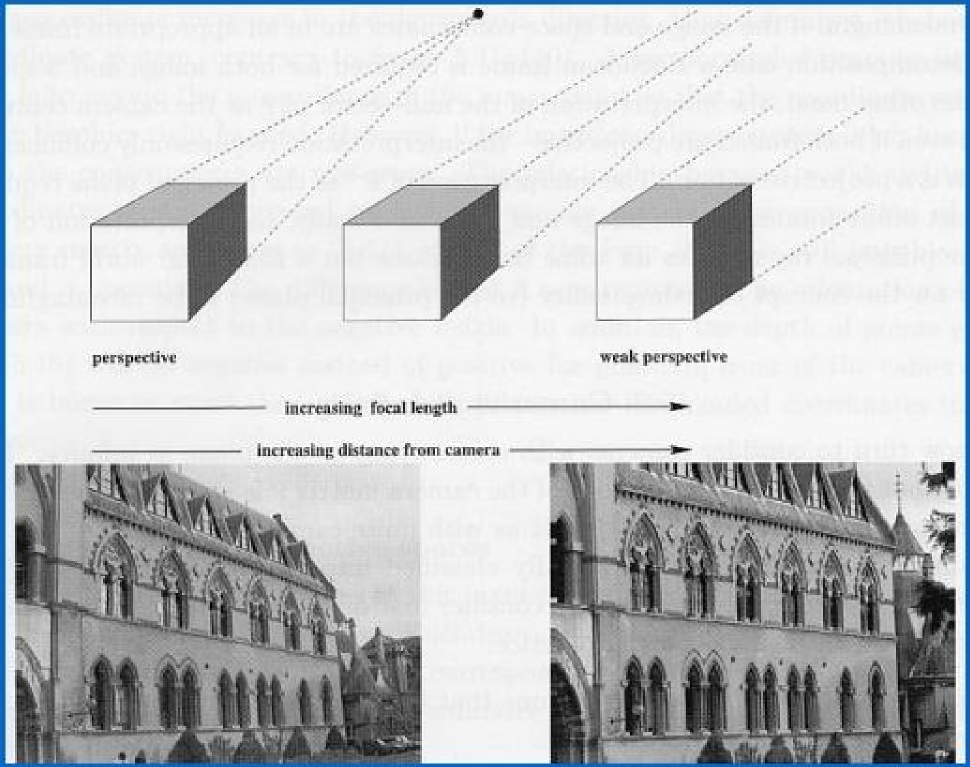


telephoto

<u>Practical significance</u>: we can approximate perspective projection using a simpler model when using telephoto lens to view a distant object that has a relatively small range of depth.

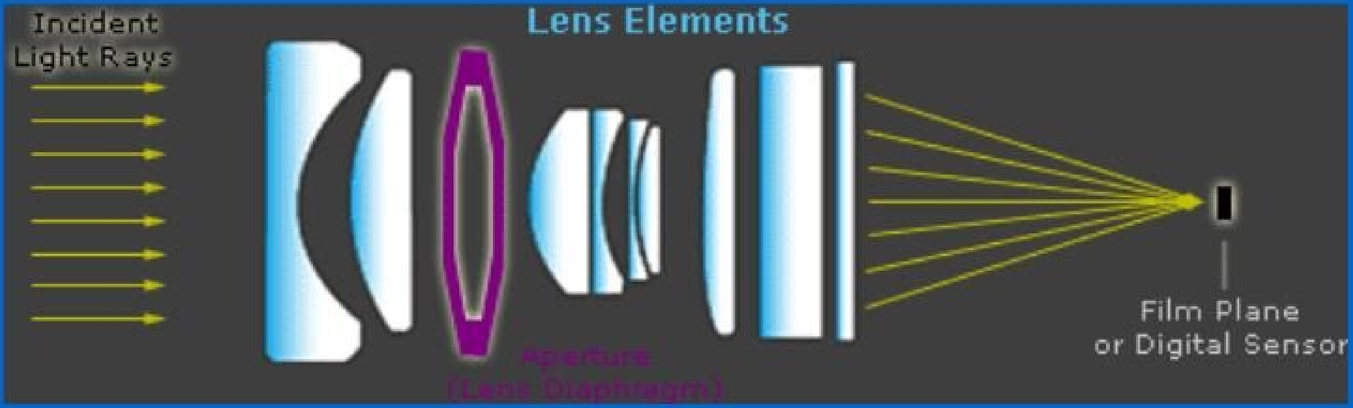
Approximating an “affine” camera

Center of projection is at infinity!



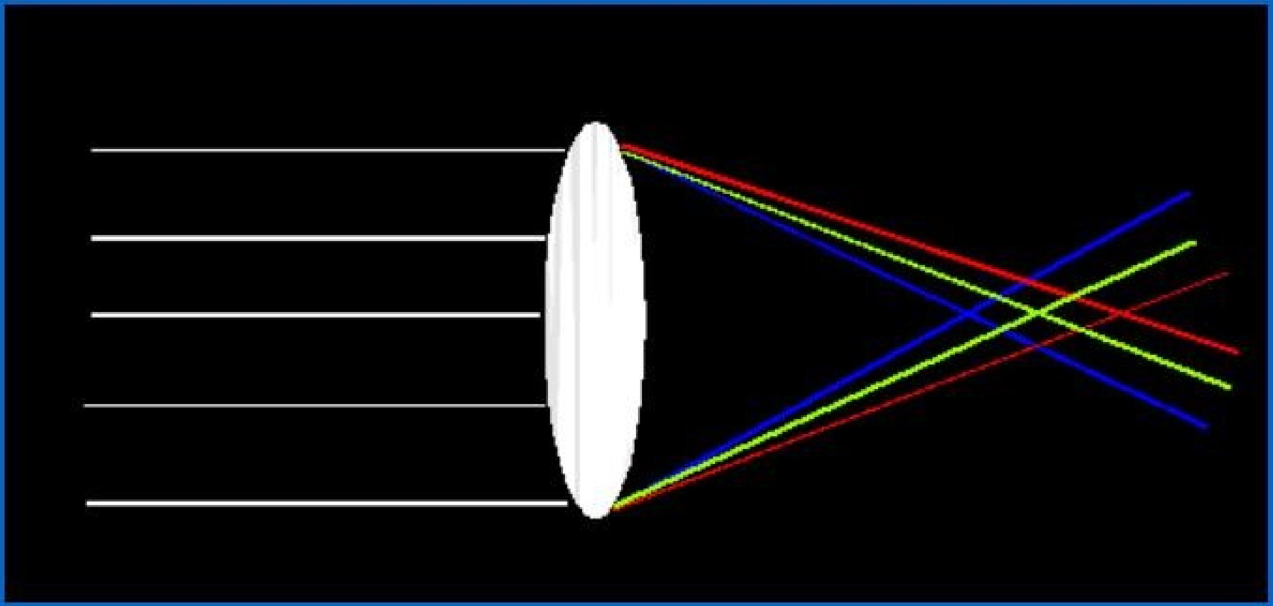
Real lenses

- All but the simplest cameras contain lenses which are actually comprised of several "lens elements."
- Each element aims to direct the path of light rays such that they recreate the image as accurately as possible on the digital sensor.



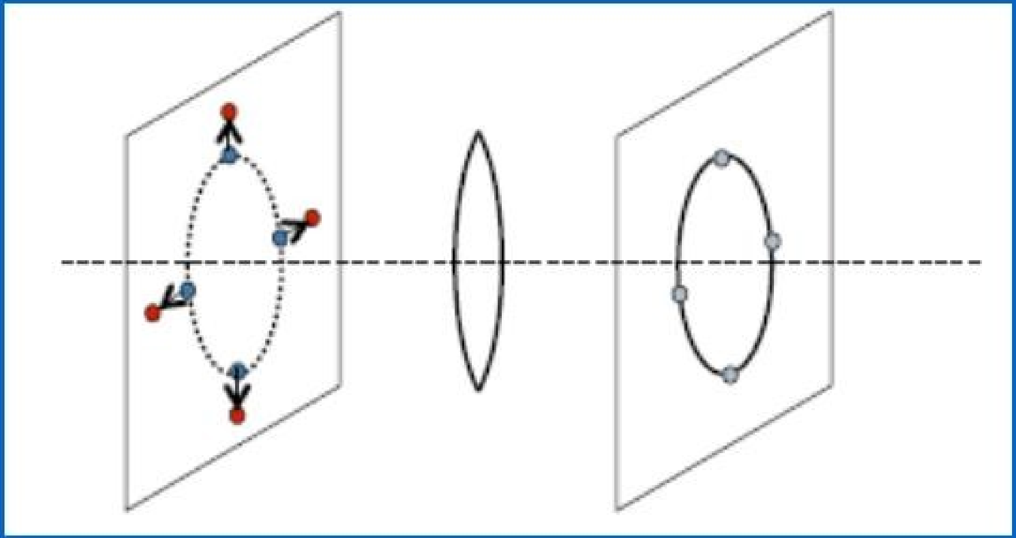
Lens Flaws: Chromatic Aberration

- Lens has different refractive indices for different wavelengths.
- Could cause color fringing:
  - i.e., lens cannot focus all the colors at the same point.



