Introduction to Computer Vision

CS-484 & CS-555

Sedat Ozer
Department of Computer Engineering
Bilkent University
sedat@cs.bilkent.edu.tr

Last Week

- Introduction to (Computer) Vision Applications
  - introduction to the field & to the course
  - an overview of various vision applications

Announcement

- Good news everyone: Besides Google and Bing, now you have a new TA:

- Alexa or Siri

  - Teaching Assistant: Mr. Aydamir Mirzayev
    - Office: EA505

Projects and Papers

- 2019, 2018 CVPR, ICCV or ECCV conferences
- http://openaccess.thecvf.com/CVPR2019.py
- Check Paper titles first,
- Pick a couple to read.
  - Check what dataset they use,
  - Do they have a github code? If not, ignore!

- Need a GPU to run the GITHUB code typically.

Test: what do you see in this image?

Image source: Antonio Torralba

What do you see in this image?

Now can you see it?

Credit: *Thompson, Basic Vision, Oxford Press, 2012.*

Subjective contours

Question:

How would you model (represent) an object in visual recognition? Ideas?

Remember the image:

01010100 01101000 01101001 01110011
00100000 01101001 01110011 00100000
01110100 01101000 01100101 00100000
01110100 01110101 01110100 01101111
01110010 01101001 01100001 01101100
00100000 01110100 01101111 00100000
01101100 01100101 01100001 01110010
01101110 00100000 01100010 01101001
01101110 01100001 01110010 01111001
00101110 00100000 01001001 00100000
01101000 01101111 01110000 01100101
00100000 01111001 01101111 01110101
00100000 01100101 01101110 01101010
01101111 01111001 00100000 01101001
01110100 00100001

This is what computers get (see?)!

This is what you see…

Our common units: Pixels

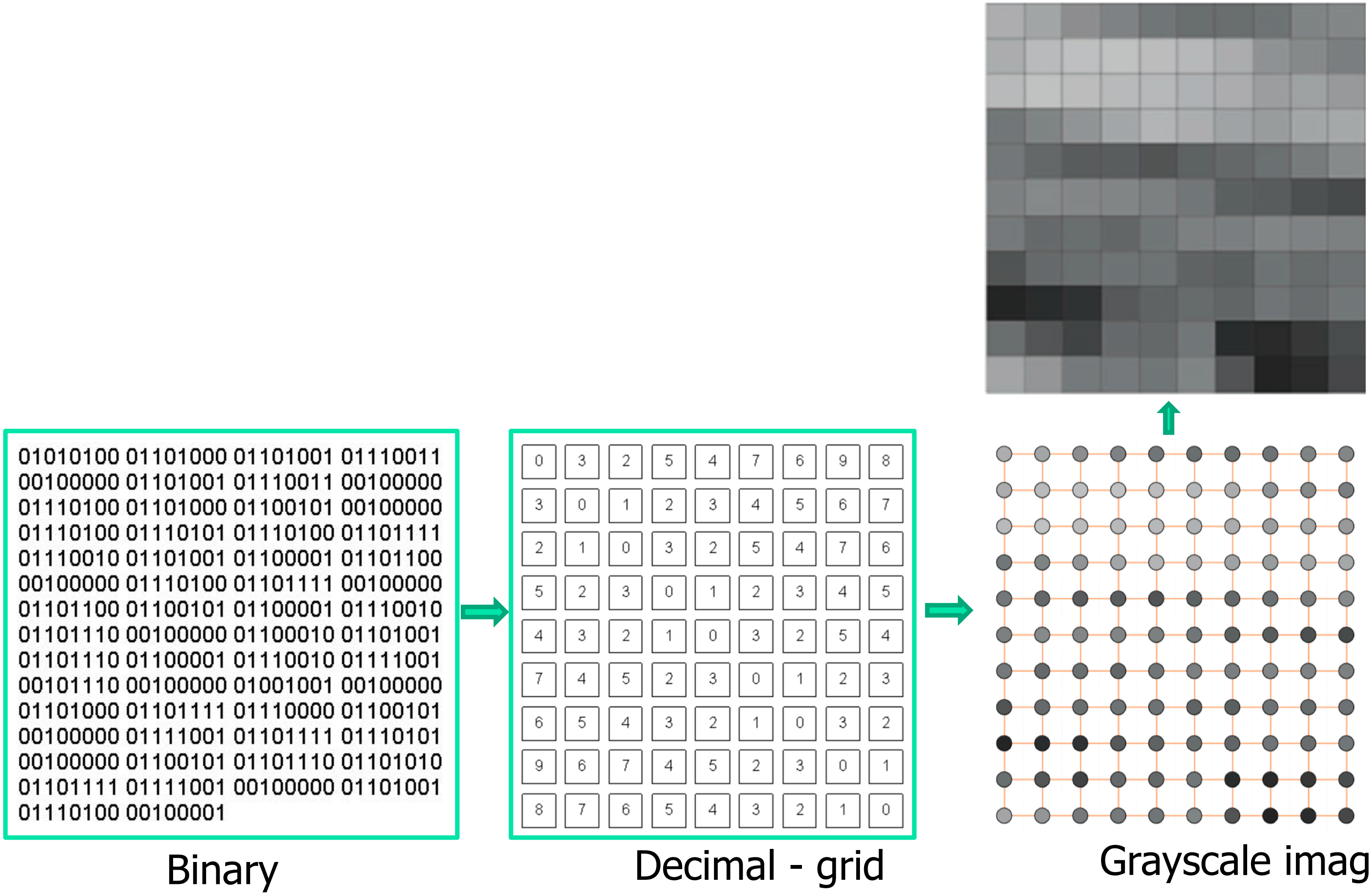
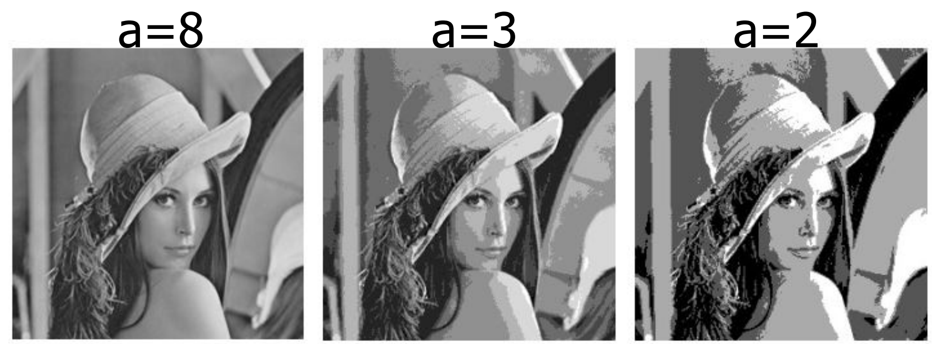


Image Types: (Gray)Scalar and Binary



Original (256 colors) | 8 colors | 4 colors

- A scalar image has integer values

$$u \in \{0, 1, ..., 2^a - 1\}$$

a: level (bit)

Ex. If 8 bit (a=8), image spans from 0 to 255

0 black

255 white

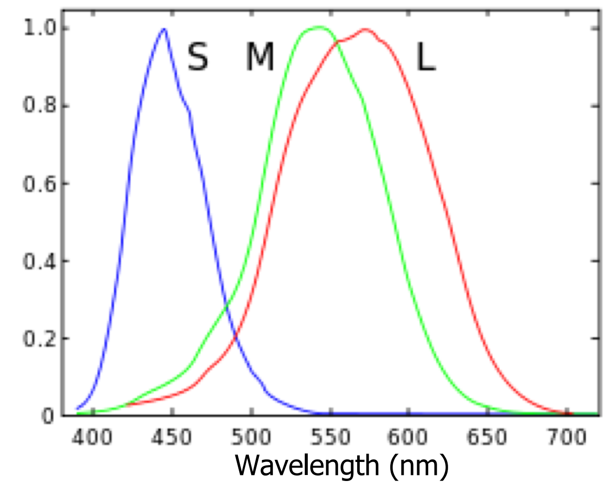
Ex. If 1 bit (a=1), it is binary image, 0 and 1 only.

Image Type: RGB (red, green, blue)

- Image has three channels (bands), each channel spans a-bit values.

R G B

Human Cone-cells (normalized) responsivity spectra



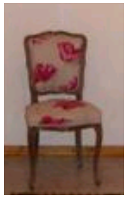
(Some people might have 4 cone-types!)

So…. How do we detect an object in an image?

ANY IDEAS ???

Naïve approach: Template Matching

This is a chair



Find the chair in this image

Output of correlation

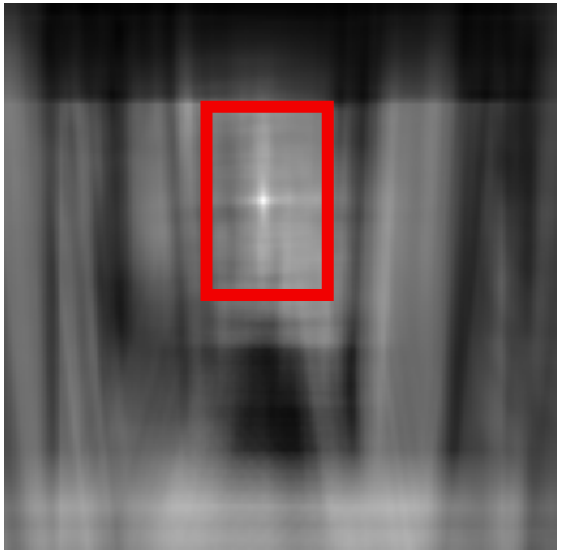
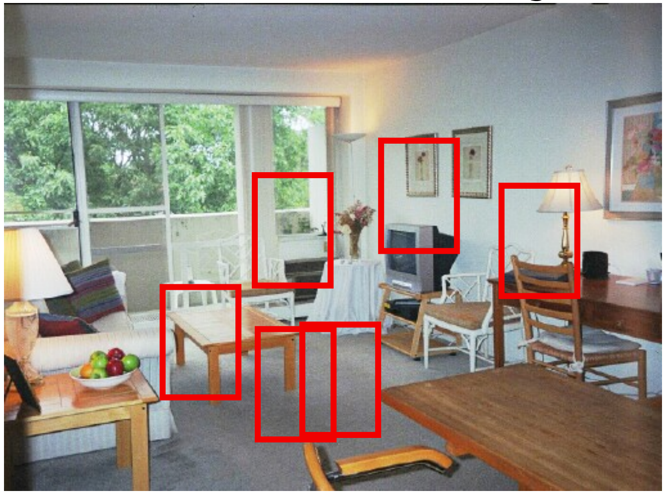
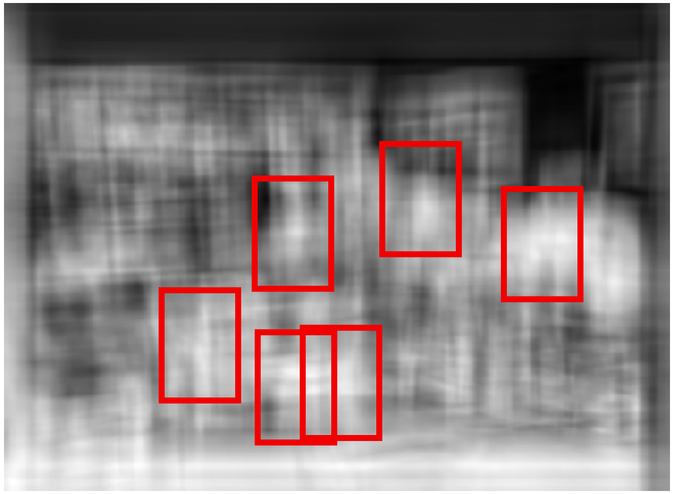


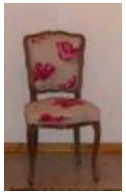
Image source: Antonio Torralba

Template Matching

Find the chair in this image







Epic fail!
Simple template matching is not going to make it

Image source: Antonio Torralba

Idea:

- Instead of comparing raw image pixels:
  - first map those pixels into another (more robust) form,
  - and then compare those mapped forms.
  - Finally, select the closest image map (how do you define “closest”? Metrics).
- Features
  - Examples: compute edges, compute color histograms, Gradients, HOG, SIFT, …

"Bag-of-Words" Representation

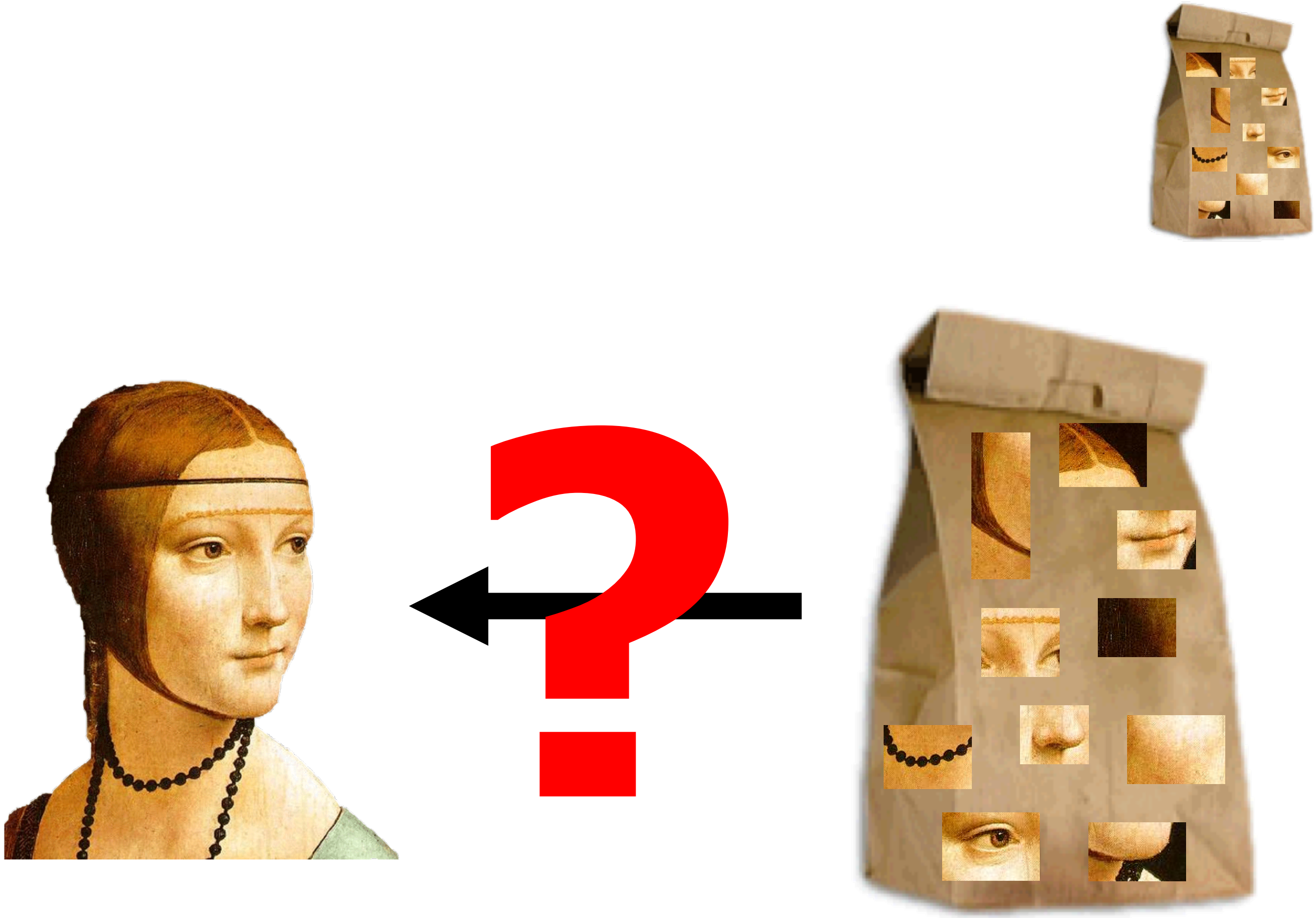
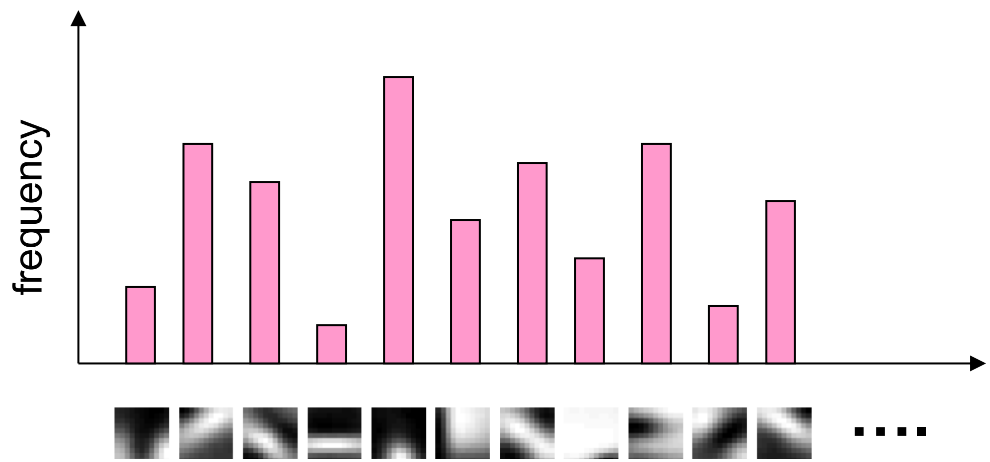


Image source: Antonio Torralba

"Bag-of-Words" (BoW) Histograms

BoW Histogram

Image



codewords

Image Features → BoW Histogram → Classifier → Result

Recipe

BoW Representation

- All have equal probability for bag-of-words methods,
- Location (spatial) information is important but lost.

Image source: Antonio Torralba

QUESTIONS?

Stages of computer vision

- Low-level

  image → image

- Mid-level

  image → features / attributes

- High-level

  features → “making sense”, recognition

Low-level

original image

Canny

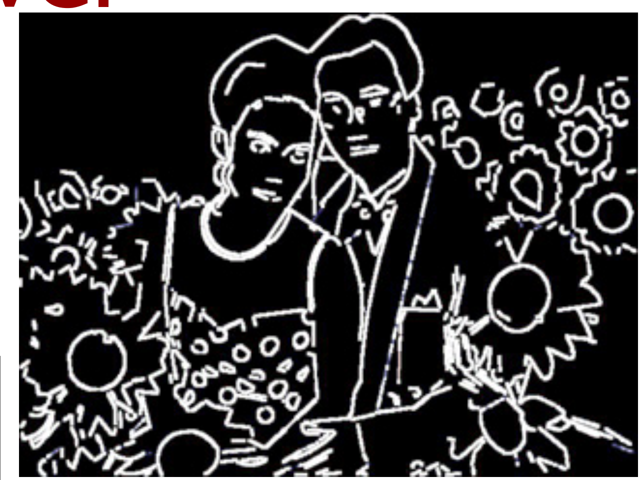
edge image

Mid-level

edge image

ORT

data structure



circular arcs and line segments

Adapted from Linda Shapiro, U of Washington

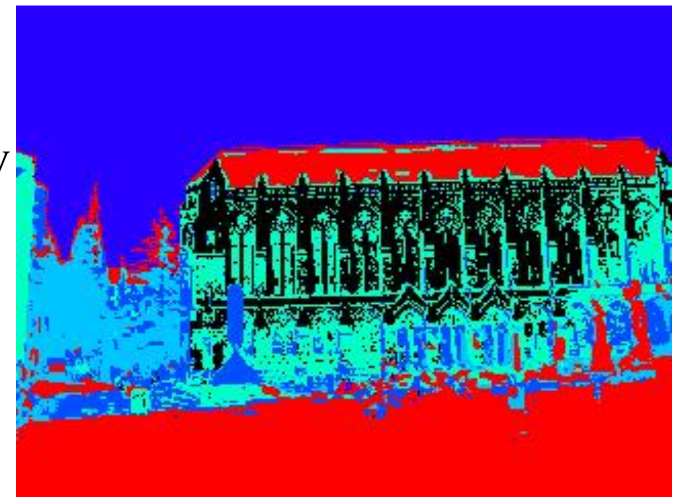
Mid-level

original color image

K-means clustering

(followed by connected component analysis)

data structure



regions of homogeneous color

Adapted from Linda Shapiro, U of Washington

Low-level to high-level

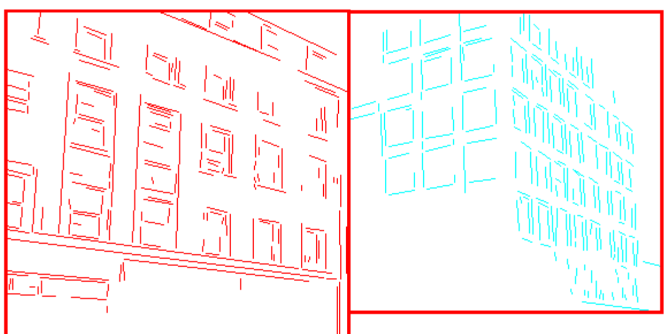
low-level

edge image

mid-level

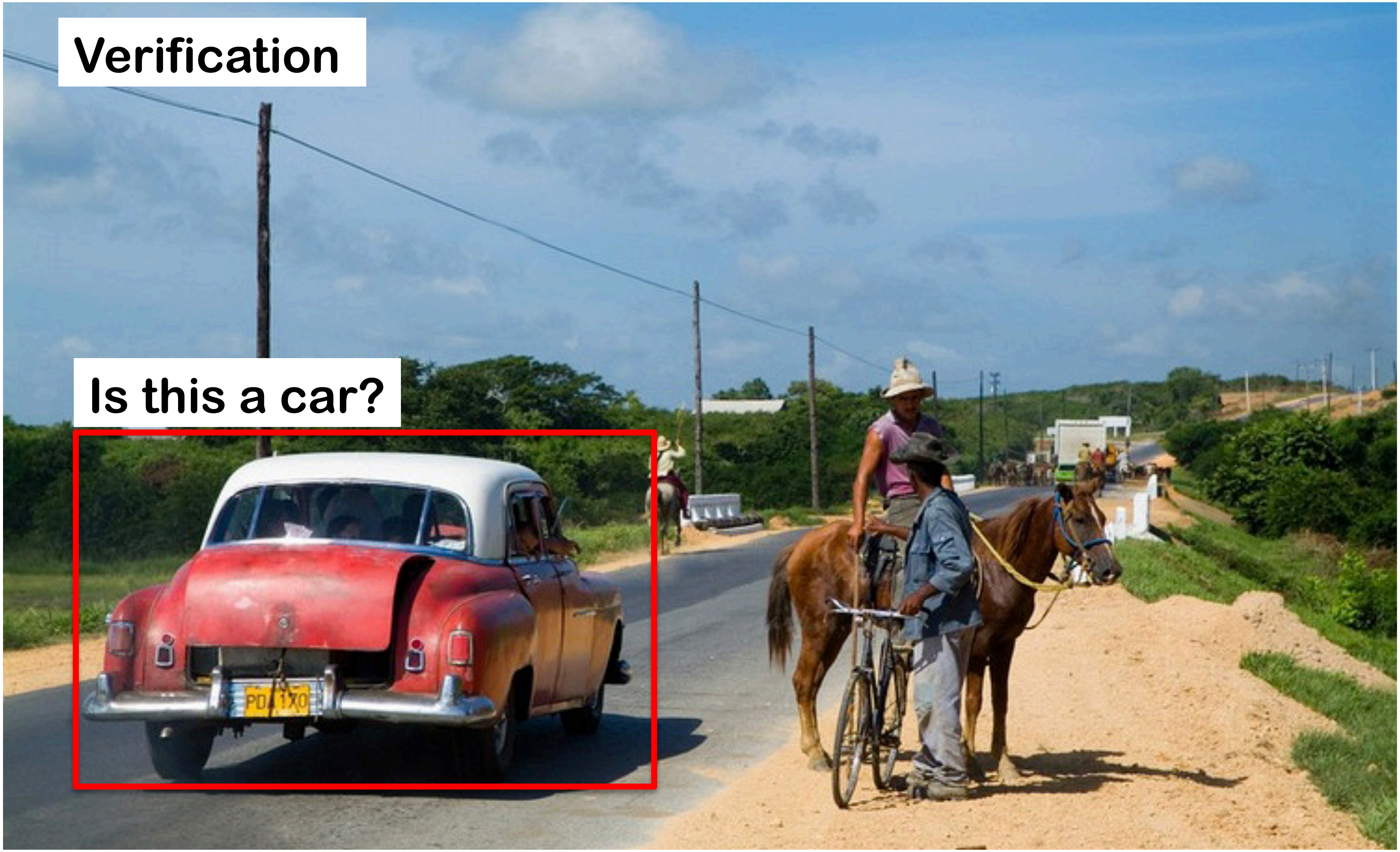
consistent line clusters

high-level



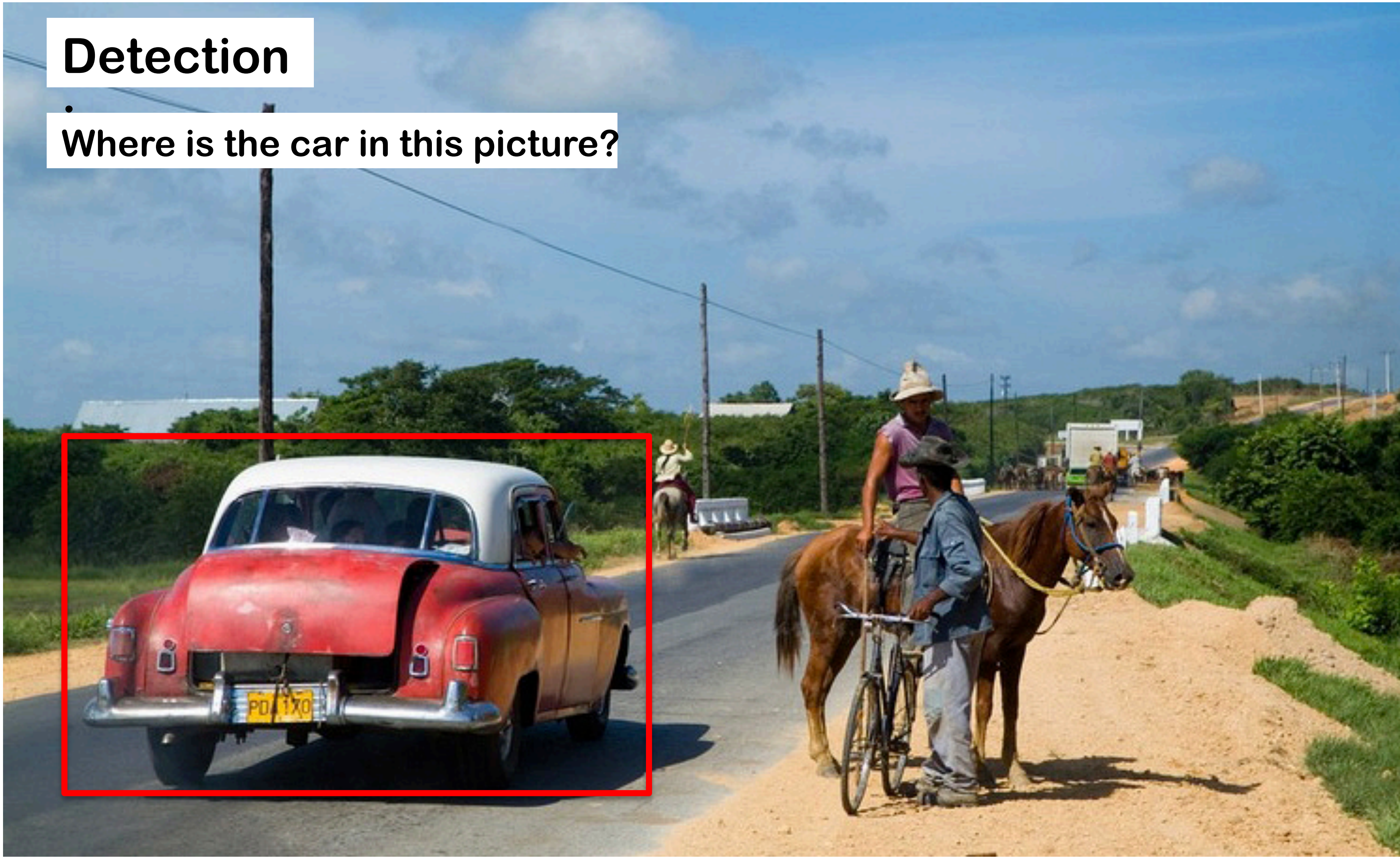
Adapted from Linda Shapiro, U of Washington