Chromatic Aberration - Example

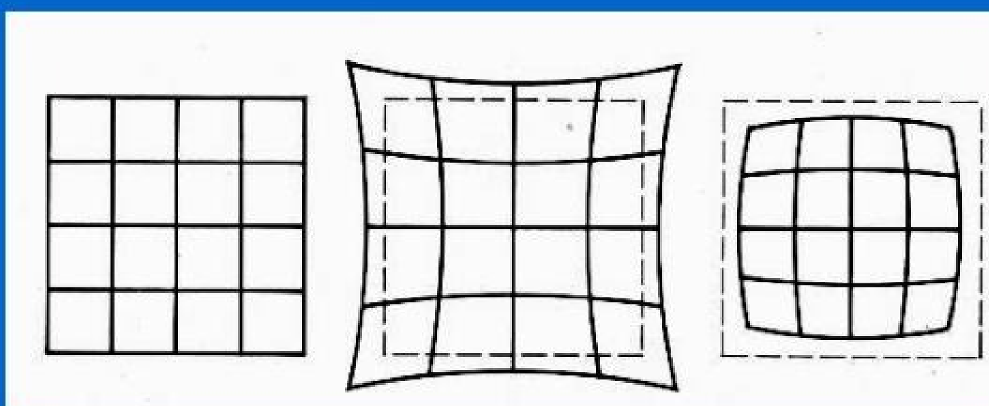
Lens Flaws: Radial Distortion



- Straight lines become distorted as we move further away from the center of the image.
- Deviations are most noticeable for rays that pass through the edge of the lens.

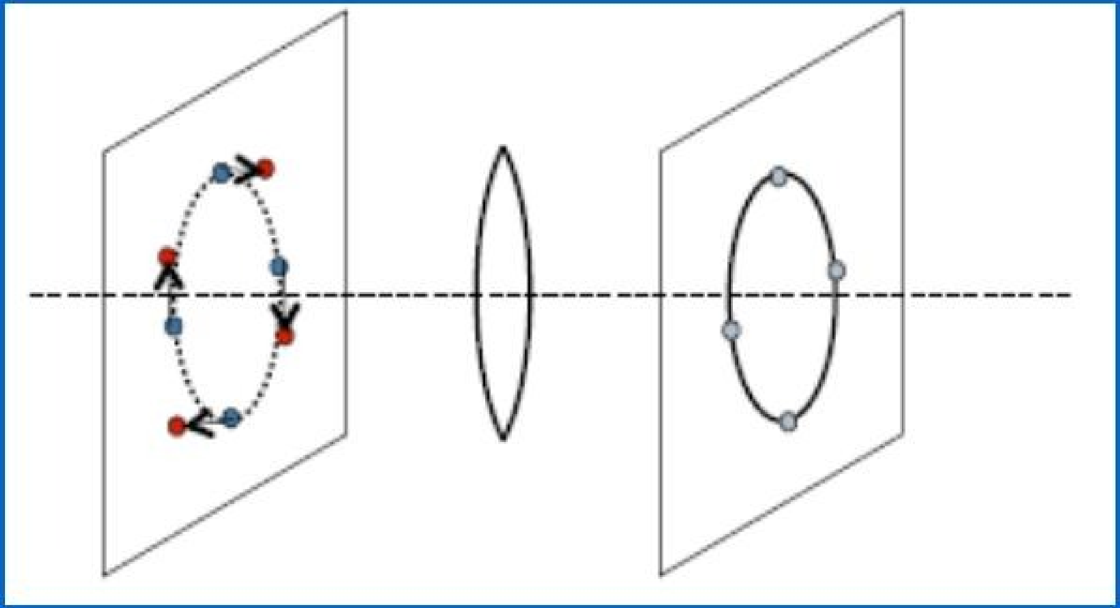
Lens Flaws: Radial Distortion (cont'd)

No distortion Pin cushion Barrel



Lens Flaws: Tangential Distortion

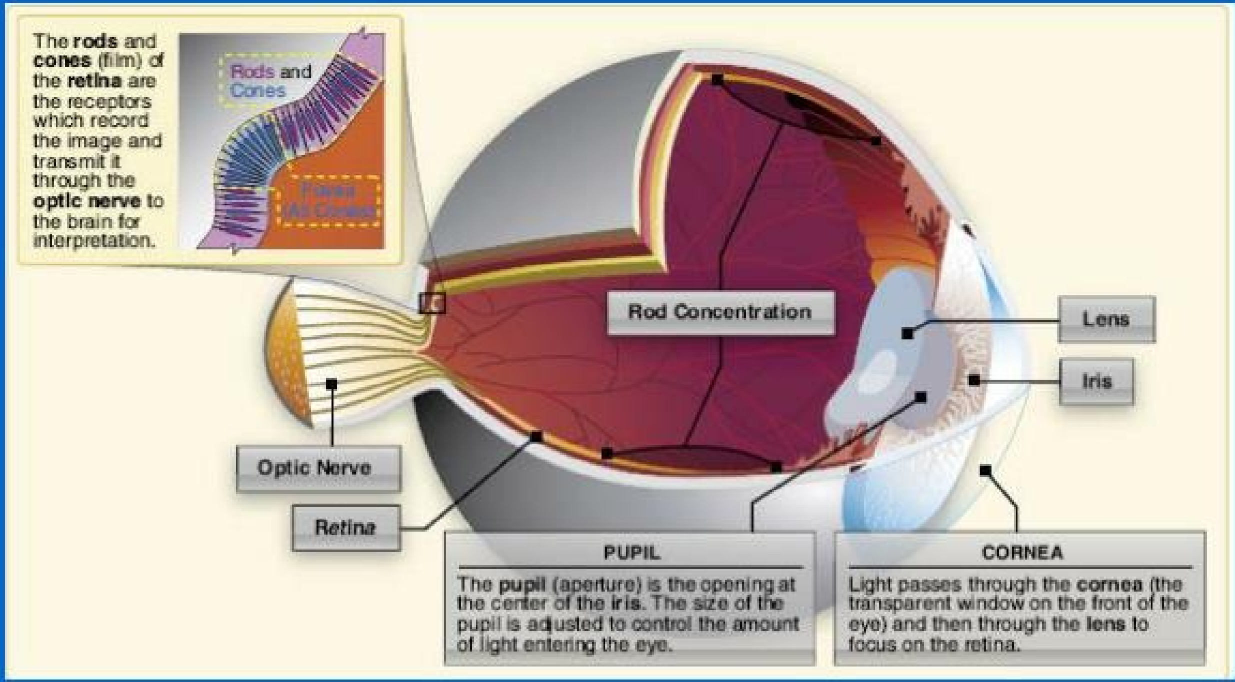
- Lens is not exactly parallel to the imaging plane!



Human Eye

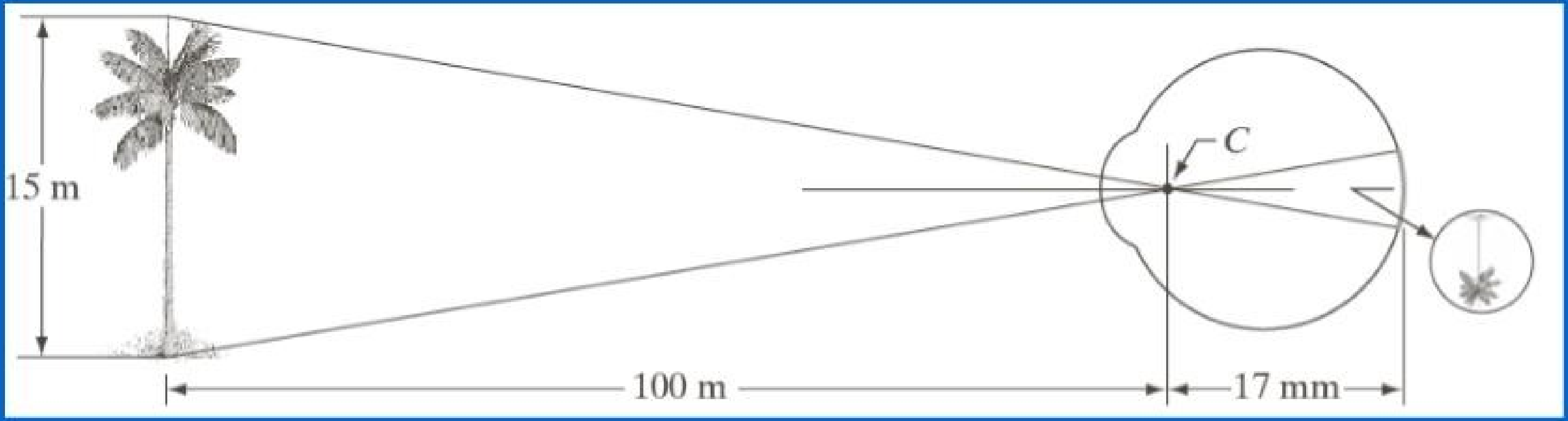
- Functions much like a camera:

  aperture (i.e., pupil), lens, mechanism for focusing (zoom in/out) and surface for registering images (i.e., retina)



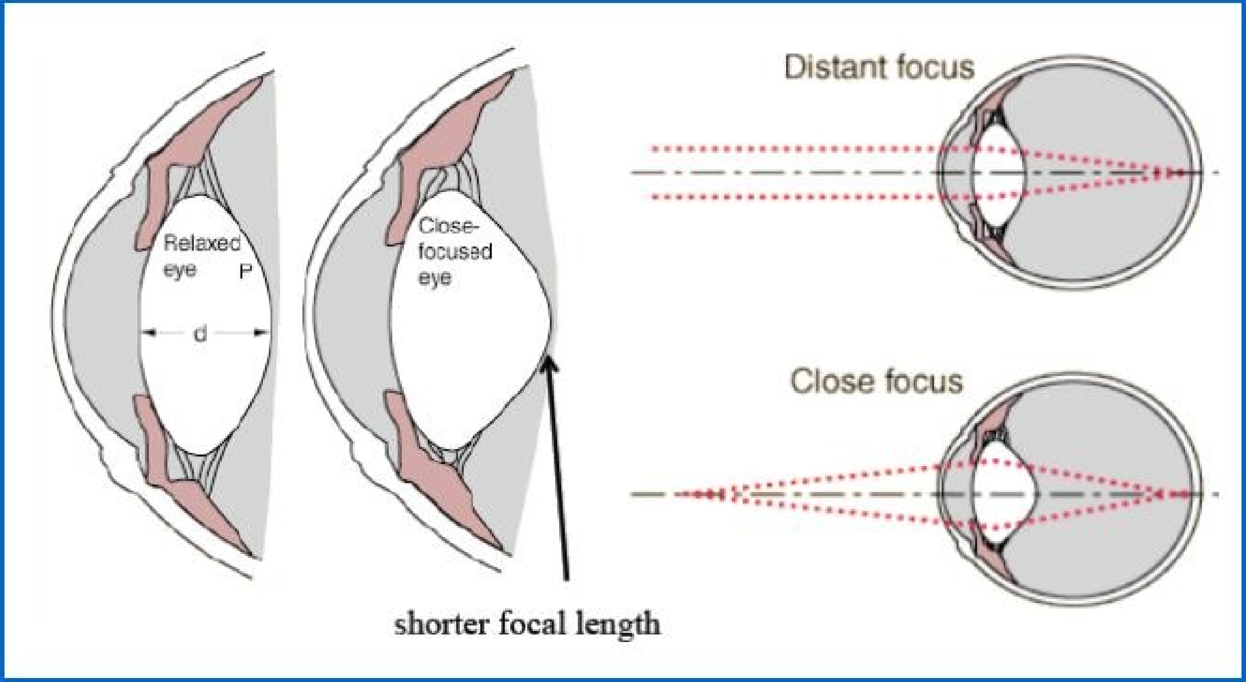
Human Eye (cont'd)