Visual recognition

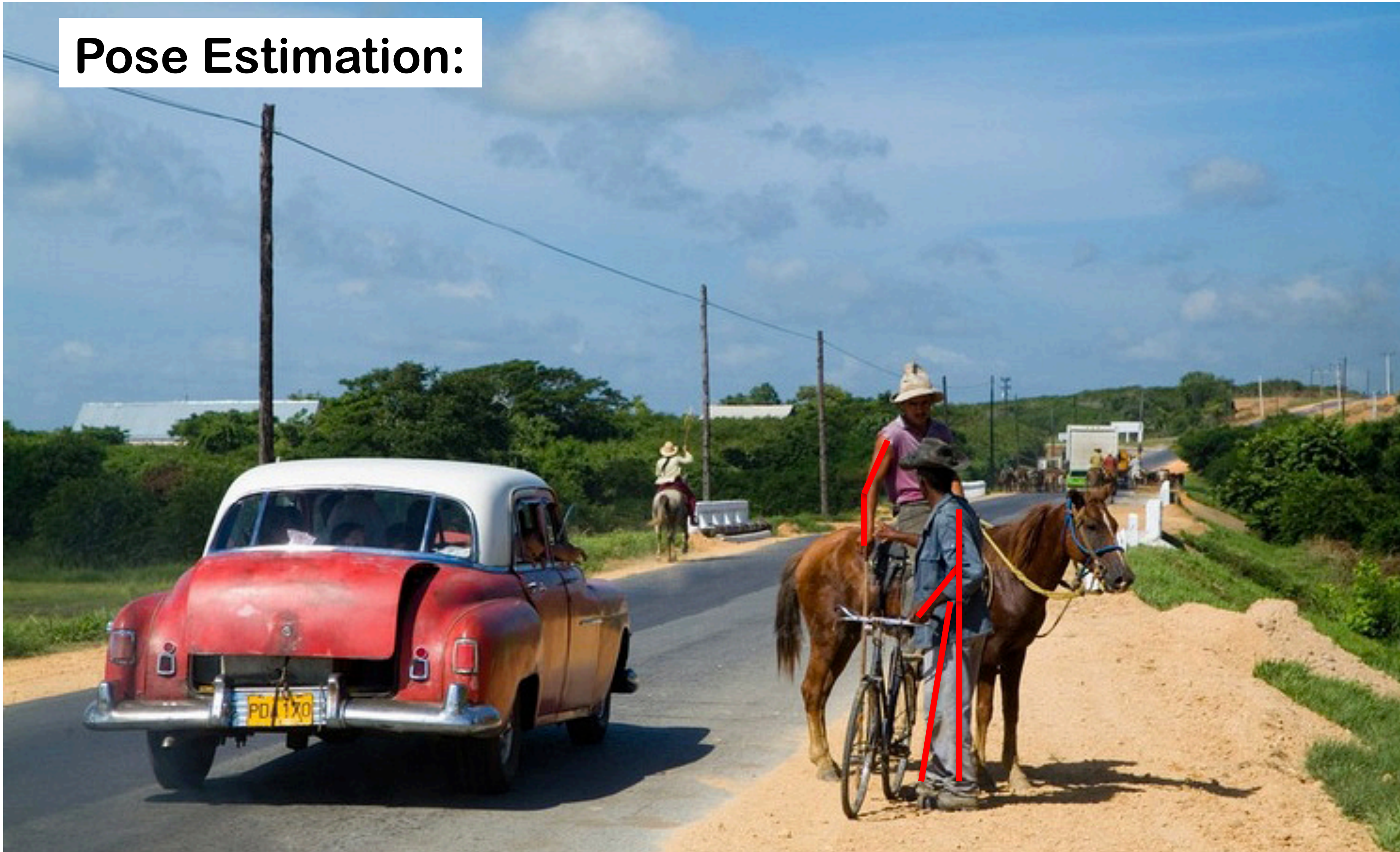


Visual recognition

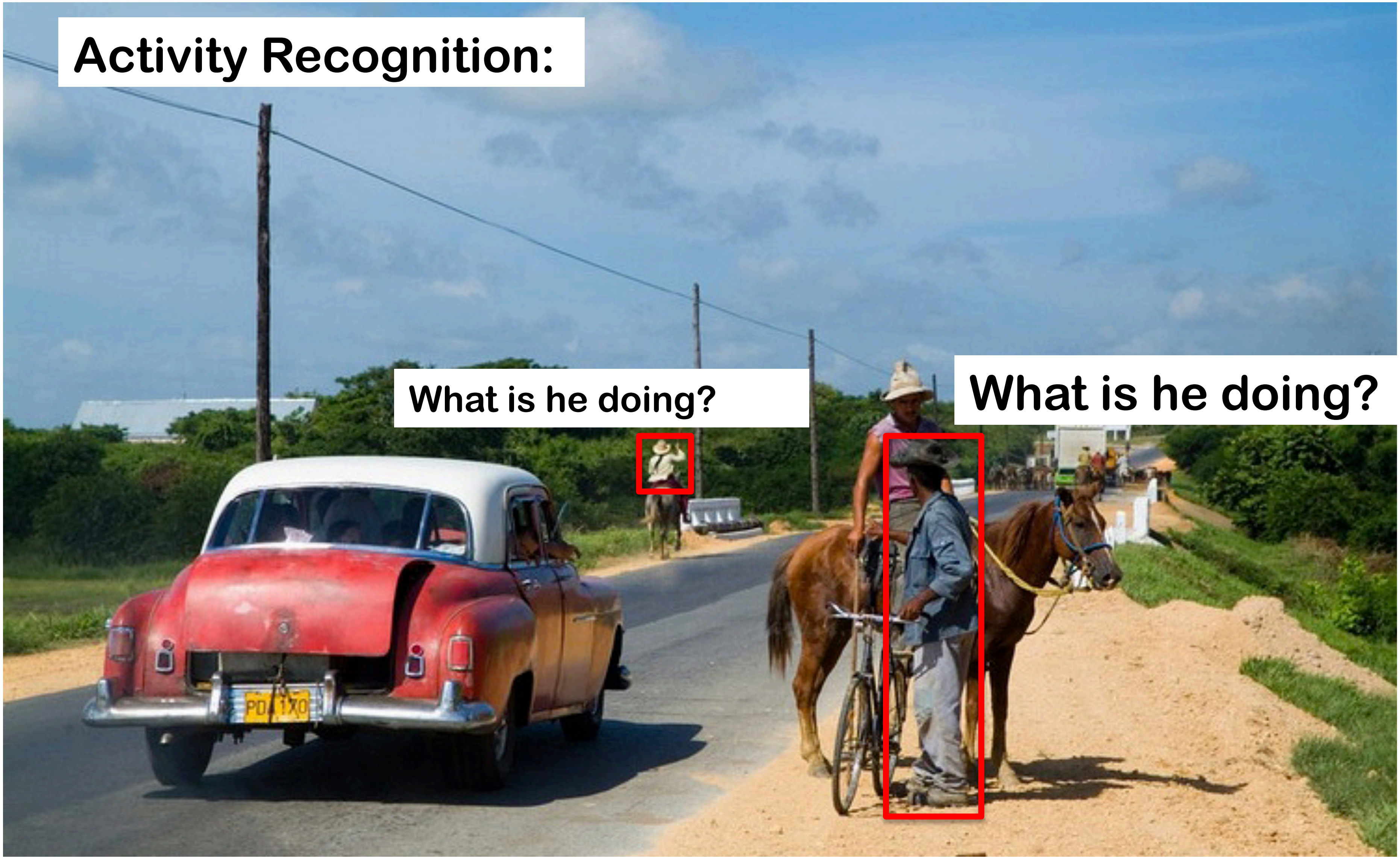
Visual recognition



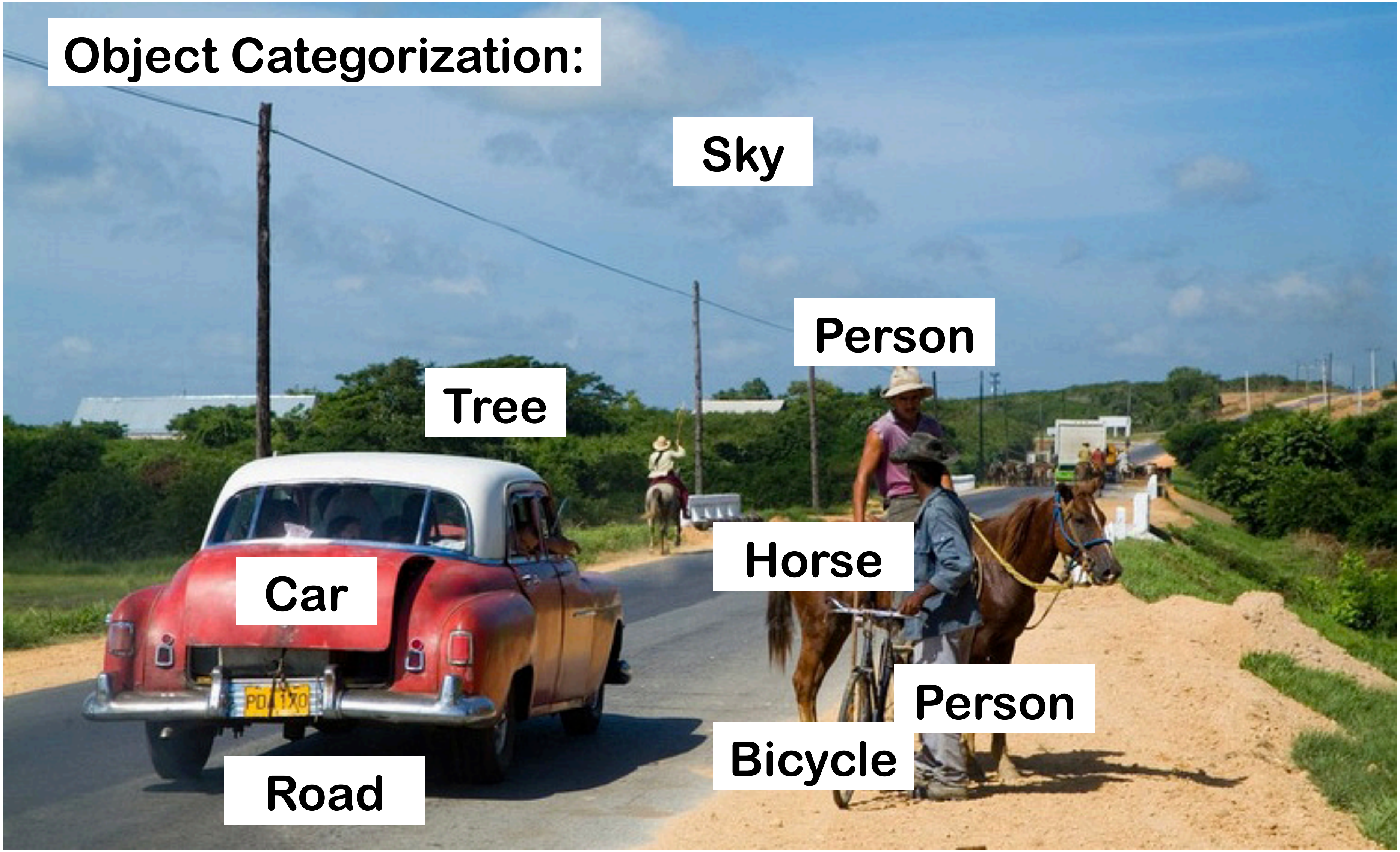
Visual recognition



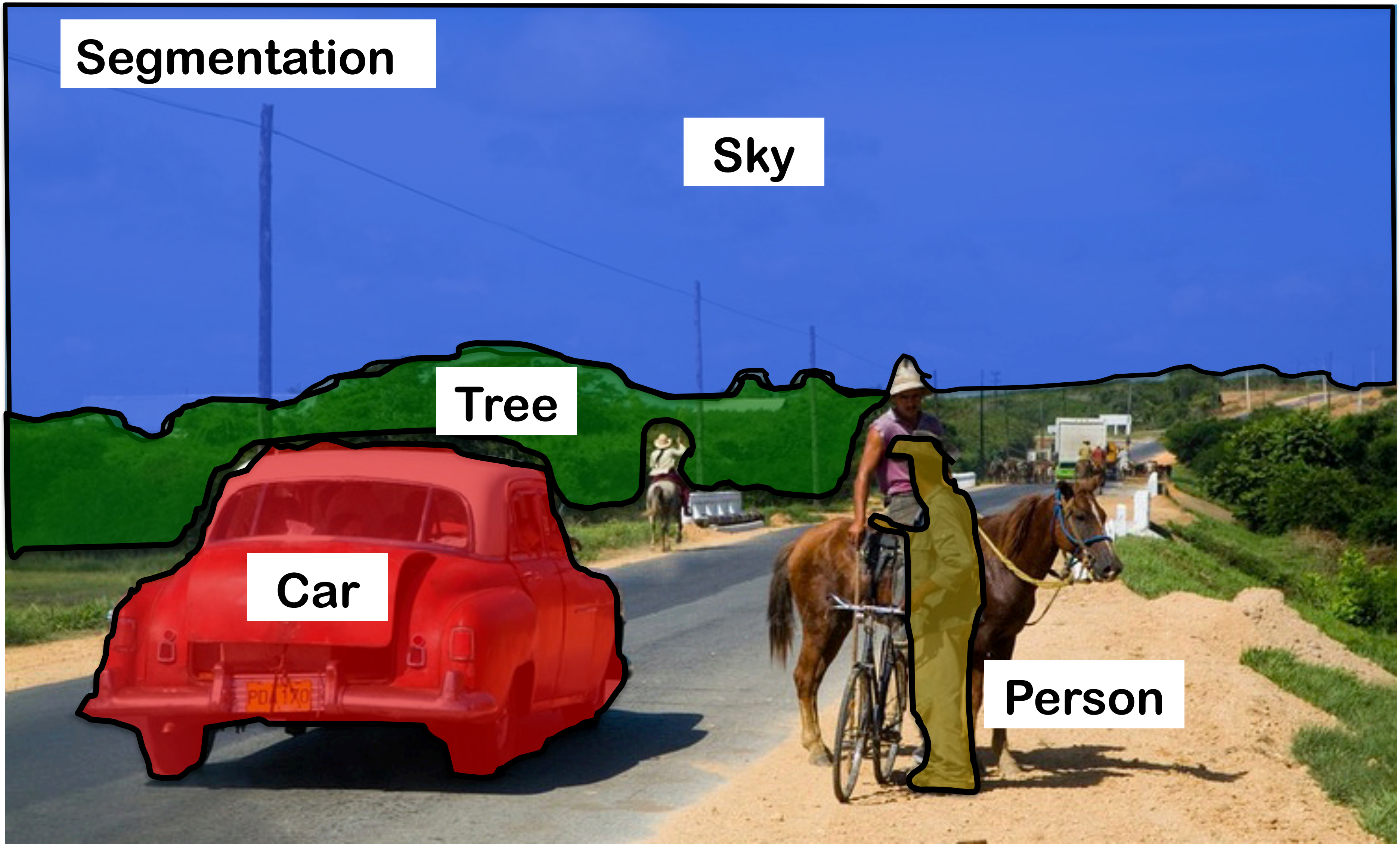
Visual recognition



Visual recognition



Visual recognition



Imaging process

- Light reaches surfaces in 3D.
- Surfaces reflect.
- Sensor element receives light energy.
- Intensity is important.
- Angles are important.
- Material is important.

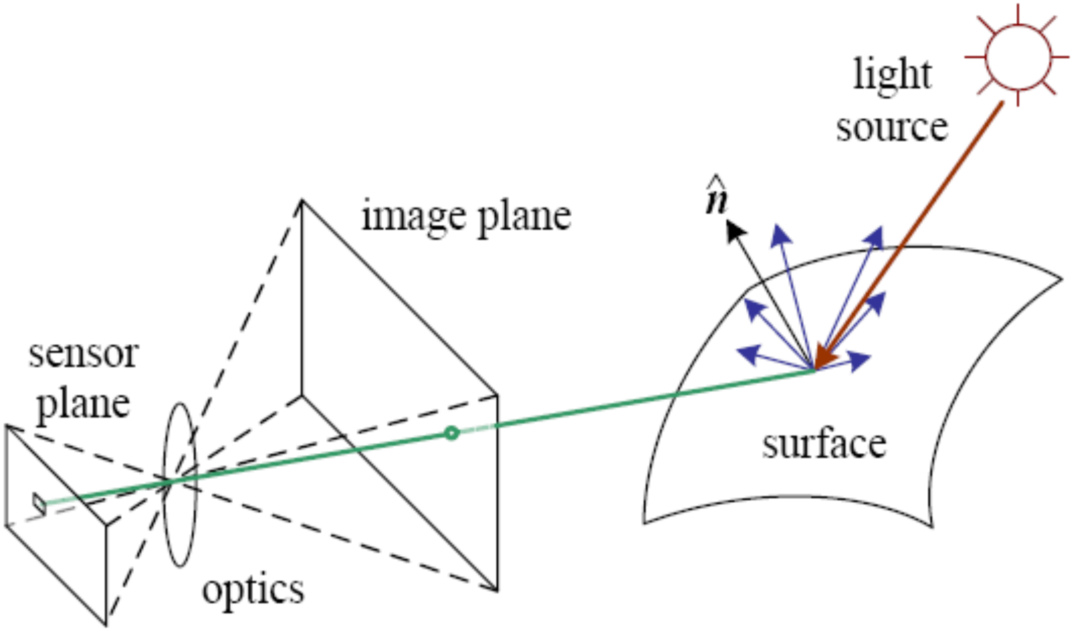


Figure 2.14: *A simplified model of photometric image formation. Light is emitted by one or more light sources, and is then reflected from an object's surface. A portion of this light is directed towards the camera. This simplified model ignores multiple reflections, which often occur in real-world scenes.*

Adapted from Rick Szeliski

Physical parameters

- Geometric
  - Type of projection
  - Camera pose
- Optical
  - Sensor’s lens type
  - Focal length, field of view, aperture
- Photometric
  - Type, direction, intensity of light reaching sensor
  - Surfaces’ reflectance properties
- Sensor
  - Sampling, etc.

Adapted from Trevor Darrell, UC Berkeley

Image acquisition

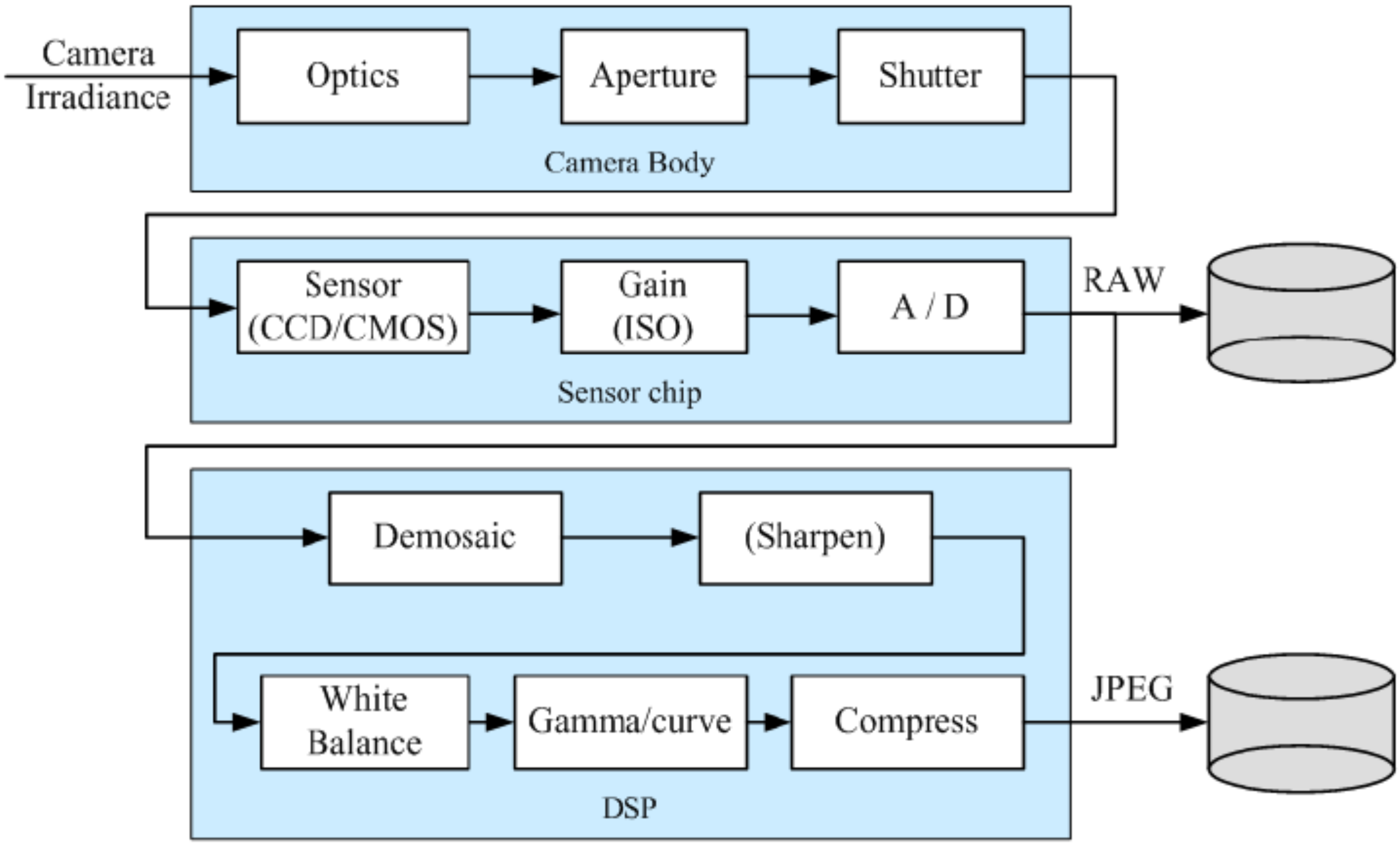


Figure 2.26: *Image sensing pipeline, showing the various sources of noise as well as the typical digital post-processing steps.*

Adapted from Rick Szeliski

Image & Camera (Chicken & Egg)

- What is an image?
- What is an image for a camera?
- Input?
- Output?
- Relation?
- Projection?
- Compression,
- etc.?

Projection?


MAKE POVERTY HISTORY
JULIAN BEEVER
CoolOpticalIllusions.com

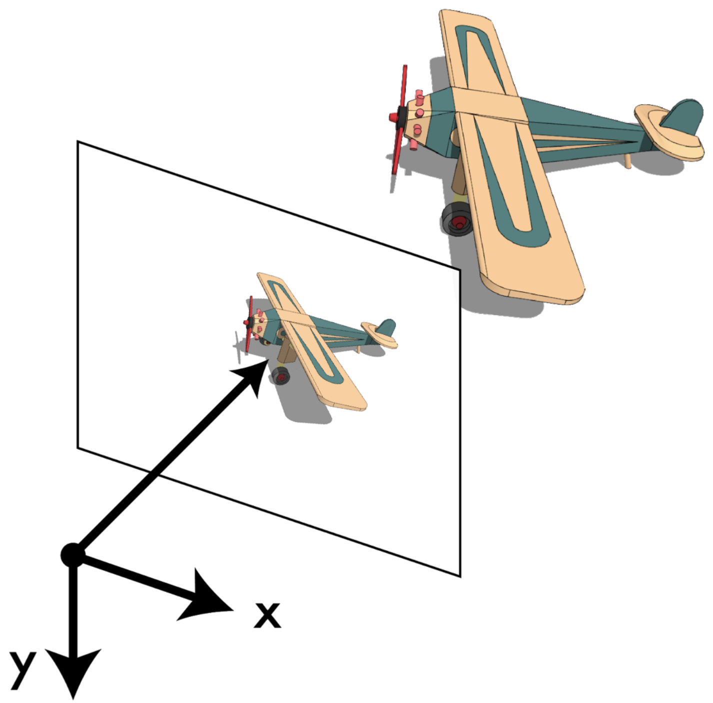

CoolOpticalIllusions.com

Projection ?

Chalk-art

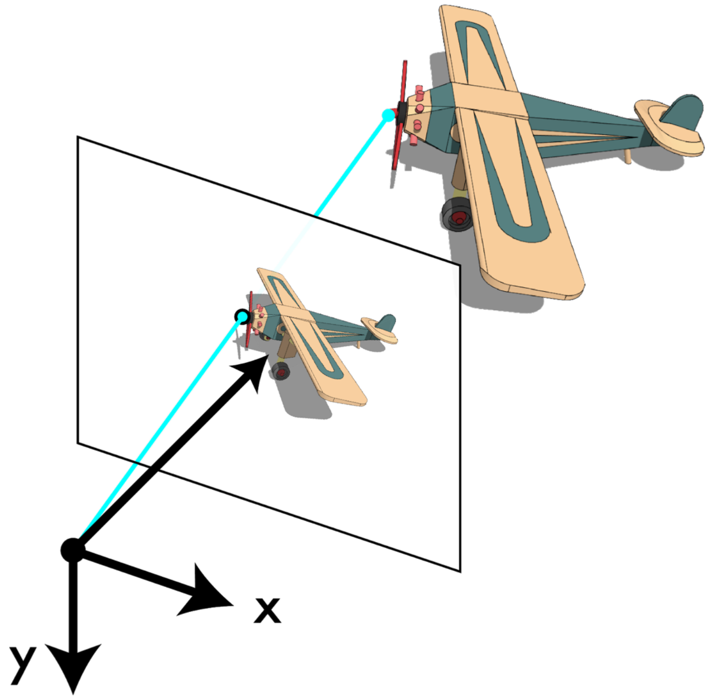

coolopticalillusions.com

Image: 3d->2d projection of the world



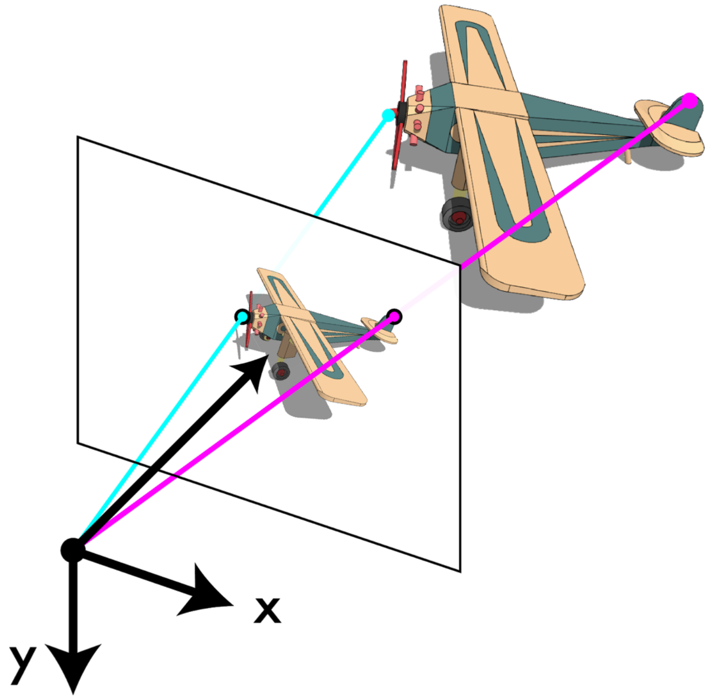
Adapted from Joseph Redmon, U of Washington

Image: 3d->2d projection of the world



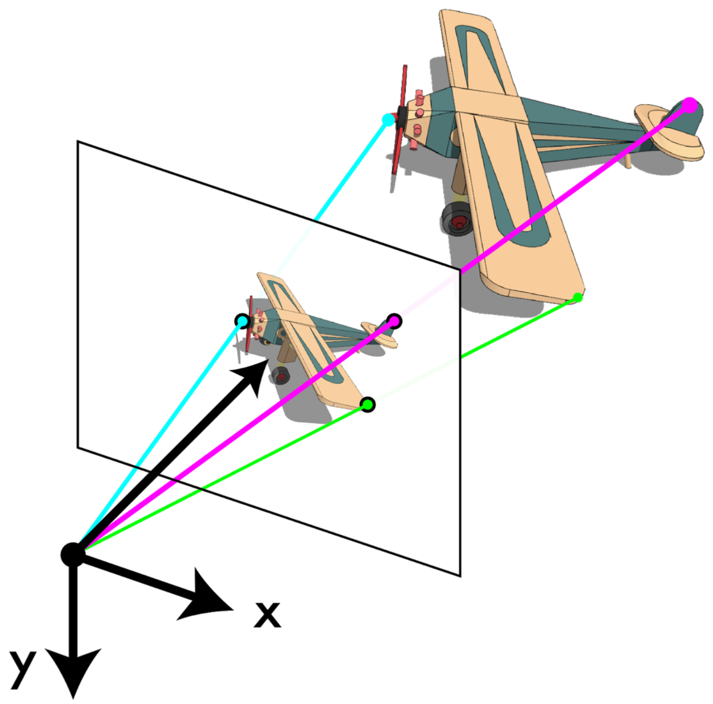
Adapted from Joseph Redmon, U of Washington

Image: 3d->2d projection of the world



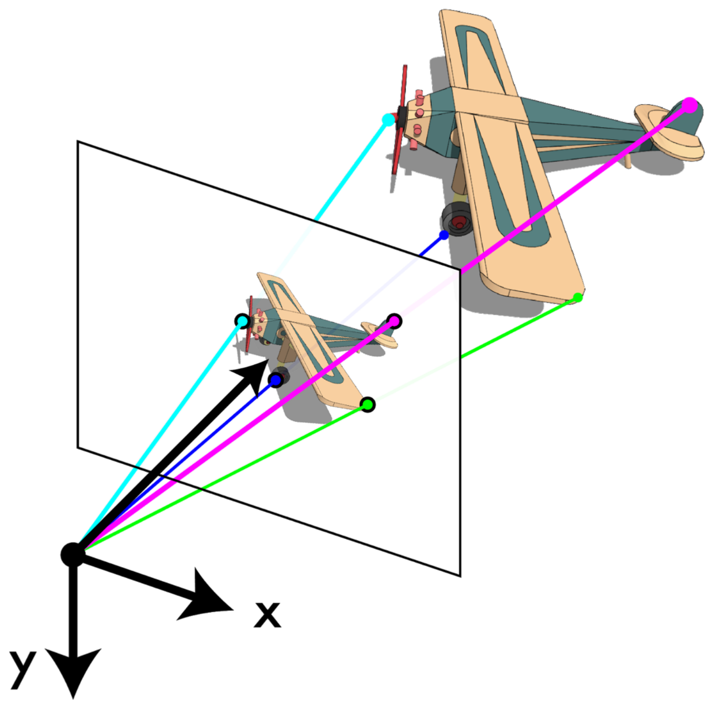
Adapted from Joseph Redmon, U of Washington

Image: 3d->2d projection of the world



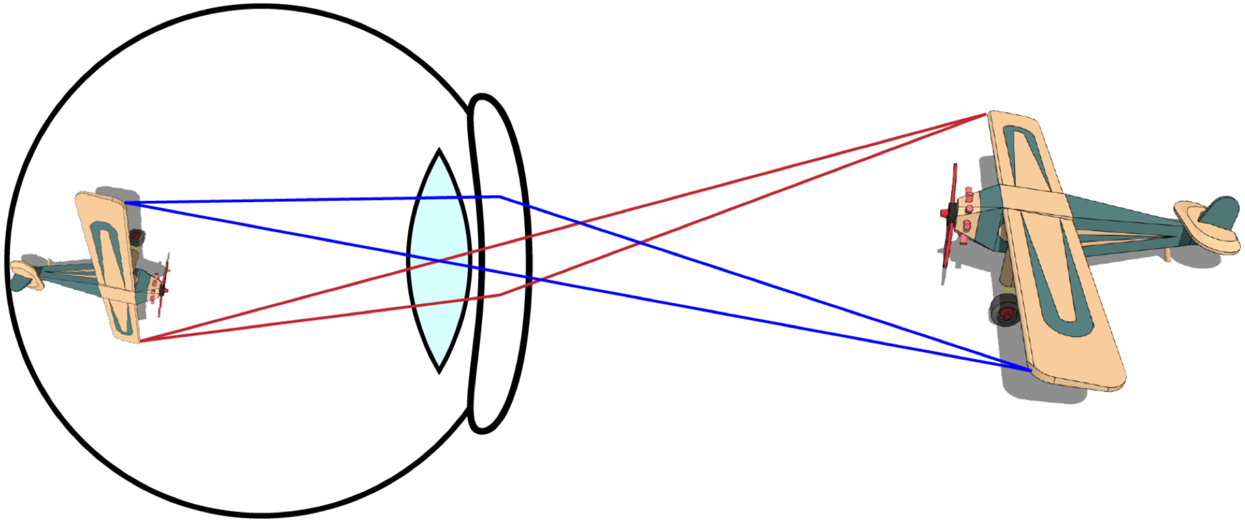
Adapted from Joseph Redmon, U of Washington

Image: 3d->2d projection of the world



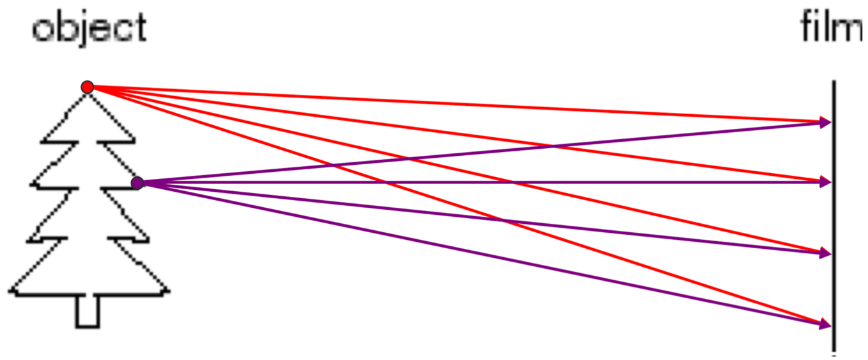
Adapted from Joseph Redmon, U of Washington

Eyes: projection onto retina



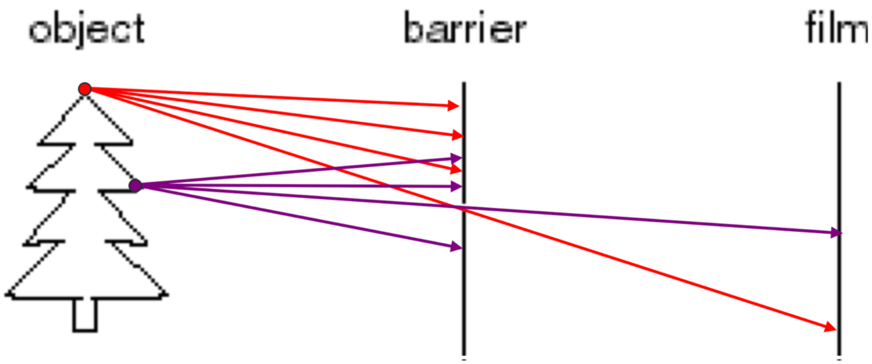
Adapted from Joseph Redmon, U of Washington

How to model a camera?



Source:https://www.cc.gatech.edu/~afb/classes/CS4495-Fall2013/slides/CS4495-05-CameraModel.pdf

How to model a camera?



Source:https://www.cc.gatech.edu/~afb/classes/CS4495-Fall2013/slides/CS4495-05-CameraModel.pdf

Pinhole Camera (Dark Chamber Model)

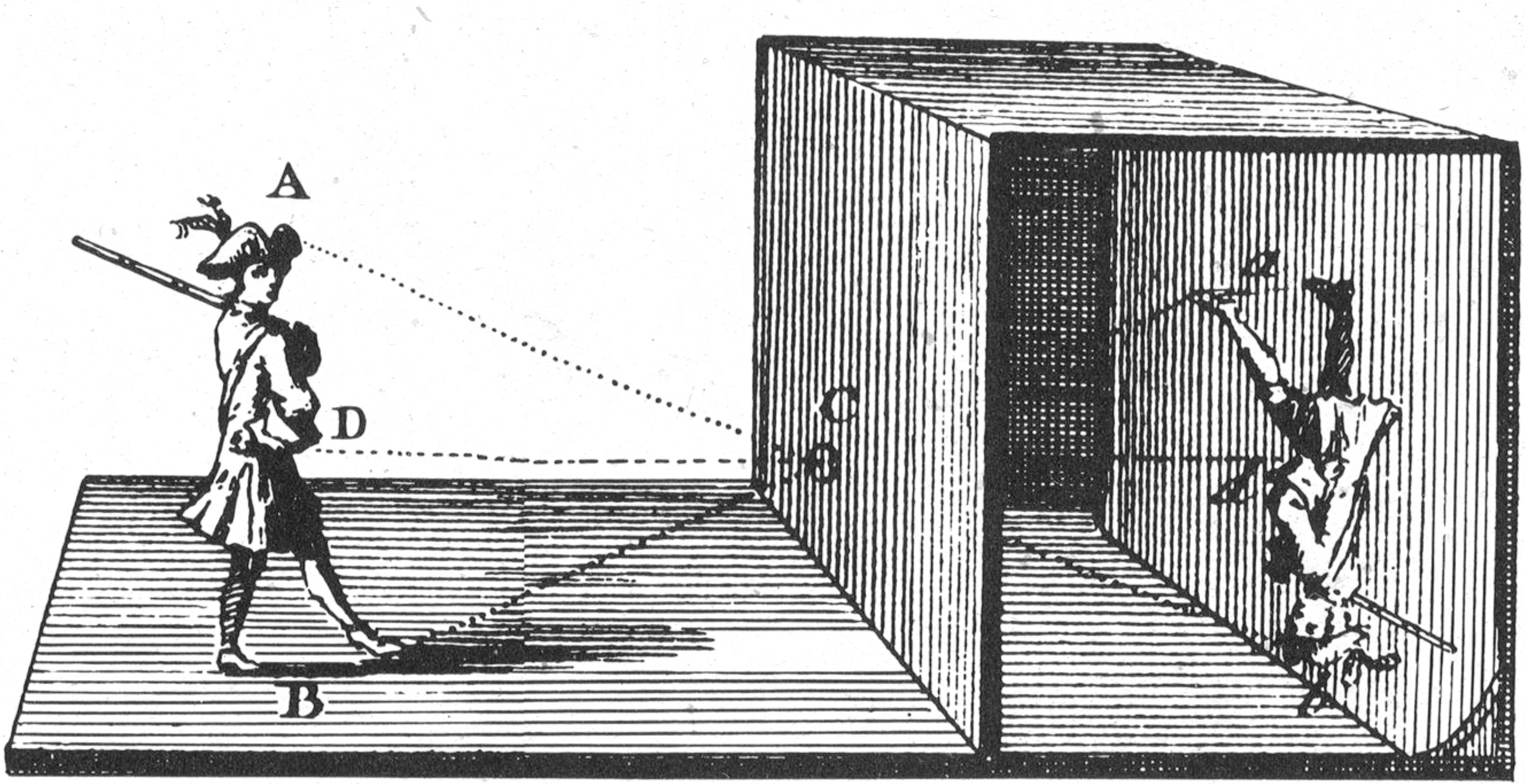
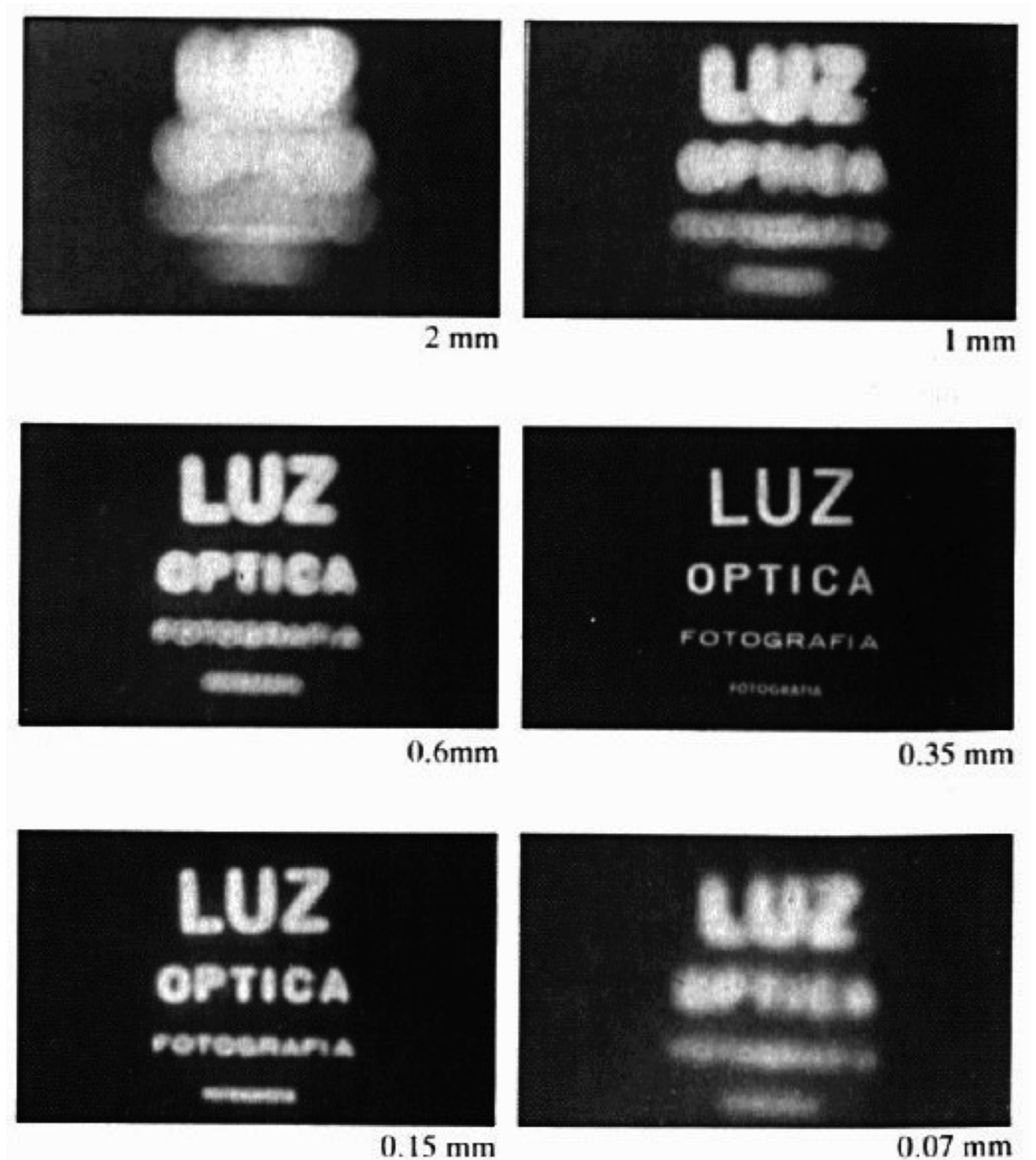