- In a camera, focusing at various distances is achieved by varying the distance between the lens and the imaging plane.
- In the human eye, the distance between the lens and the retina is fixed (i.e., 14mm to 17mm).



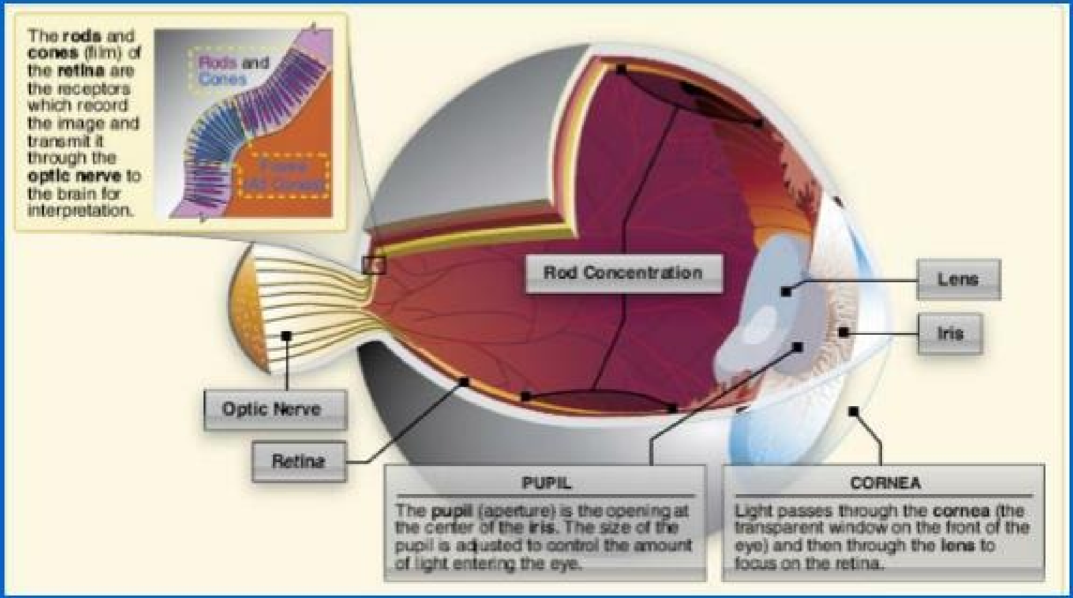
Human Eye (cont'd)

- Focusing is achieved by varying the shape of the lens (i.e., flattening of thickening).



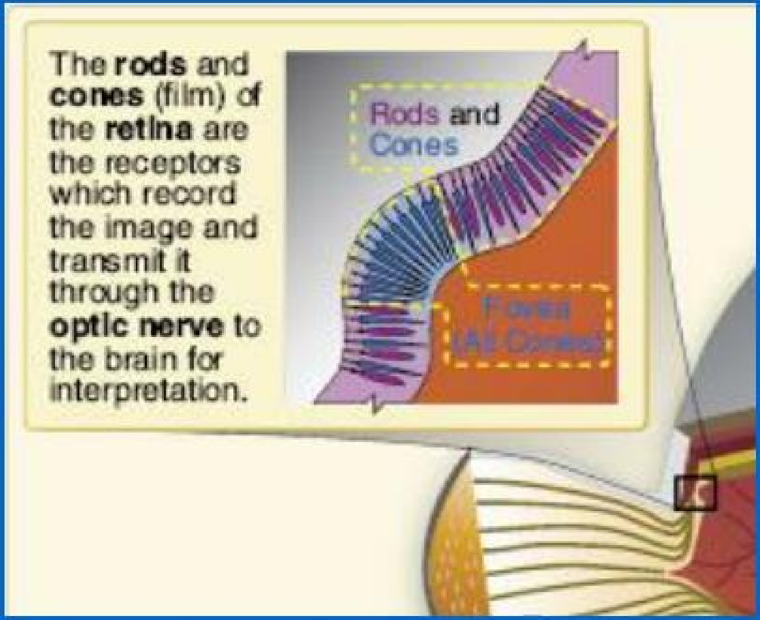
Human Eye (cont'd)

- Retina contains light sensitive cells that convert light energy into electrical impulses that travel through nerves to the brain.
- Brain interprets the electrical signals to form images.



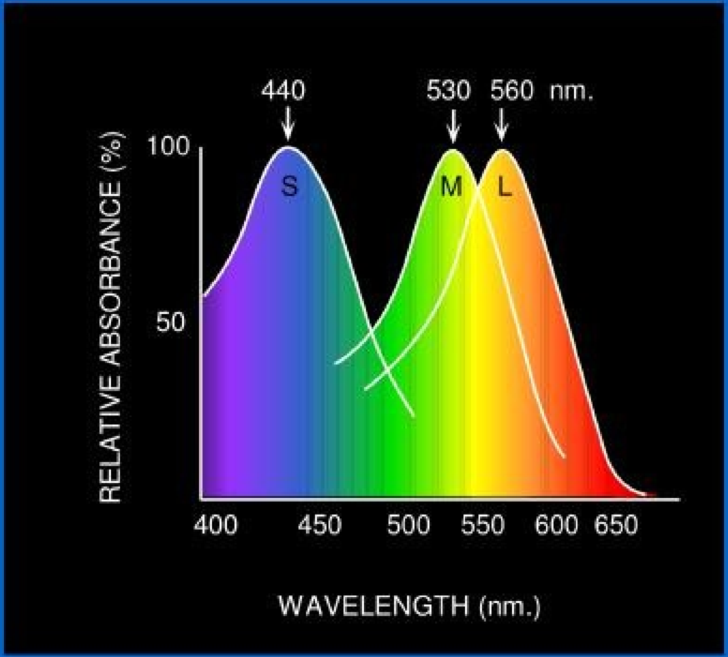
Human Eye (cont'd)

- Two kinds of light-sensitive cells: rods and cone (unevenly distributed).
- Cones (6 – 7 million) are responsible for all color vision and are present throughout the retina, but are concentrated toward the center of the field of vision at the back of the retina.
- Fovea – special area
  - Mostly cones.
  - Detail, color sensitivity, and resolution are highest.



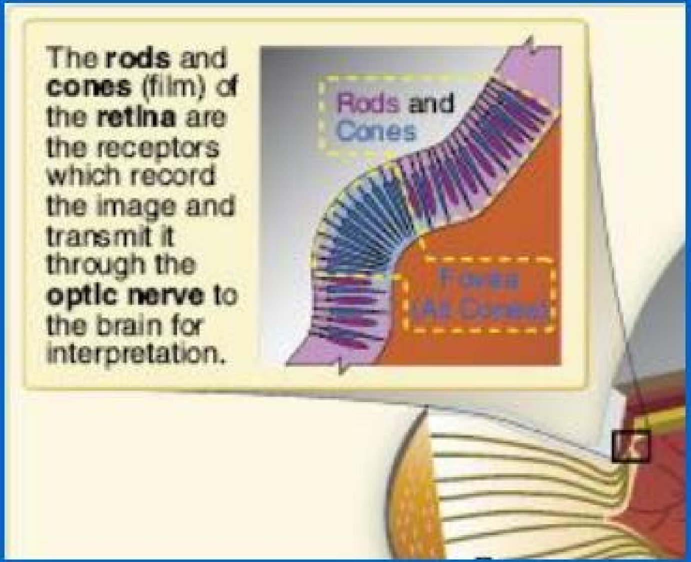
Human Eye (cont'd)

- Three different types of cones; each type has a special pigment that is sensitive to wavelengths of light in a certain range:
  - Short (S) corresponds to blue
  - Medium (M) corresponds to green
  - Long (L) corresponds to red
- Ratio of L to M to S cones:
  - approx. 10:5:1
- Almost no S cones in the center of the fovea



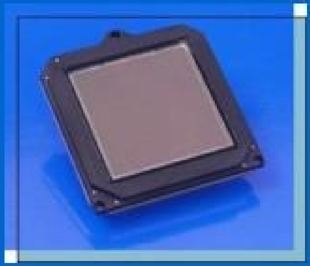
Human Eye (cont'd)

- Rods (120 million) more sensitive to light than cones but cannot discern color.
  - Primary receptors for <u>night vision</u> and <u>detecting motion</u>.
  - Large amount of light overwhelms them, and they take a long time to "reset" and adapt to the dark again.
  - Once fully adapted to darkness, the rods are 10,000 times more sensitive to light than the cones