Image source: https://commons.wikimedia.org/wiki/File%3A001_a01_camera_obscura_abrazolas.jpg
https://en.wikipedia.org/wiki/Camera_obscura

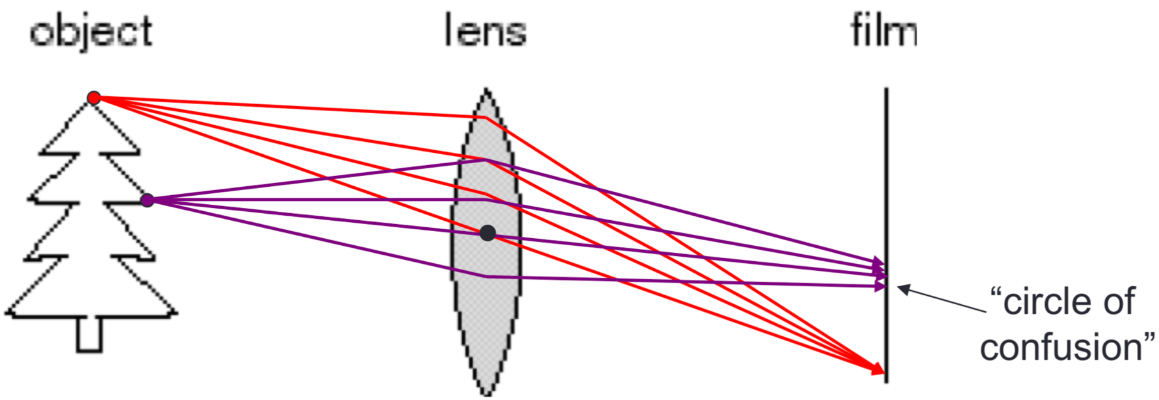
Aperture size

Sources:
https://web.stanford.edu/class/cs231a/course_notes/01-camera-models.pdf

https://www.cc.gatech.edu/~afb/classes/CS4495-Fall2013/slides/CS4495-05-CameraModel.pdf



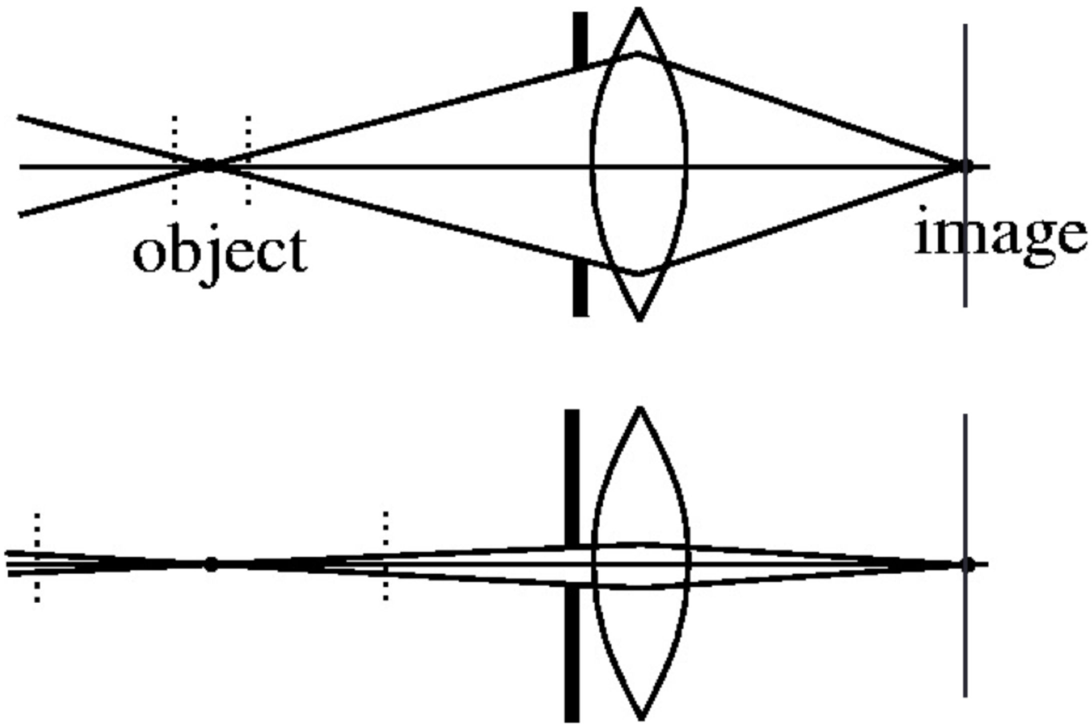
Camera with a lens



- A lens focuses more light onto the film
- There is a specific distance at which objects are "in focus" and other points project to a "circle of confusion" in the image
- Changing the shape (or thickness, radius or parameters) of the lens changes this distance

Source:https://www.cc.gatech.edu/~afb/classes/CS4495-Fall2013/slides/CS4495-05-CameraModel.pdf

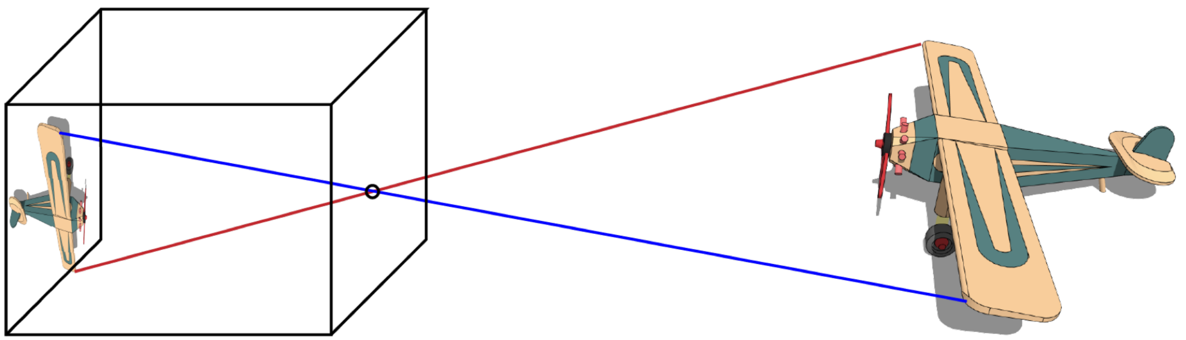
Depth of field



$f / 5.6$

$f / 32$

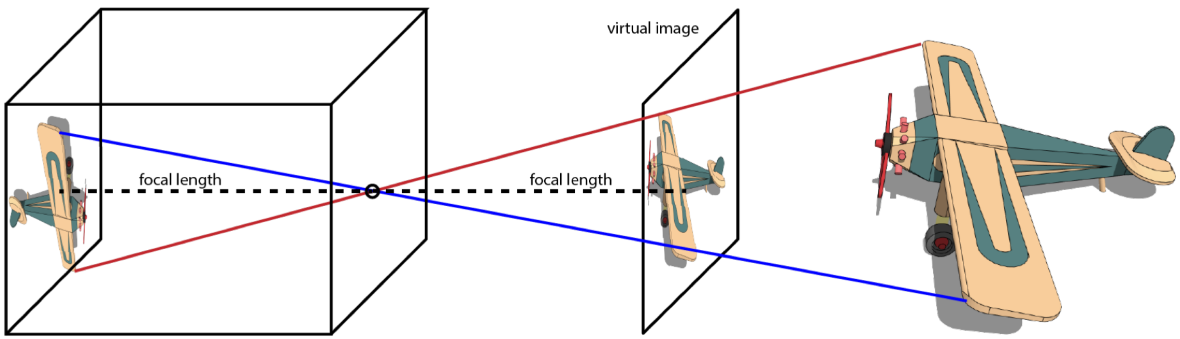
Model: pinhole camera



No lens
No distortion model to correct any possible distortion

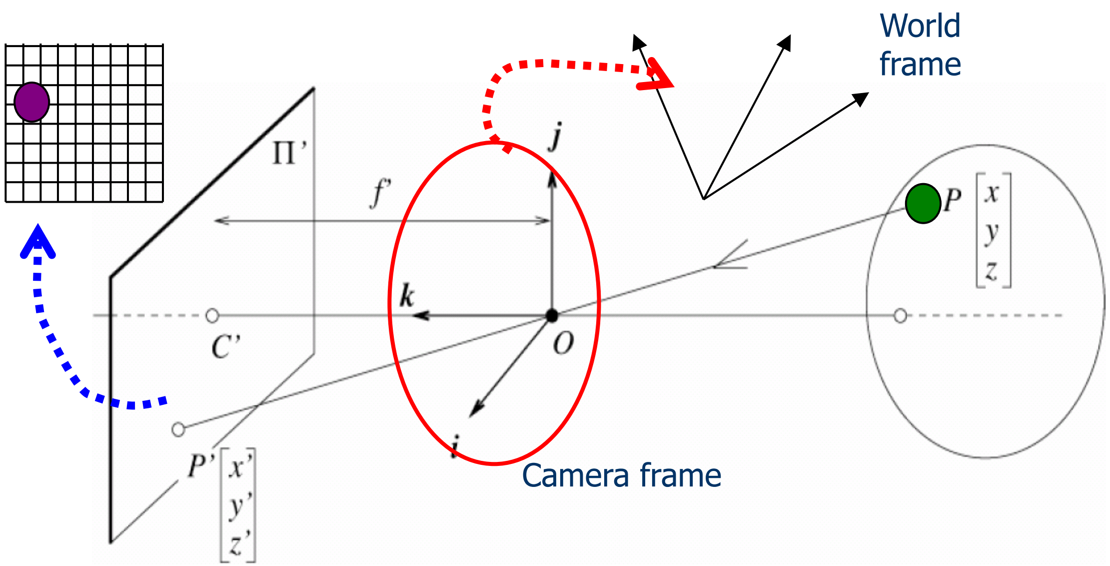
Adapted from Joseph Redmon, U of Washington

Model: pinhole camera



Adapted from Joseph Redmon, U of Washington

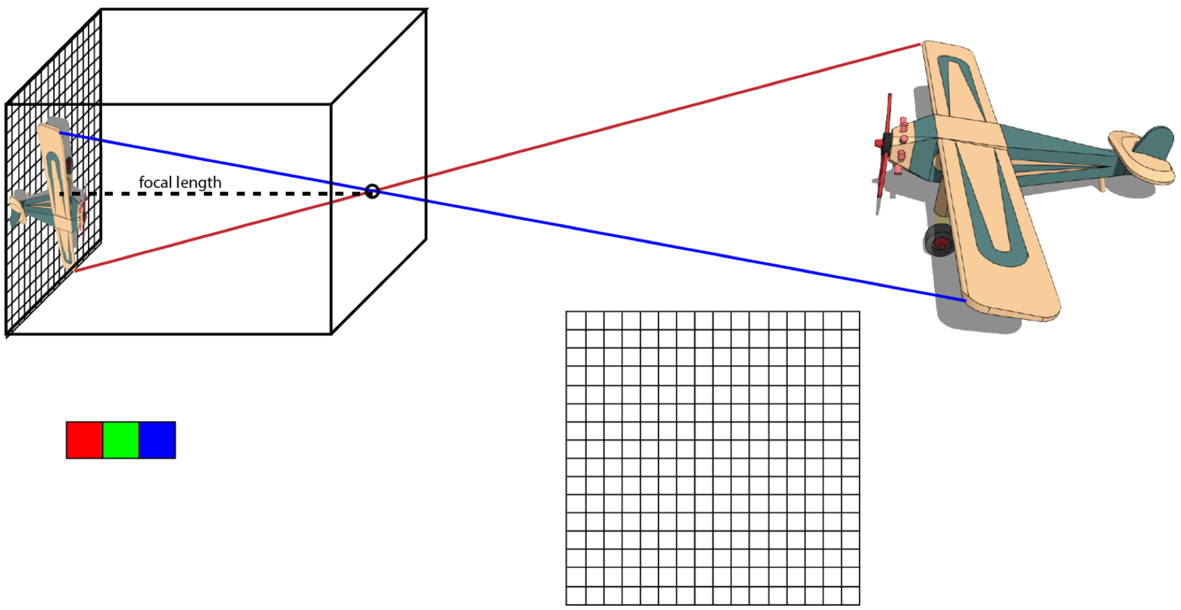
Camera calibration



- Camera's extrinsic and intrinsic parameters are needed to calibrate the geometry.
- Extrinsic: camera frame $\leftrightarrow$ world frame
- Intrinsic: image coordinates relative to camera $\leftrightarrow$ pixel coordinates

Adapted from Trevor Darrell, UC Berkeley

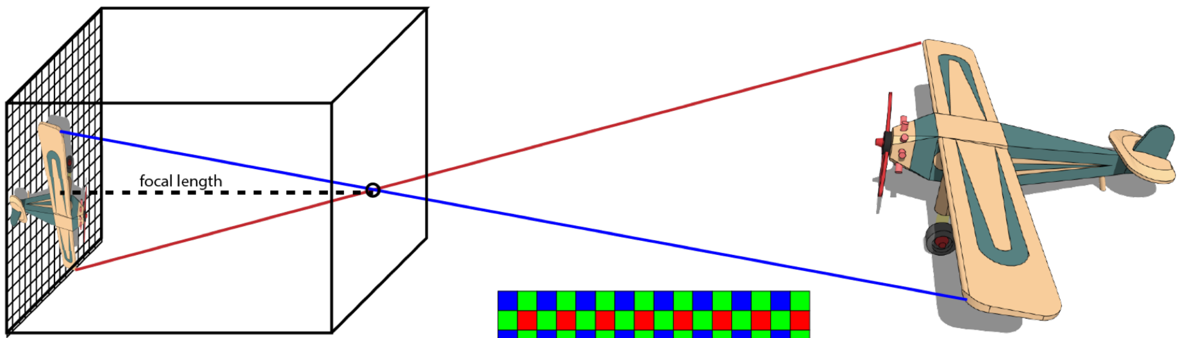
How do we record color?

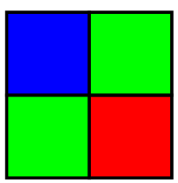


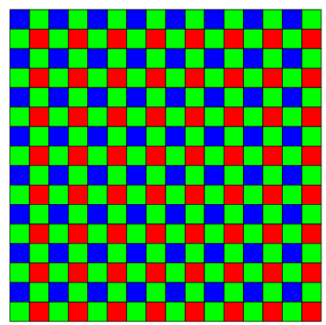
At each point we record incident light

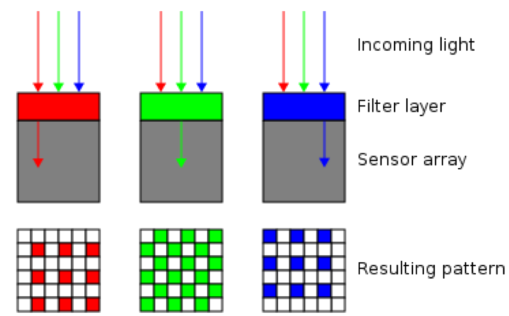
Adapted from Joseph Redmon, U of Washington

How do we record color?







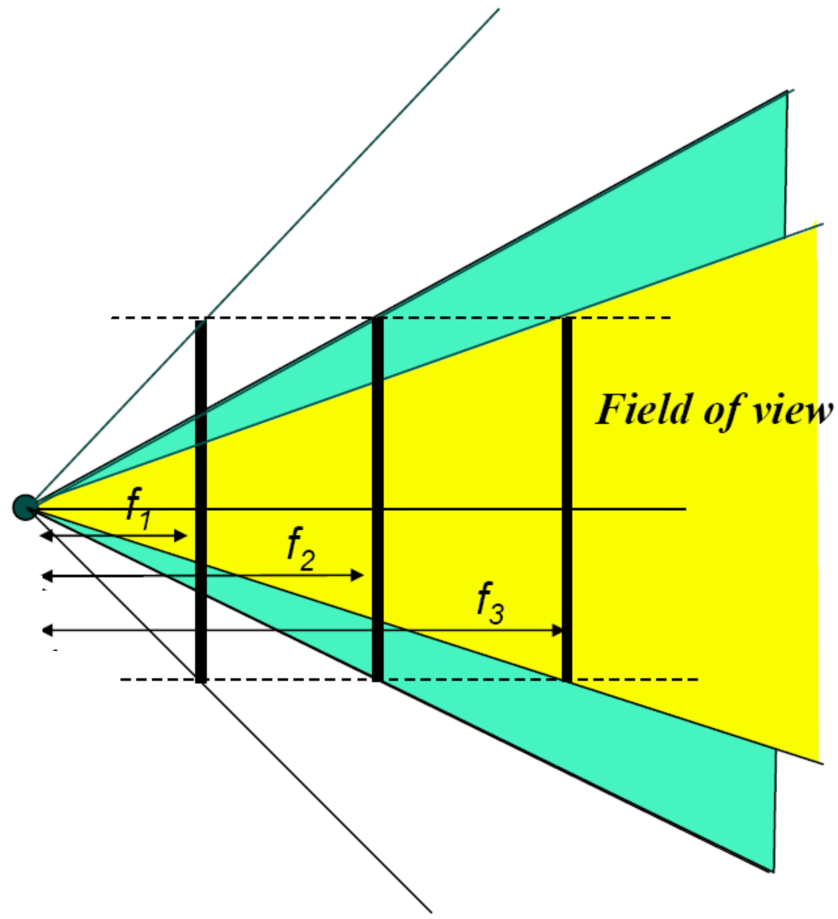


Bayer pattern for CMOS sensors

Adapted from Joseph Redmon, U of Washington

Focal length

- Field of view depends on focal length.
- As **f** gets smaller, image becomes more *wide* angle
  - more world points project onto the finite image plane
- As **f** gets larger, image becomes more *telescopic*
  - smaller part of the world projects onto the finite image plane



Adapted from Trevor Darrell, UC Berkeley

Focal length

28 mm

35 mm

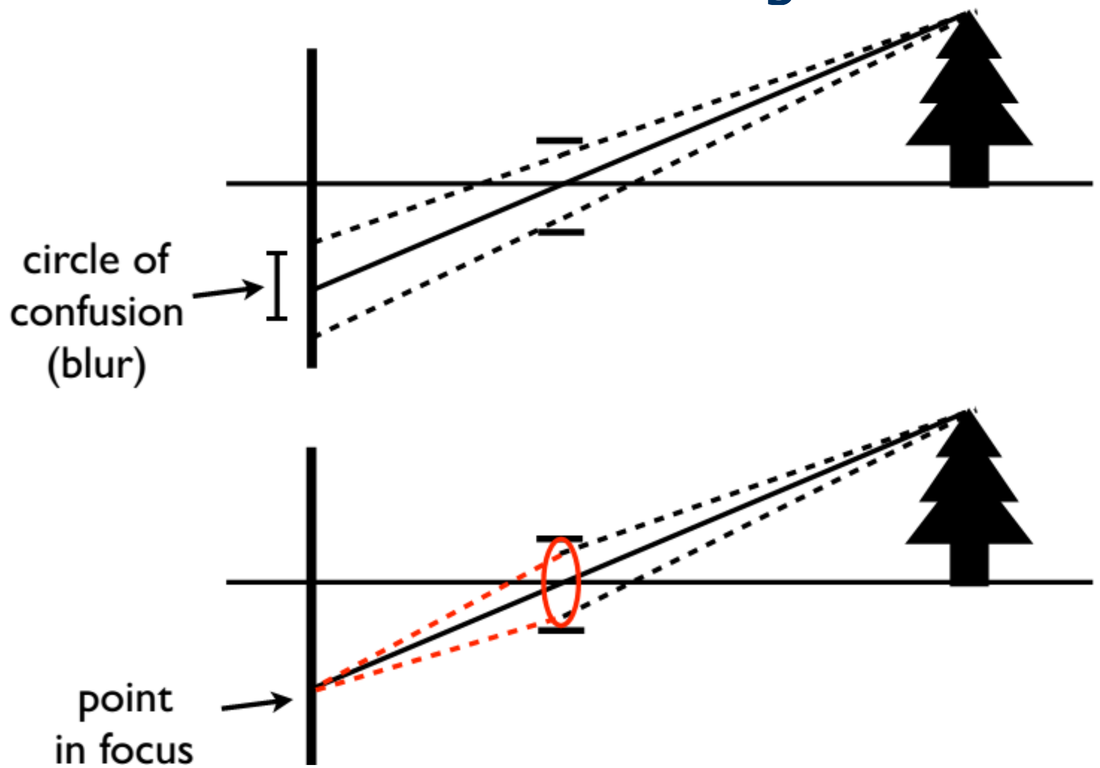
50 mm

70 mm

Adapted from Matthew Brown, U of Washington

Aperture

- A real camera must have a finite aperture to get enough light, but this causes blur in the image

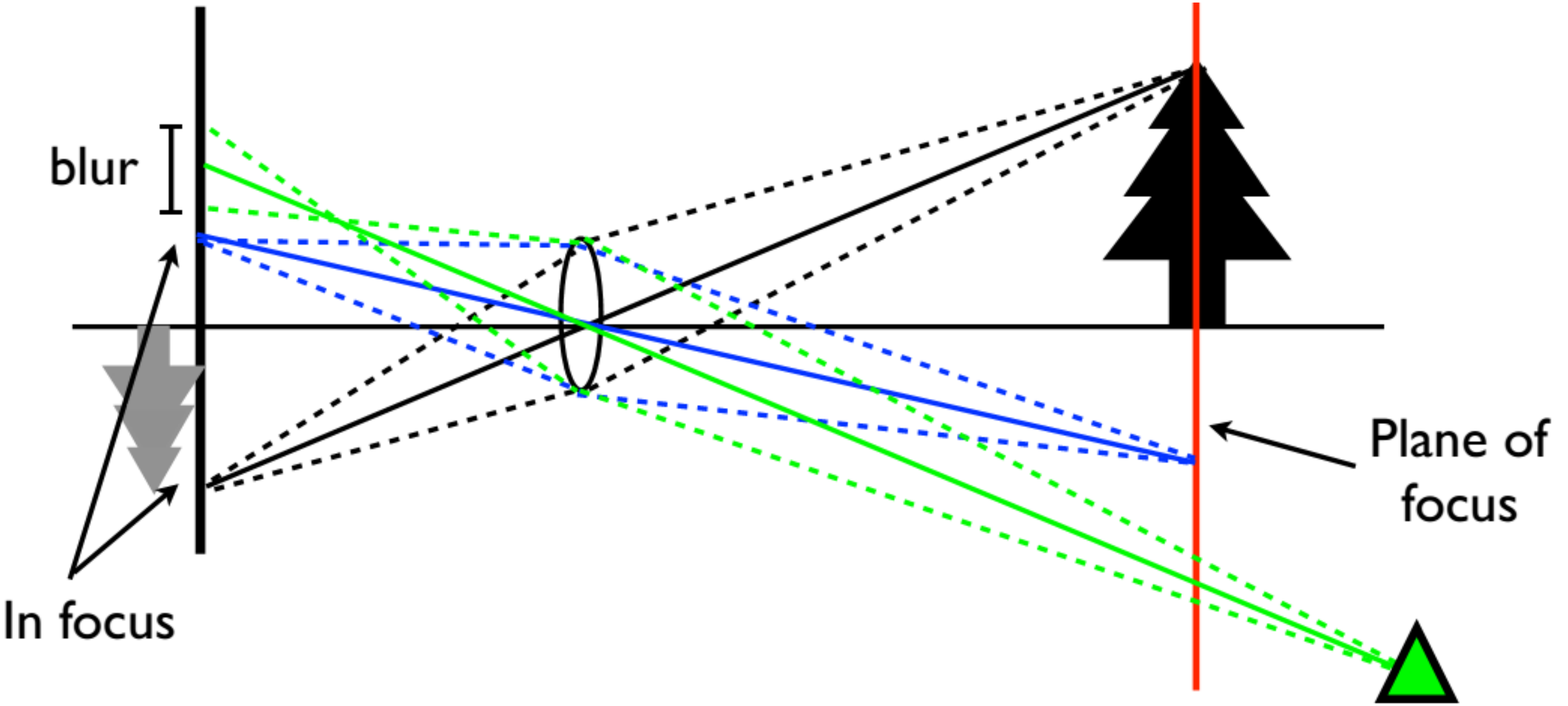


- Solution: use a lens to focus light onto the image plane

Adapted from Matthew Brown, U of Washington

Aperture

- Note that lenses focus all rays from a plane in the world
- Objects off the plane are blurred depending on distance



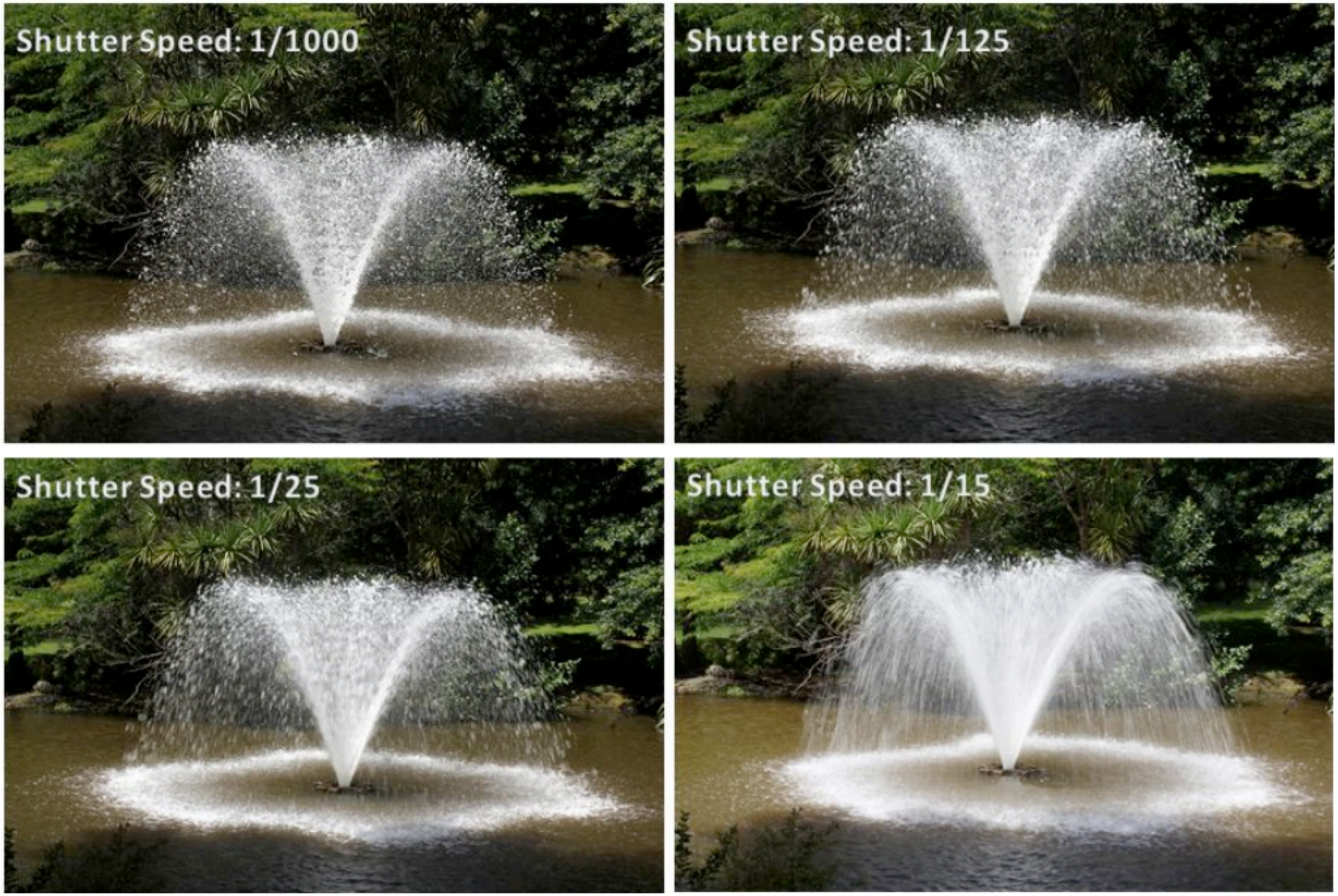
Adapted from Matthew Brown, U of Washington

Aperture

- Smaller aperture -> smaller blur, larger depth of field

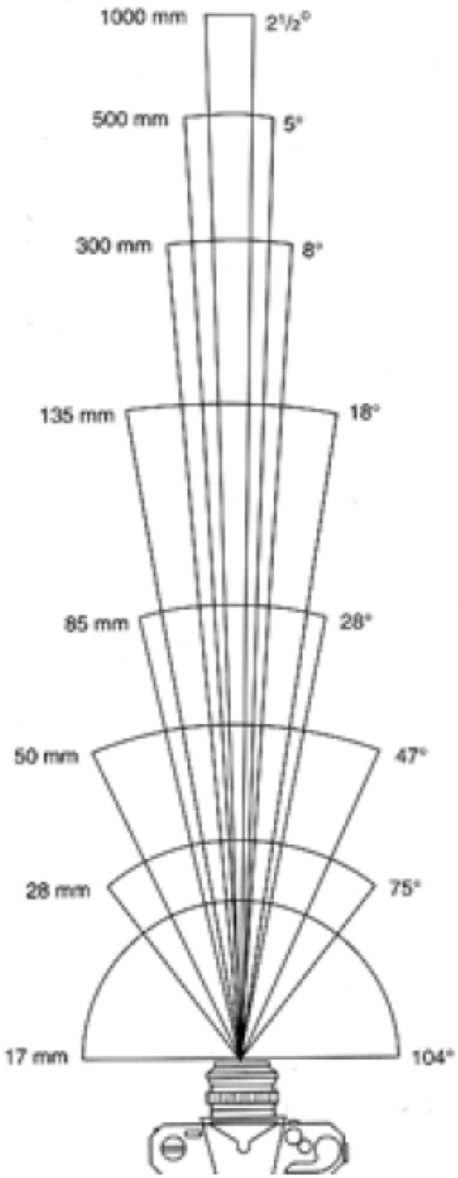
Adapted from Matthew Brown, U of Washington

Shutter speed



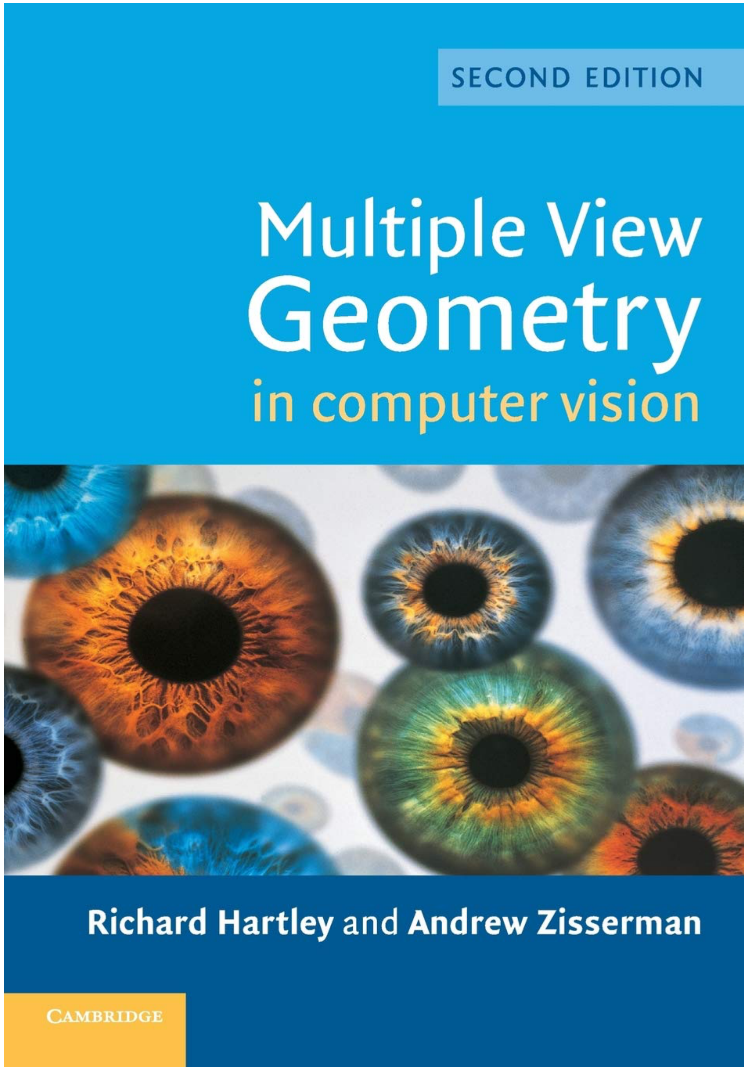
Adapted from Matthew Brown, U of Washington

Field of View (zoom)



From London and Upton

Theory of Cameras



There is a huge math about cameras and matching them between them and to the actual world.

There is a (text) book focusing on this topic alone.

(and there are courses focusing on this topic alone)

Further details on the geometry of multiple cameras are available in this text book.