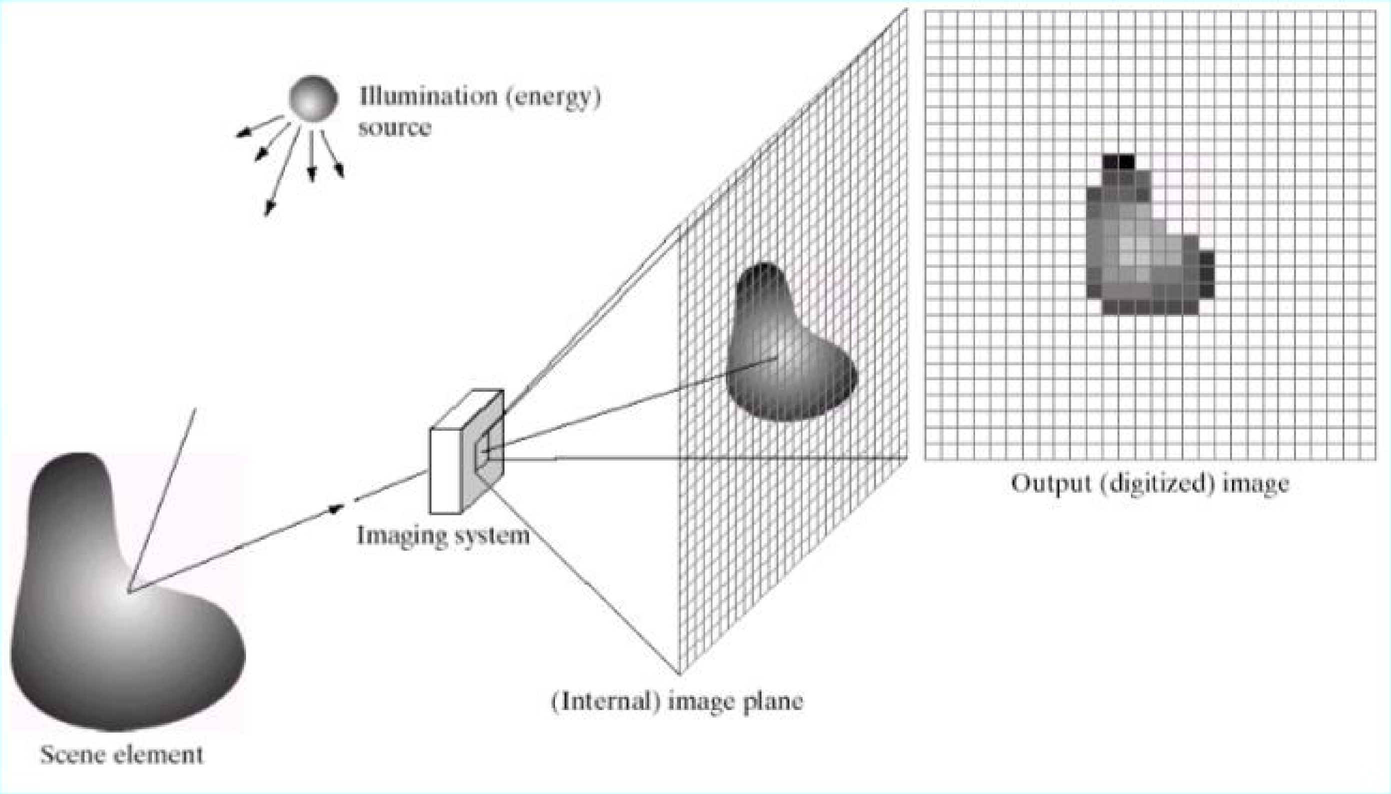
Digital cameras

- 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)



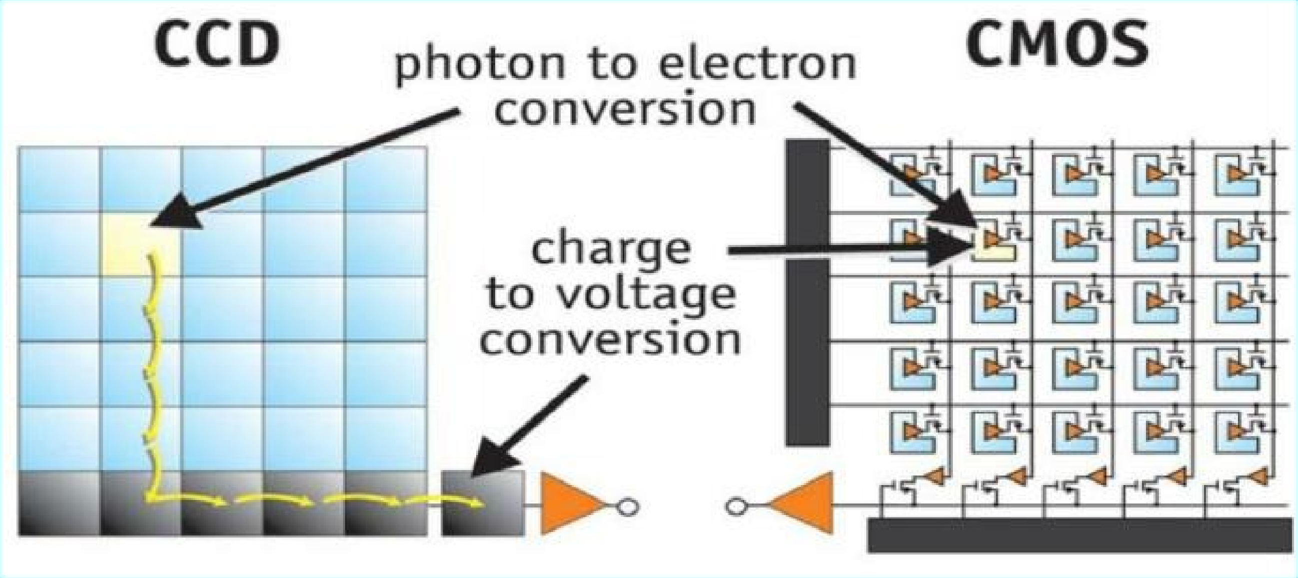
http://electronics.howstuffworks.com/digital-camera.htm

Digital cameras (cont'd)



CCD Cameras

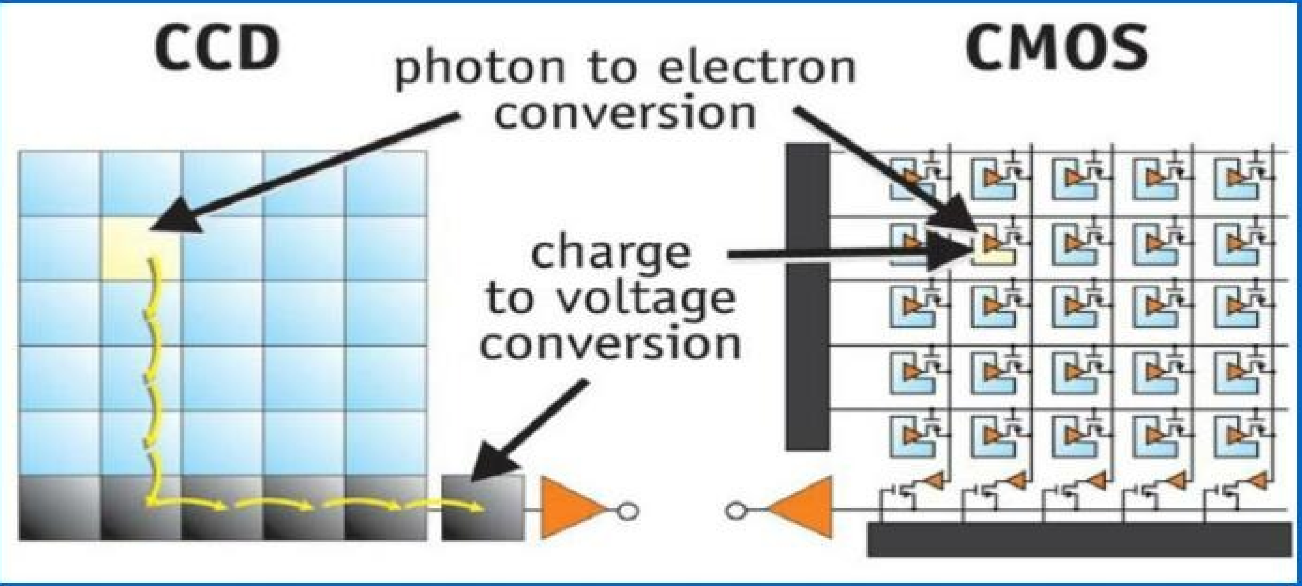
- CCDs move photogenerated charge from pixel to pixel and convert it to voltage at an output node.
- An **analog-to-digital converter (ADC)** then turns each pixel's value into a digital value.



http://www.dalsa.com/shared/content/pdfs/CCD_vs_CMOS_Litwiller_2005.pdf

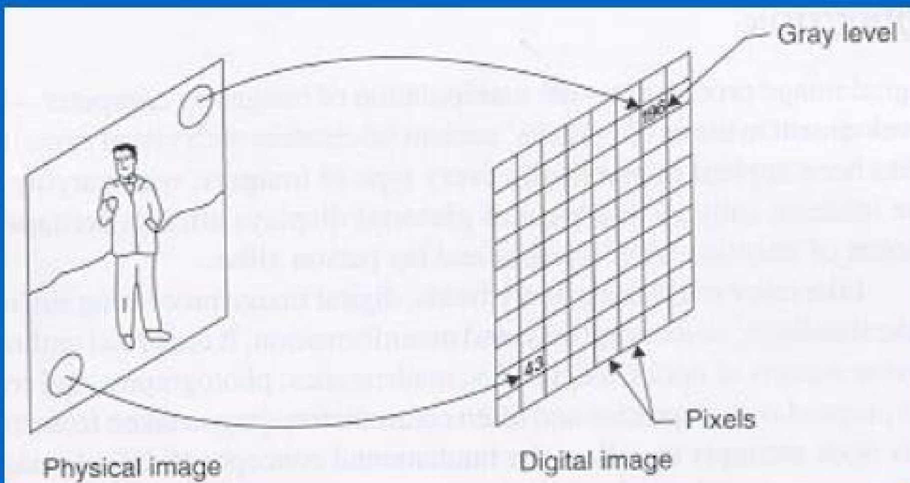
CMOS Cameras

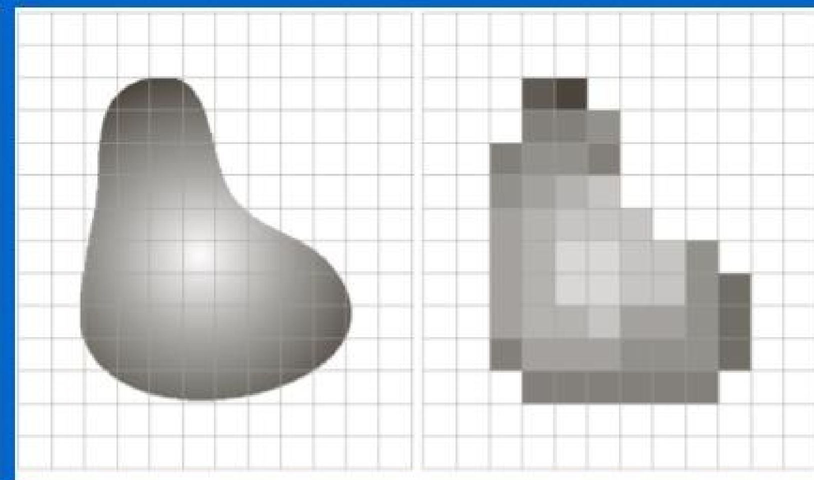
- CMOs convert charge to voltage inside each element.
- Uses several transistors at each pixel to amplify and move the charge using more traditional wires.
- The CMOS signal is digital, so it needs no ADC.



http://www.dalsa.com/shared/content/pdfs/CCD_vs_CMOS_Litwiller_2005.pdf

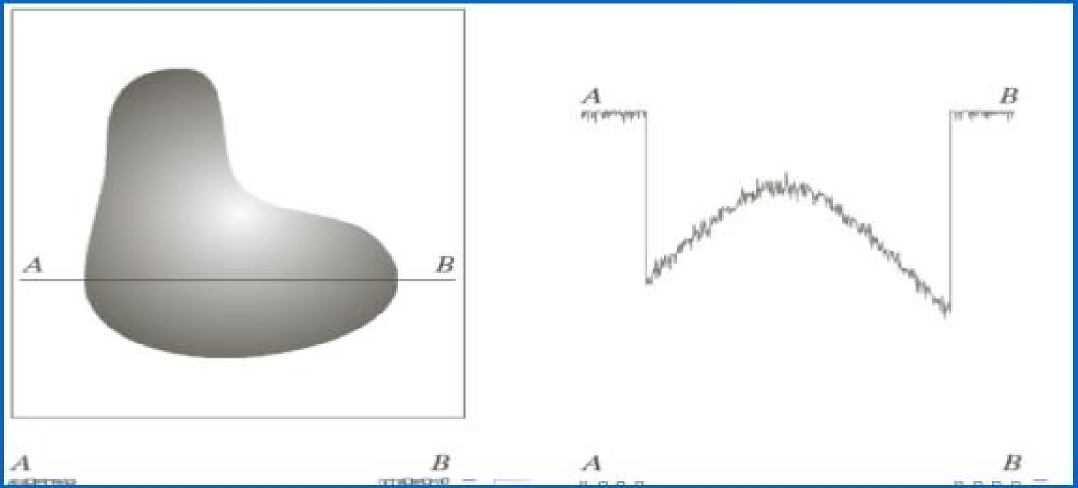
Image digitization



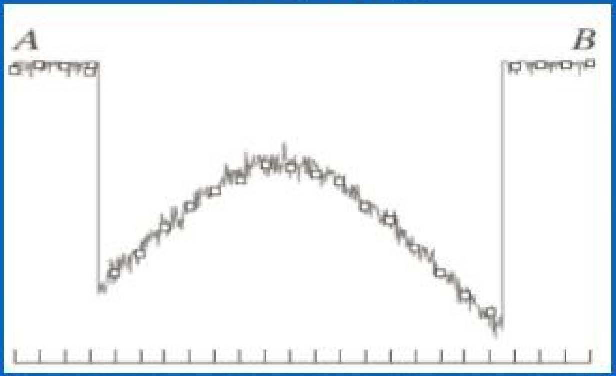


- **Sampling**: measure the value of an image at a finite number of points.
- **Quantization**: represent measured value (i.e., voltage) at the sampled point by an integer.

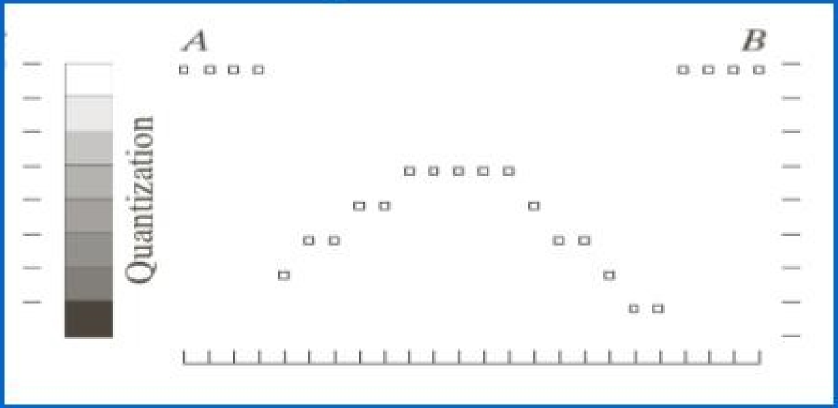
Image digitization (cont'd)



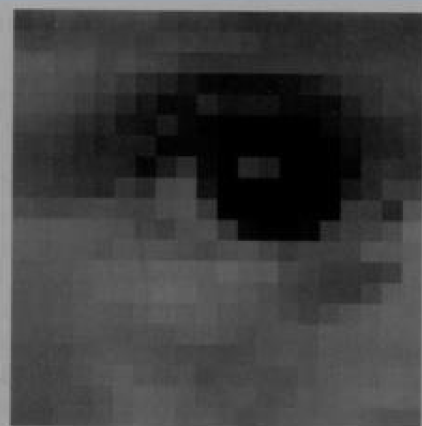
Sampling



Quantization

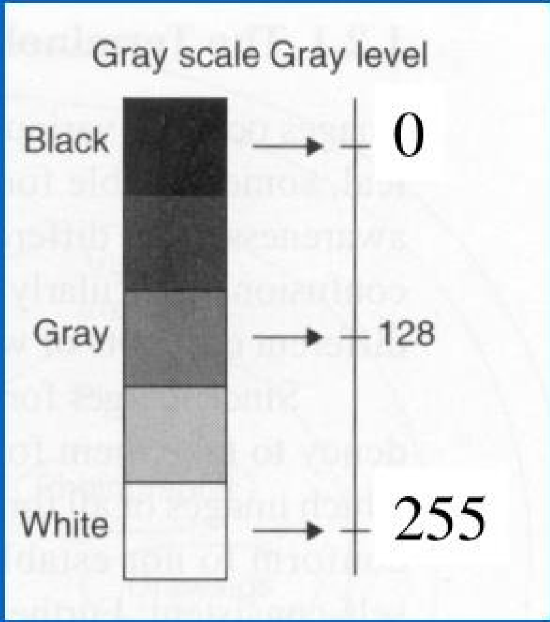


What is an image?



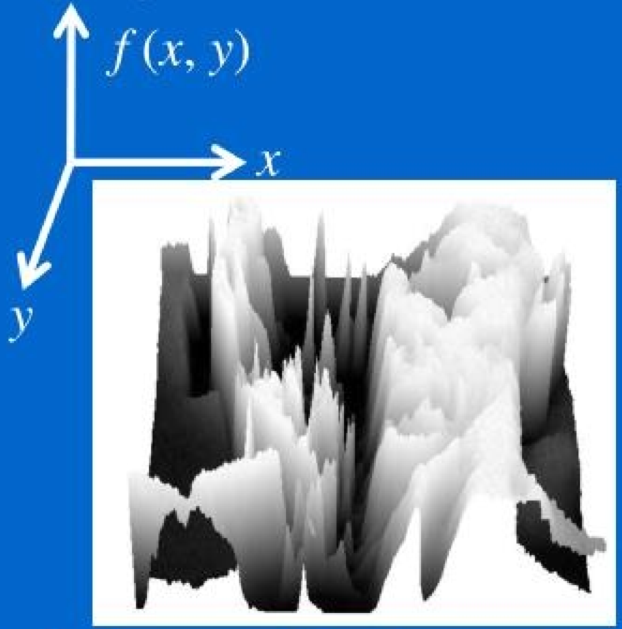
117	125	133	127	130	130	133	121	116	115	100	91	93	94	99	103	112	105	109	106
134	133	138	138	132	134	130	133	128	123	121	113	106	102	99	106	113	109	109	113
146	147	138	140	125	134	124	115	102	96	93	94	99	96	99	100	103	110	109	110
144	141	136	130	120	108	88	74	53	37	31	37	35	39	53	79	93	100	109	116
139	136	129	119	102	85	58	31	41	77	51	53	53	33	37	41	69	94	105	108
132	127	117	102	87	57	49	77	42	28	17	15	13	13	17	41	53	69	88	100
124	120	108	94	72	74	72	31	35	31	15	13	15	11	15	13	46	75	83	96
125	115	102	93	88	82	42	79	113	41	19	100	82	11	11	17	31	91	99	100
124	116	109	99	91	113	99	140	144	57	20	20	15	11	15	17	63	87	119	124
136	133	133	135	138	133	132	144	150	120	24	17	15	15	17	20	115	113	88	150
158	157	157	154	149	145	133	127	146	150	116	35	20	19	28	105	124	128	141	171
155	154	156	155	146	155	154	154	147	139	148	150	138	120	128	129	130	151	156	165
150	151	154	162	166	167	169	174	172	167	177	166	164	140	134	120	121	120	127	172
145	149	151	157	165	169	173	179	176	166	166	157	145	136	129	124	120	136	163	168
144	148	153	160	159	158	165	172	165	169	157	151	149	141	130	140	151	162	169	167
144	141	147	155	154	149	156	151	157	157	151	144	147	147	149	159	158	159	166	165
139	140	140	150	153	151	150	146	140	139	138	140	145	151	149	156	156	162	162	161
136	134	138	146	156	164	153	146	145	136	139	139	140	141	149	157	159	161	169	166
136	133	136	135	144	159	168	159	151	142	141	145	139	146	153	156	164	167	172	168
133	129	140	142	146	159	167	165	154	151	146	141	147	154	156	160	161	157	153	154