Sampling and quantization

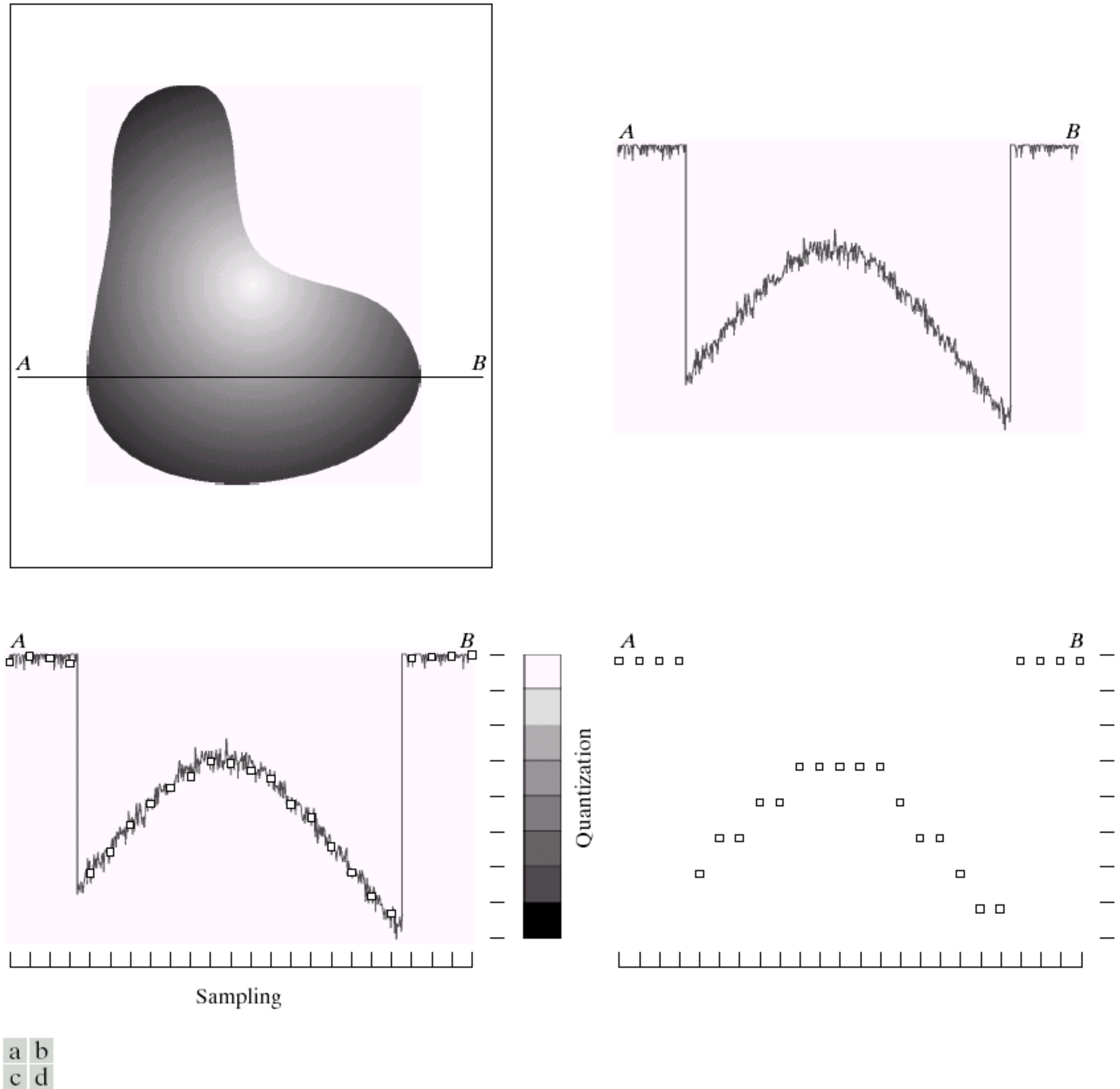
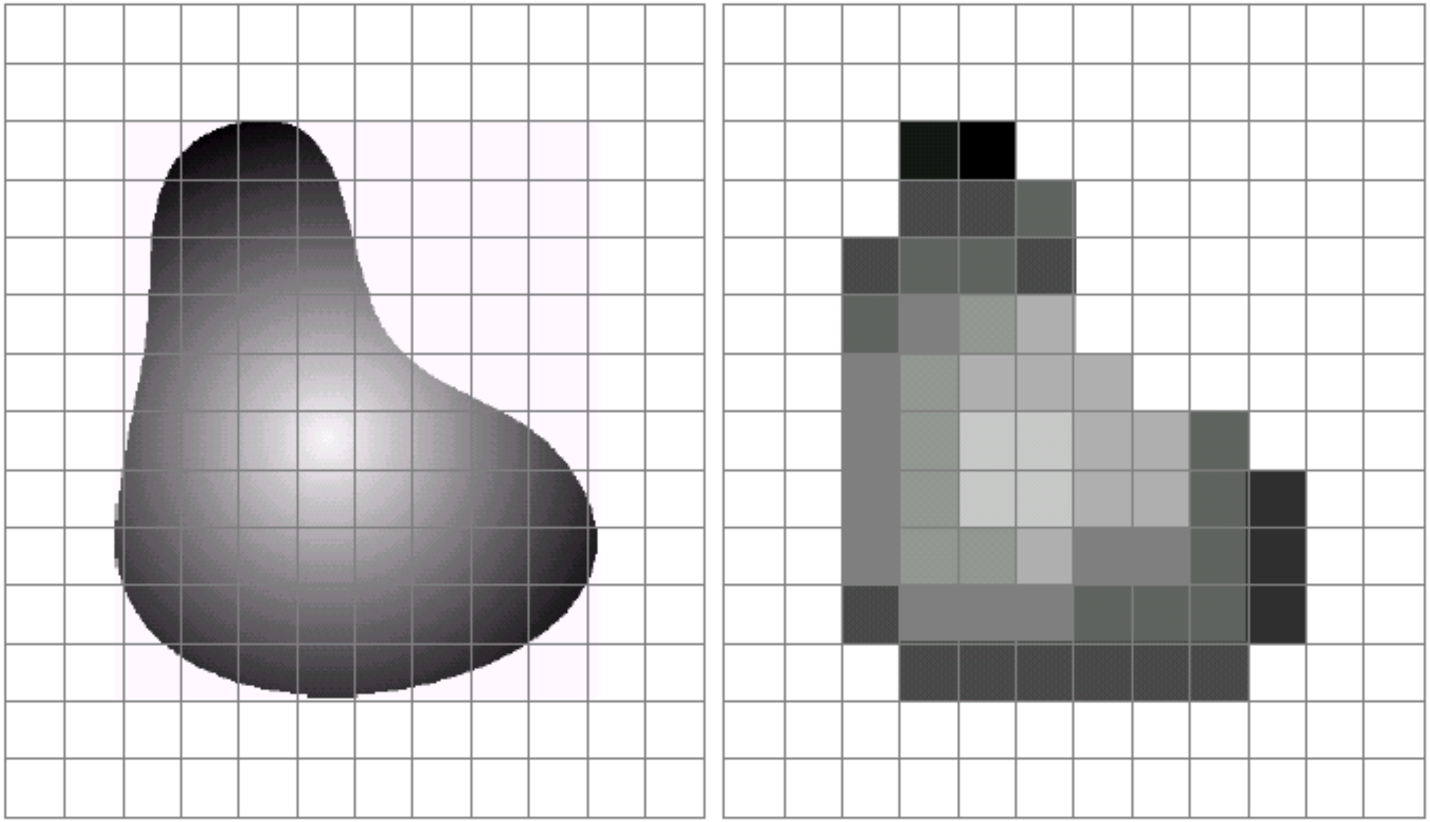


FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from A to B in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

Sampling and quantization

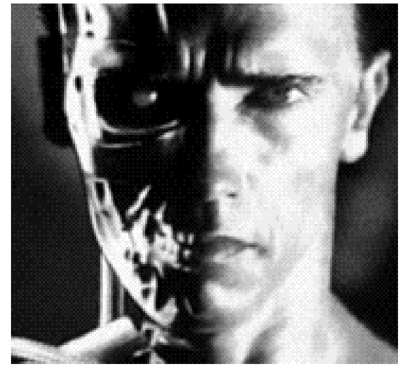


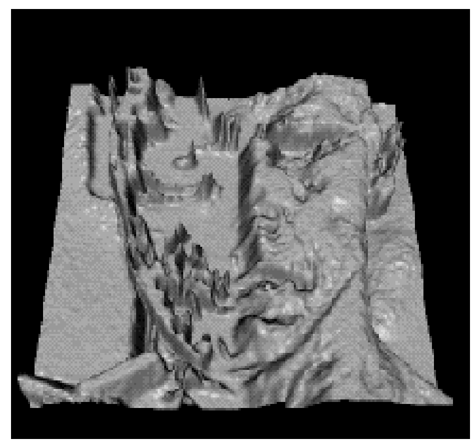
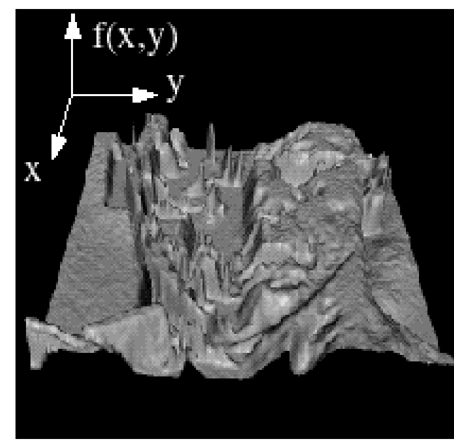
a b

FIGURE 2.17 (a) Continuos image projected onto a sensor array. (b) Result of image sampling and quantization.

Image representation

- Images can be represented by 2D functions of the form f(x,y).
- The physical meaning of the value of f at spatial coordinates (x,y) is determined by the source of the image.

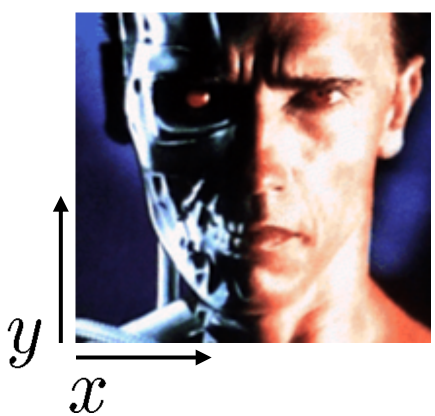


Adapted from Shapiro and Stockman

Image representation

- In a digital image, both the coordinates and the image value become discrete quantities.
- Images can now be represented as 2D arrays (matrices) of integer values: I[i,j] (or I[r,c]).
- The term gray level is used to describe monochromatic intensity.



j

i

62	79	23	119	120	105	4	0
10	10	9	62	12	78	34	0
10	58	197	46	46	0	0	48
176	135	5	188	191	68	0	49
2	1	1	29	26	37	0	77
0	89	144	147	187	102	62	208
255	252	0	166	123	62	0	31
166	63	127	17	1	0	99	30

Spatial resolution

- Spatial resolution is the smallest discernible detail in an image.
- Sampling is the principal factor determining spatial resolution.

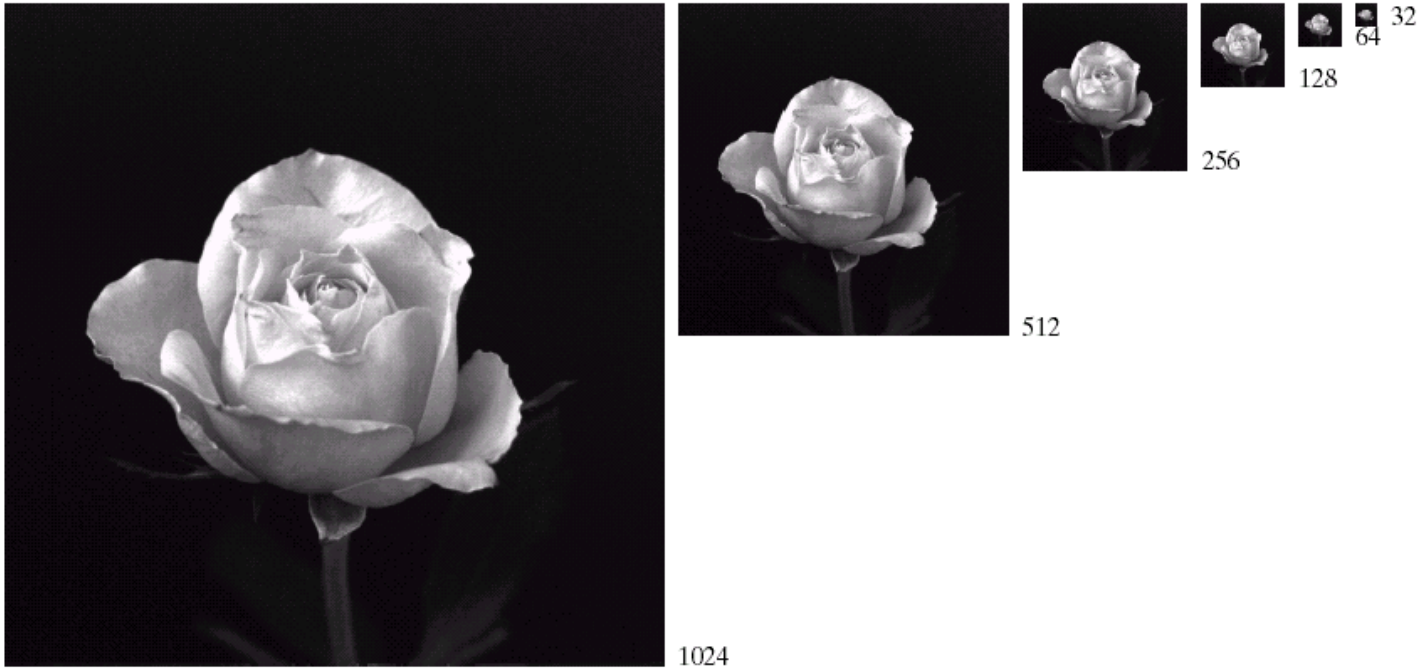
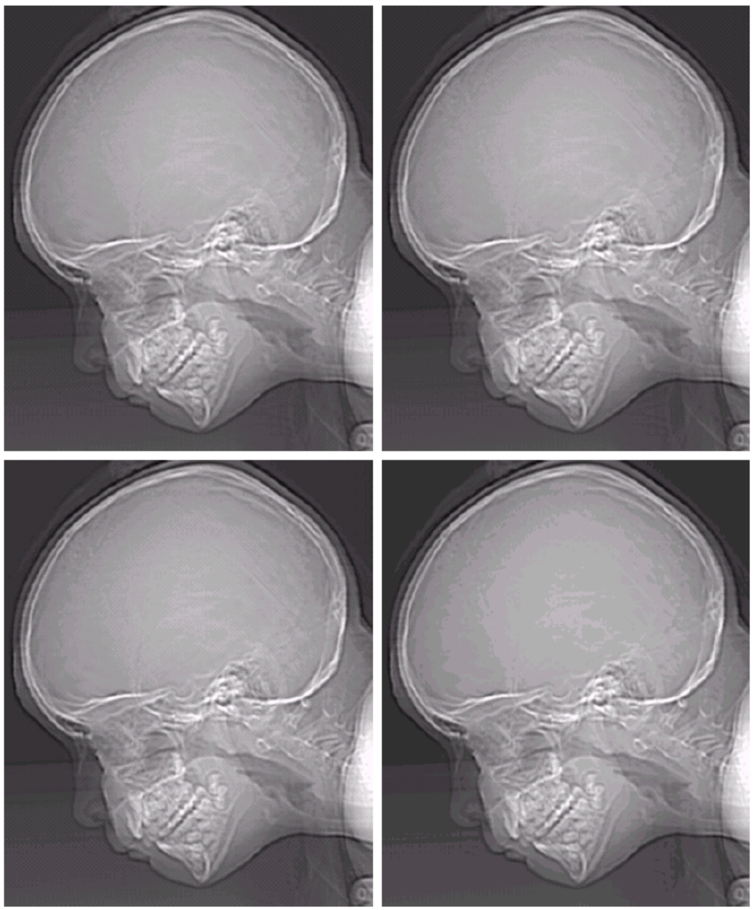


FIGURE 2.19 A 1024 × 1024, 8-bit image subsampled down to size 32 × 32 pixels. The number of allowable gray levels was kept at 256.

Gray level resolution

- Gray level resolution refers to the smallest discernible change in gray level (often power of 2).

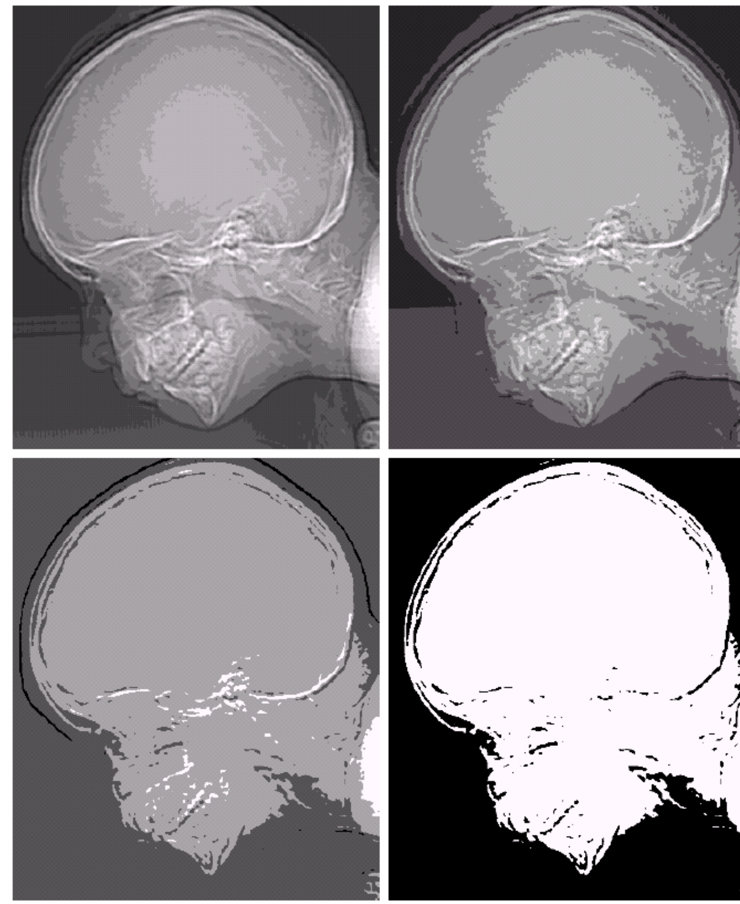


a b
c d

FIGURE 2.21 (a) 452 × 374, 256-level image. (b)–(d) Image displayed in 128, 64, and 32 gray levels, while keeping the spatial resolution constant.

e f
g h

FIGURE 2.21 *(Continued)* (e)–(h) Image displayed in 16, 8, 4, and 2 gray levels. (Original courtesy of Dr. David R. Pickens, Department of Radiology & Radiological Sciences, Vanderbilt University Medical Center.)



Bit planes

a	b	c
d	e	f
g	h	i

FIGURE 3.14 (a) An 8-bit gray-scale image of size 500 × 1192 pixels. (b) through (i) Bit planes 1 through 8, with bit plane 1 corresponding to the least significant bit. Each bit plane is a binary image.