8 bits/pixel



What is an image? (cont'd)

- We can think of a (grayscale) image as a **function**, *f*, from R^2 to R (or a 2D *signal*):
  - $f(x,y)$ gives the **intensity** at position (x,y)



  - A **digital** image is a discrete (**sampled**, **quantized**) version of this function

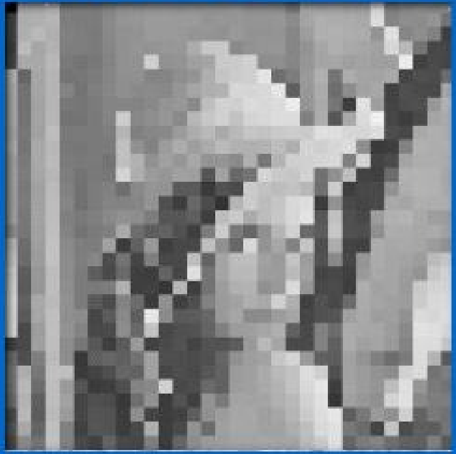
Image Sampling - Example

original image

sampled by a factor of 2

sampled by a factor of

sampled by a factor of 8



Images have been resized for easier comparison

Image Quantization - Example

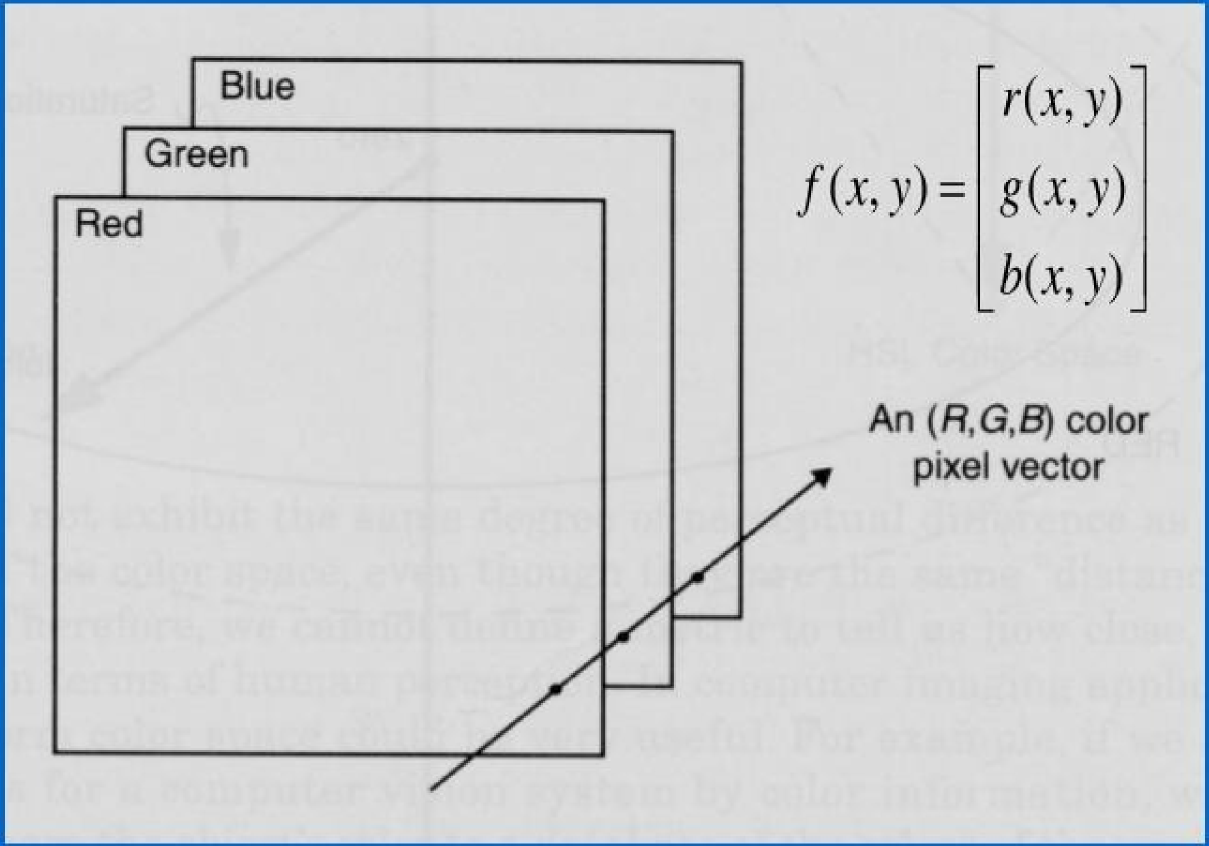
- 256 gray levels (8bits/pixel) 32 gray levels (5 bits/pixel) 16 gray levels (4 bits/pixel)

- 8 gray levels (3 bits/pixel) 4 gray levels (2 bits/pixel) 2 gray levels (1 bit/pixel)

Color Images

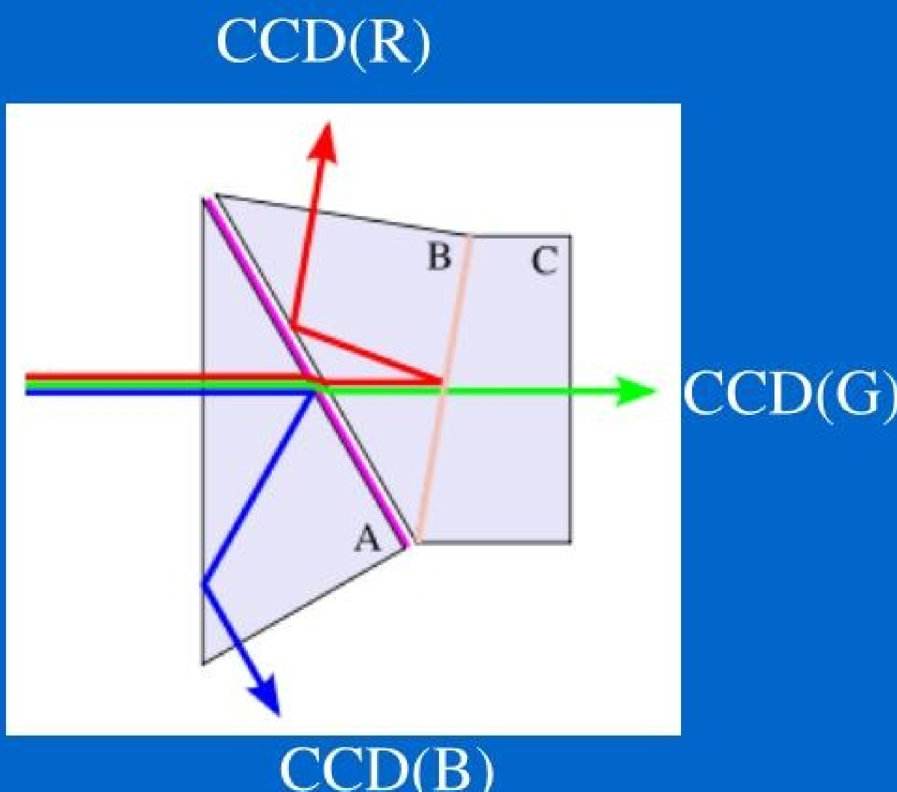
- Color images are comprised of three color channels – red, green, and, blue – which combine to create most of the colors we can see.

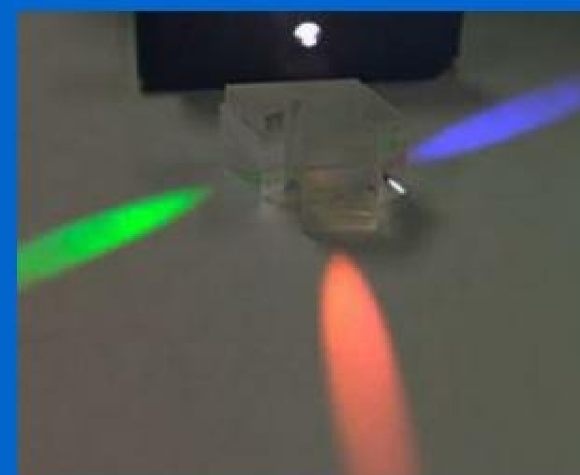
Color images



Color sensing in camera: Prism

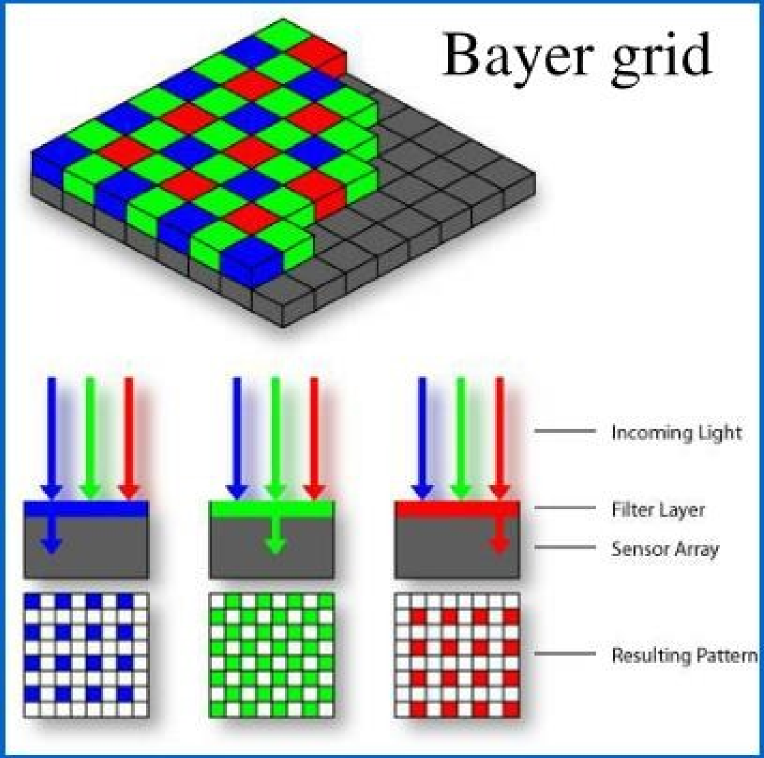
- Requires three chips and precise alignment.



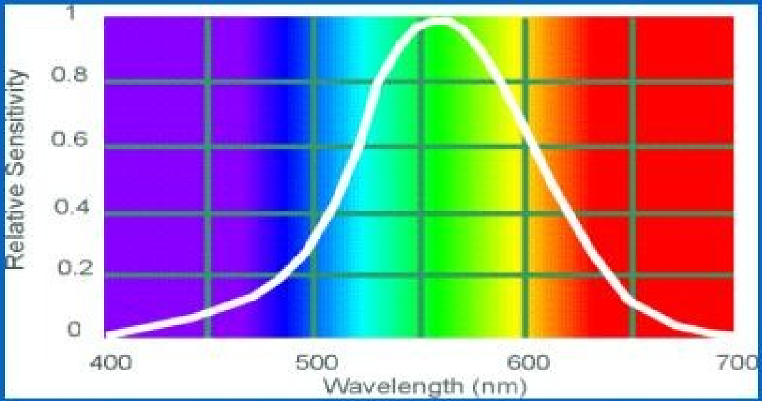


Color sensing in camera: Color filter array

- In traditional systems, color filters are applied to a single layer of photodetectors in a tiled mosaic pattern.



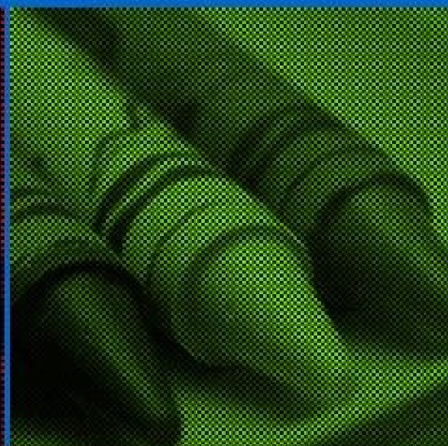
Why more green?



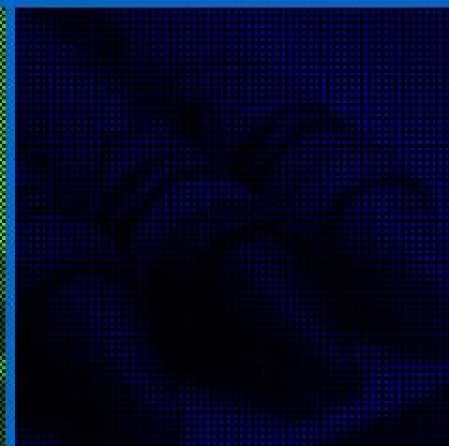
Human Luminance Sensitivity Function

Color sensing in camera: Color filter array

red



green



blue

output

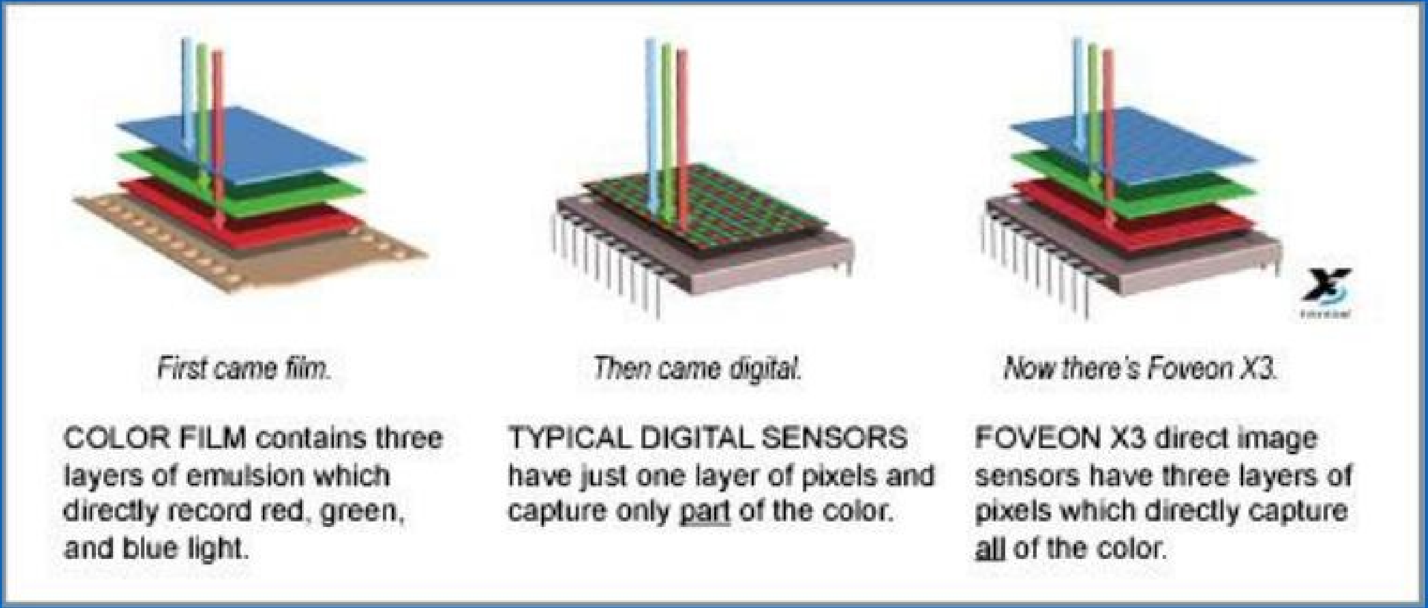
G	R	G	R
B	G	B	G
G	R	G	R
B	G	B	G

demosaicing
(interpolation)

rGb	Rgb	rGb	Rgb
rgB	rGb	rgB	rGb
rGb	Rgb	rGb	Rgb
rgB	rGb	rgB	rGb

Color sensing in camera: Foveon X3

• CMOS sensor; takes advantage of the fact that red, blue and green light silicon to different depths.



http://www.foveon.com/article.php?a=67

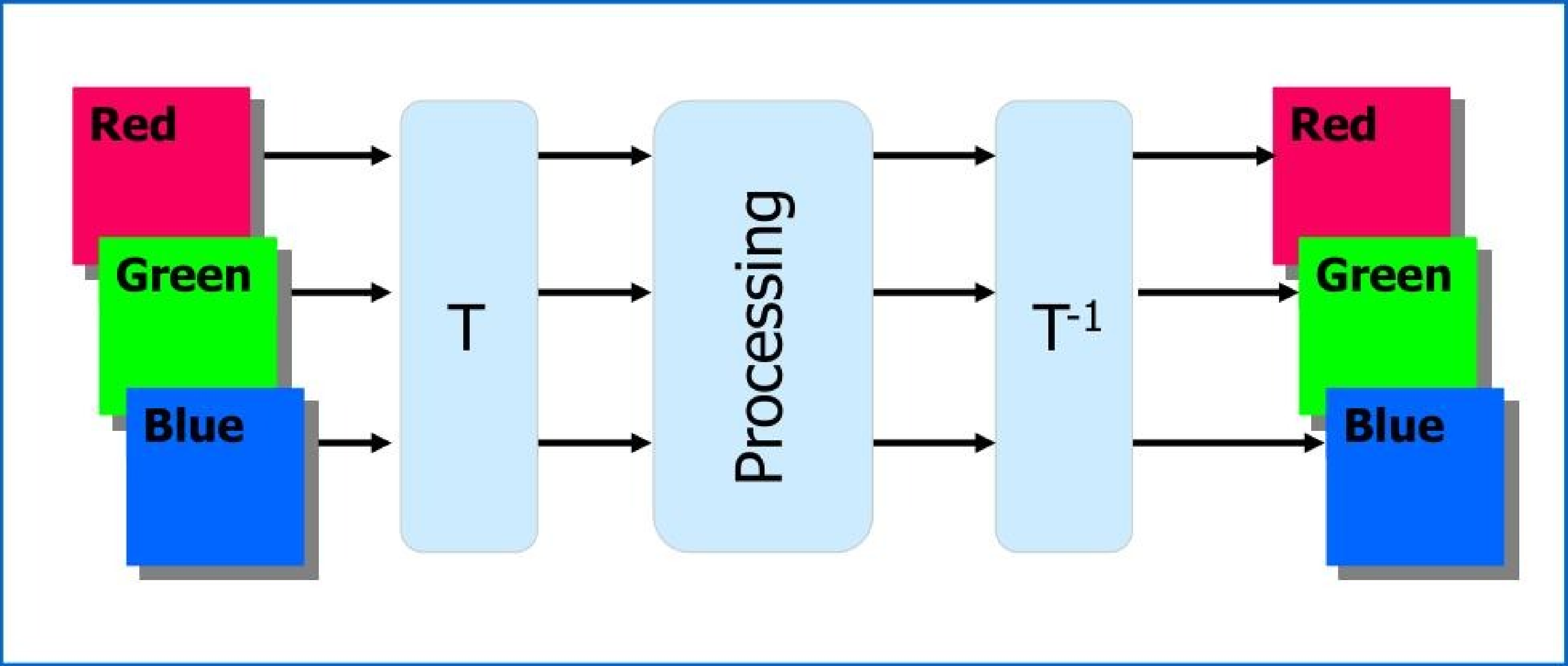
Alternative Color Spaces

- Various other color representations can be computed from RGB.
- This can be done for:
  - Decorrelating the color channels:
    - principal components.
  - Bringing color information to the fore:
    - Hue, saturation and brightness.
  - Perceptual uniformity:
    - CIELuv, CIELab, …

Alterative Color paces

- RGB (CIE), RnGnBn (TV - National Television Standard Committee)
- XYZ (CIE)
- UVW (UCS de la CIE), U*V*W* (UCS modified by the CIE)
- YUV, YIQ, YCbCr
- YDbDr
- DSH, HSV, HLS, IHS
- Munsel color space (cylindrical representation)
- CIELuv
- CIELab
- SMPTE-C RGB
- YES (Xerox)
- Kodak Photo CD, YCC, YPbPr, ...

Processing Strategy



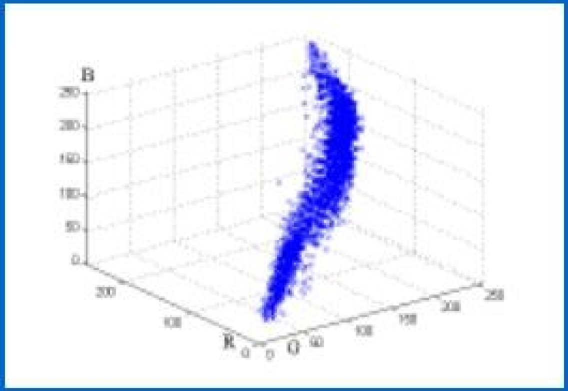
Color Transformation - Examples

$$
\begin{aligned}
H &= \arccos \frac{\frac{1}{2}((R-G)+(R-B))}{\sqrt{((R-G)^2+(R-B)(G-B))}} \\
S &= 1-3\frac{\min(R,G,B)}{R+G+B} \\
V &= \frac{1}{3}(R+G+B)
\end{aligned}
$$

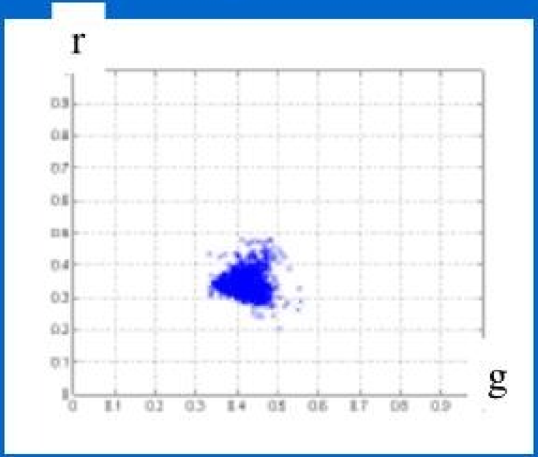
$$
\begin{aligned}
&Y = 0.299R + 0.587G + 0.114B \\
&C_r = R - Y \\
&C_b = B - Y
\end{aligned}
$$

Skin color

RGB

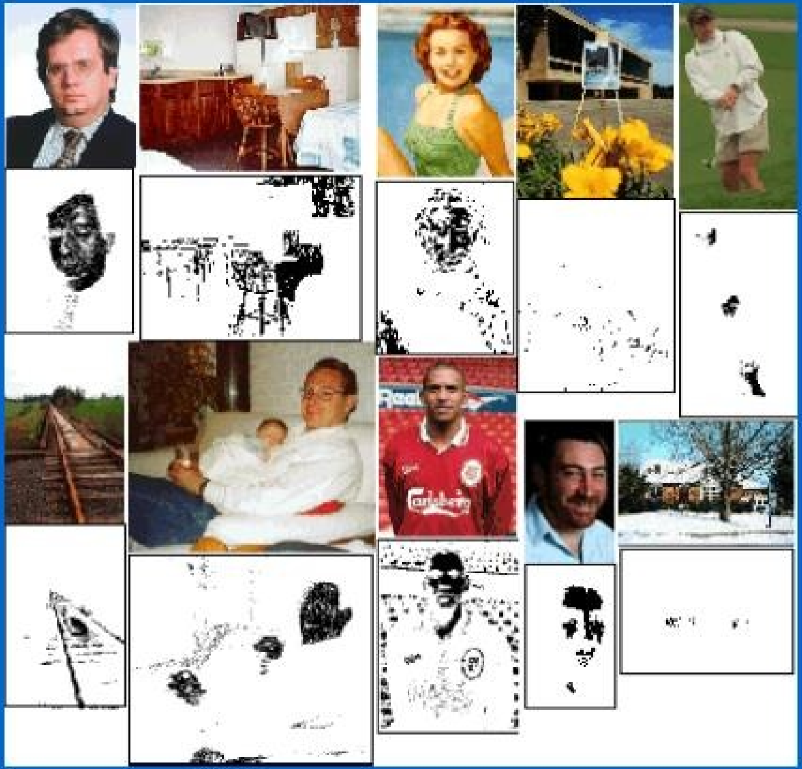


rg



$$r = \frac{R}{R+G+B} \quad g = \frac{G}{R+G+B} \quad b = \frac{B}{R+G+B}$$

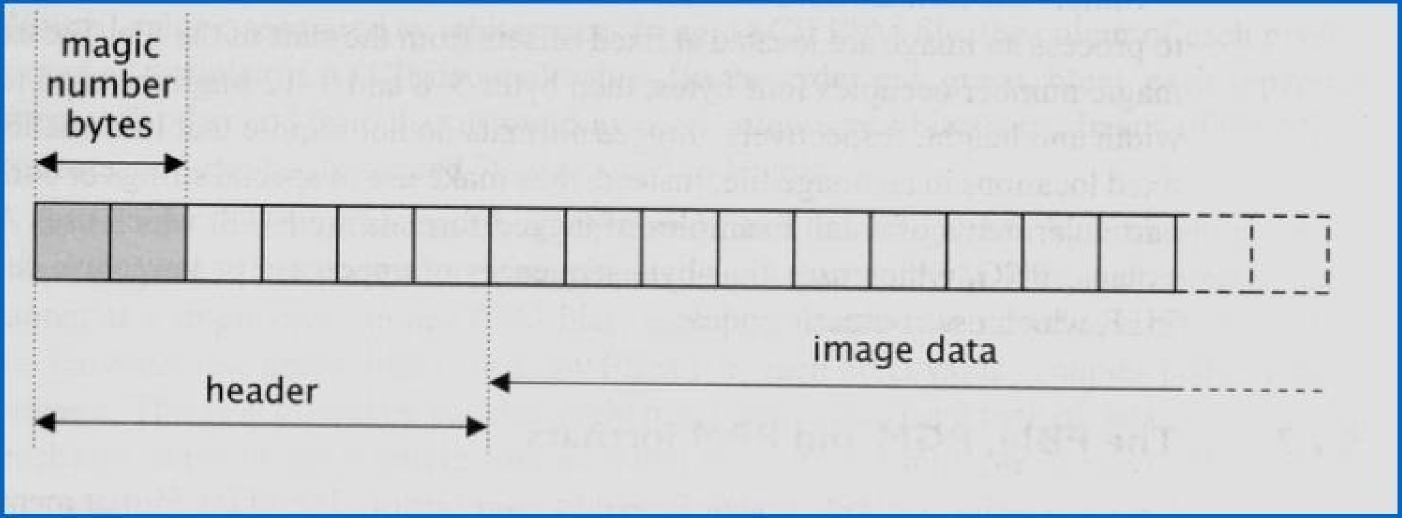
Skin detection



M. Jones and J. Rehg, Statistical Color Models with Application to Skin Detection, International Journal of Computer Vision, 2002.

Image file formats

- Many image formats adhere to the simple model shown below (line by line, no breaks between lines).
- The header contains at least the width and height of the image.
- Most headers begin with a **<u>signature</u>** or "magic number" (i.e., a short sequence of bytes for identifying the file format)



Common image file formats

- GIF (Graphic Interchange Format) -
- PNG (Portable Network Graphics)
- JPEG (Joint Photographic Experts Group)
- TIFF (Tagged Image File Format)
- PGM (Portable Gray Map)
- FITS (Flexible Image Transport System)

PBM/PGM/PPM format

- A popular format for grayscale images (8 bits/pixel)
- Closely-related formats are:
  - PBM (Portable Bitmap), for binary images (1 bit/pixel)
  - PPM (Portable Pixelmap), for color images (24 bits/pixel)

ASCII or binary (raw) storage

ASCI

```
P2
# a simple PGM image
7 7 255
 120 120 120 120 120 120 120
 120 120 120  33 120 120 120
 120 120 120  33 120 120 120
 120  33  33  33  33  33 120
 120 120 120  33 120 120 120
 120 120 120  33 120 120 120
 120 120 120 120 120 120 120
```

Binary

```
P5
# a simple PGM image
7 7 255
xxxxxxxxxx!xxxxxx!xxxx!!!!!xxxx!xxxxxx!xxxxxxxxxx
```

Signatures of the various PBM, PGM and PPM image formats.

Signature	Image type	Storage type
P1	binary	ASCII
P2	greyscale	ASCII
P3	RGB	ASCII
P4	binary	raw bytes
P5	greyscale	raw bytes
P6	RGB	raw bytes