Electromagnetic (EM) spectrum

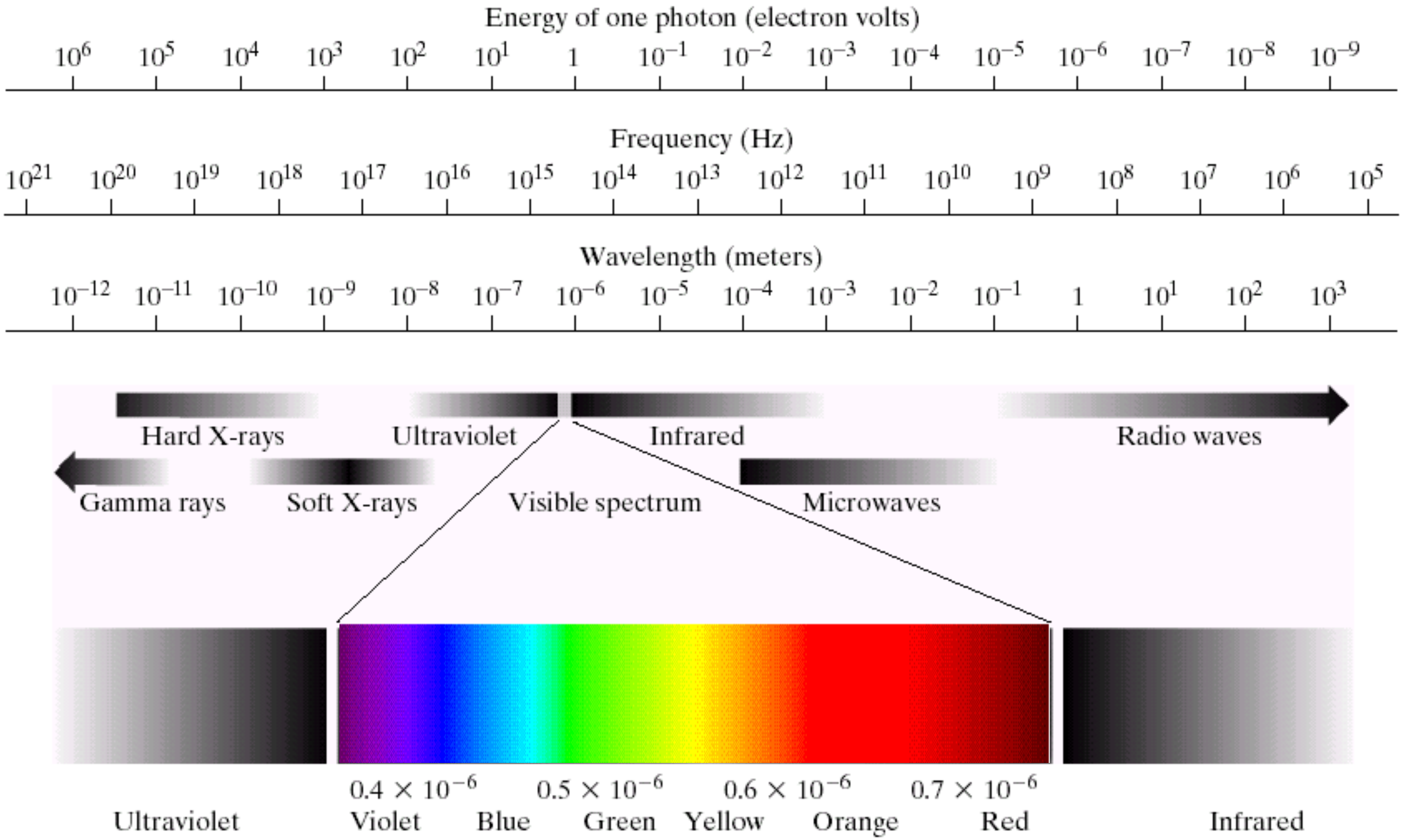
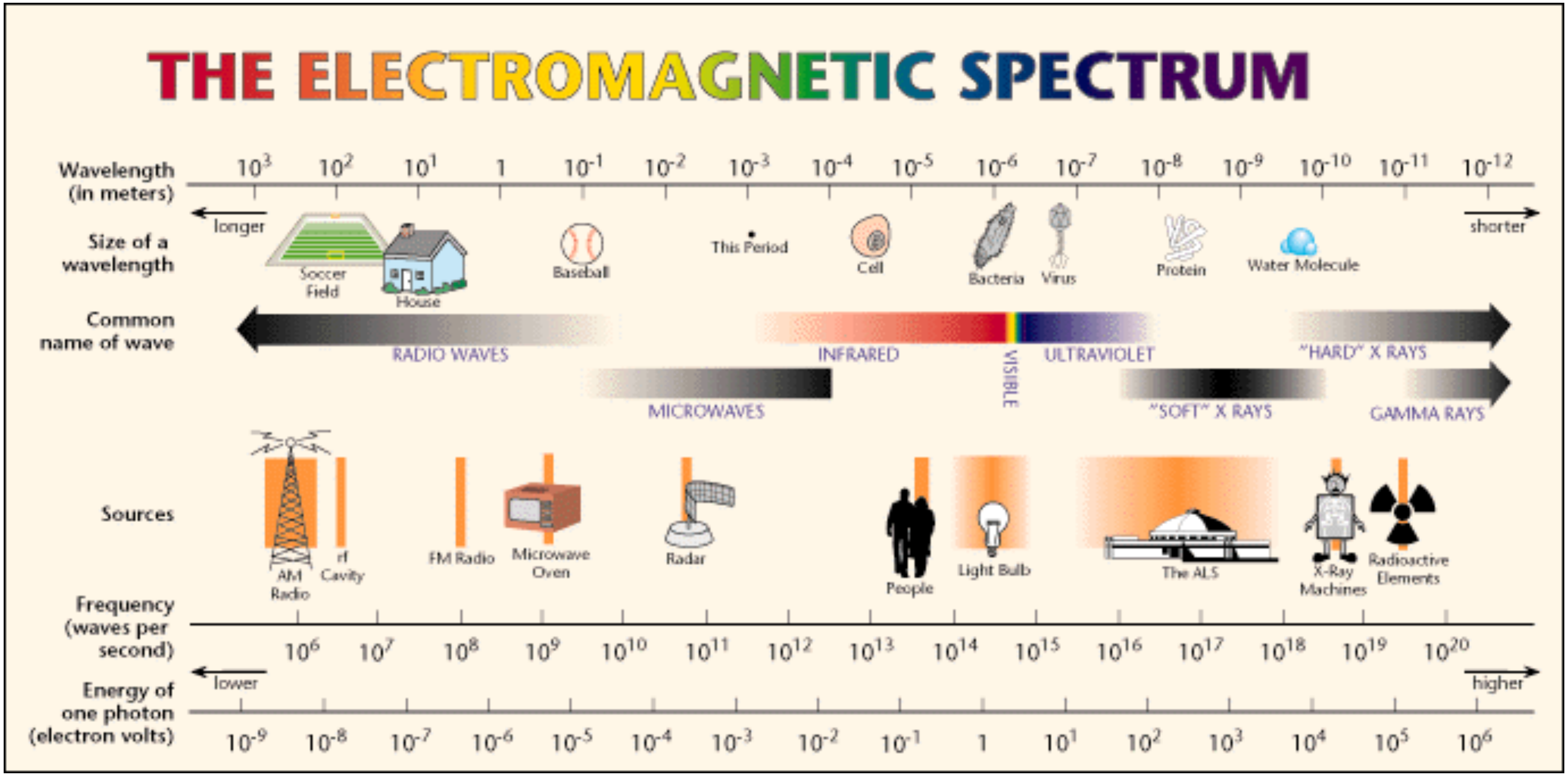


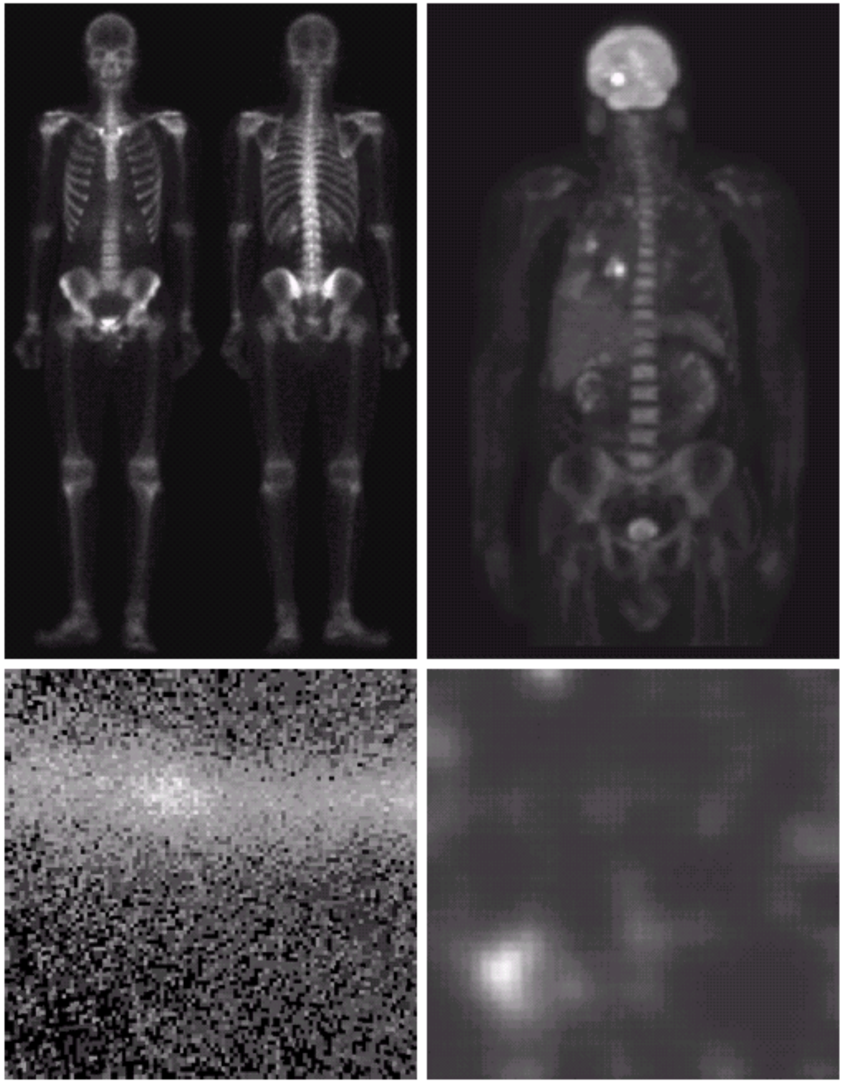
FIGURE 2.10 The electromagnetic spectrum. The visible spectrum is shown zoomed to facilitate explanation, but note that the visible spectrum is a rather narrow portion of the EM spectrum.

Electromagnetic (EM) spectrum

- The wavelength of an EM wave required to “see” an object must be of the same size as or smaller than the object.

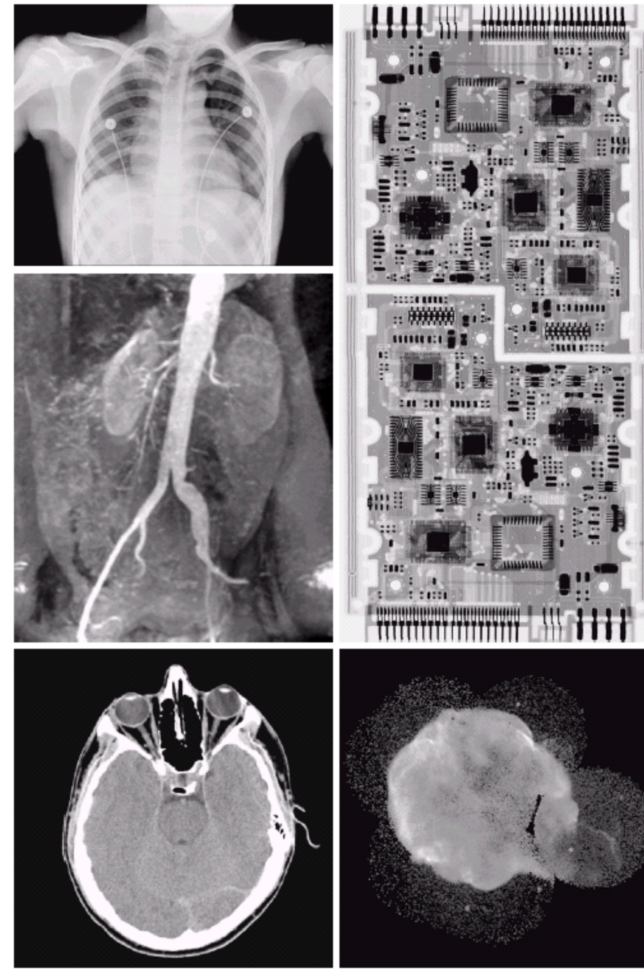


Other types of sensors



a b
c d

FIGURE 1.6 Examples of gamma-ray imaging. (a) Bone scan. (b) PET image. (c) Cygnus Loop. (d) Gamma radiation (bright spot) from a reactor valve. (Images courtesy of (a) G.E. Medical Systems, (b) Dr. Michael E. Casey, CTI PET Systems, (c) NASA, (d) Professors Zhong He and David K. Wehe, University of Michigan.)



a d
b
c e

FIGURE 1.7 Examples of X-ray imaging. (a) Chest X-ray. (b) Aortic angiogram. (c) Head CT. (d) Circuit boards. (e) Cygnus Loop. (Images courtesy of (a) and (c) Dr. David R. Pickens, Dept. of Radiology & Radiological Sciences, Vanderbilt University Medical Center, (b) Dr. Thomas R. Gest, Division of Anatomical Sciences, University of Michigan Medical School, (d) Mr. Joseph E. Pascente, Lixi, Inc., and (e) NASA.)

Other types of sensors

a b
c

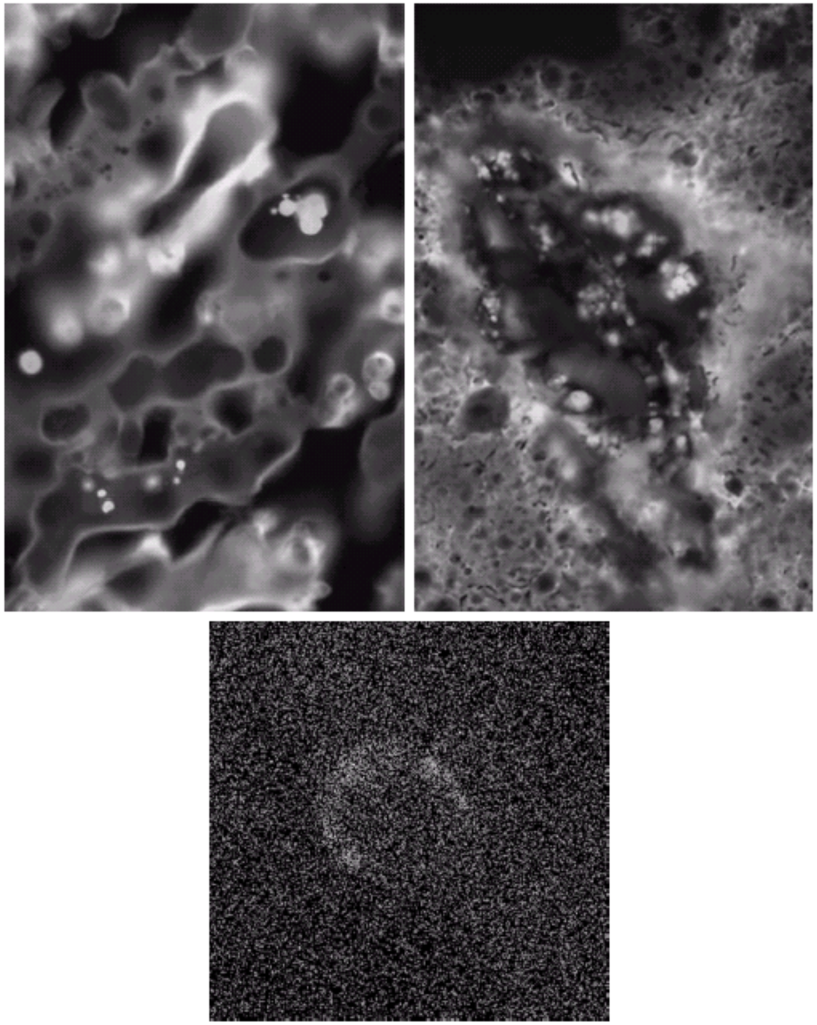
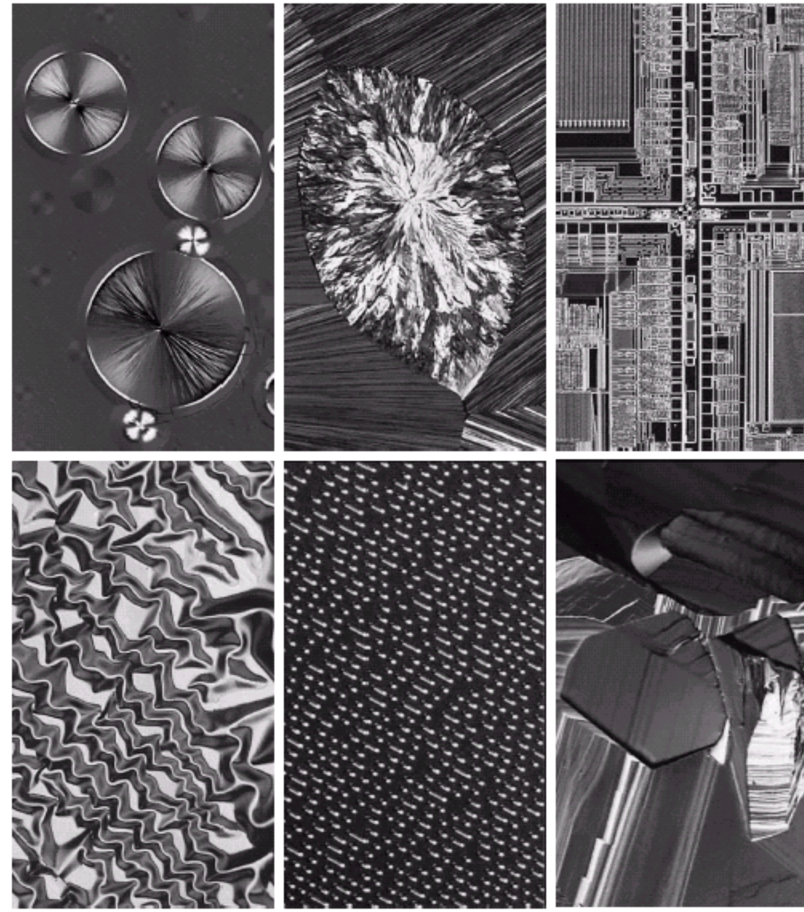


FIGURE 1.8 Examples of ultraviolet imaging. (a) Normal corn. (b) Smut corn. (c) Cygnus Loop. (Images courtesy of (a) and (b) Dr. Michael W. Davidson, Florida State University, (c) NASA.)



a b c
d e f

FIGURE 1.9 Examples of light microscopy images. (a) Taxol (anticancer agent), magnified 250×. (b) Cholesterol—40×. (c) Microprocessor—60×. (d) Nickel oxide thin film—600 ×. (e) Surface of audio CD—1750×. (f) Organic superconductor—450×. (Images courtesy of Dr. Michael W. Davidson, Florida State University.)

Other types of sensors

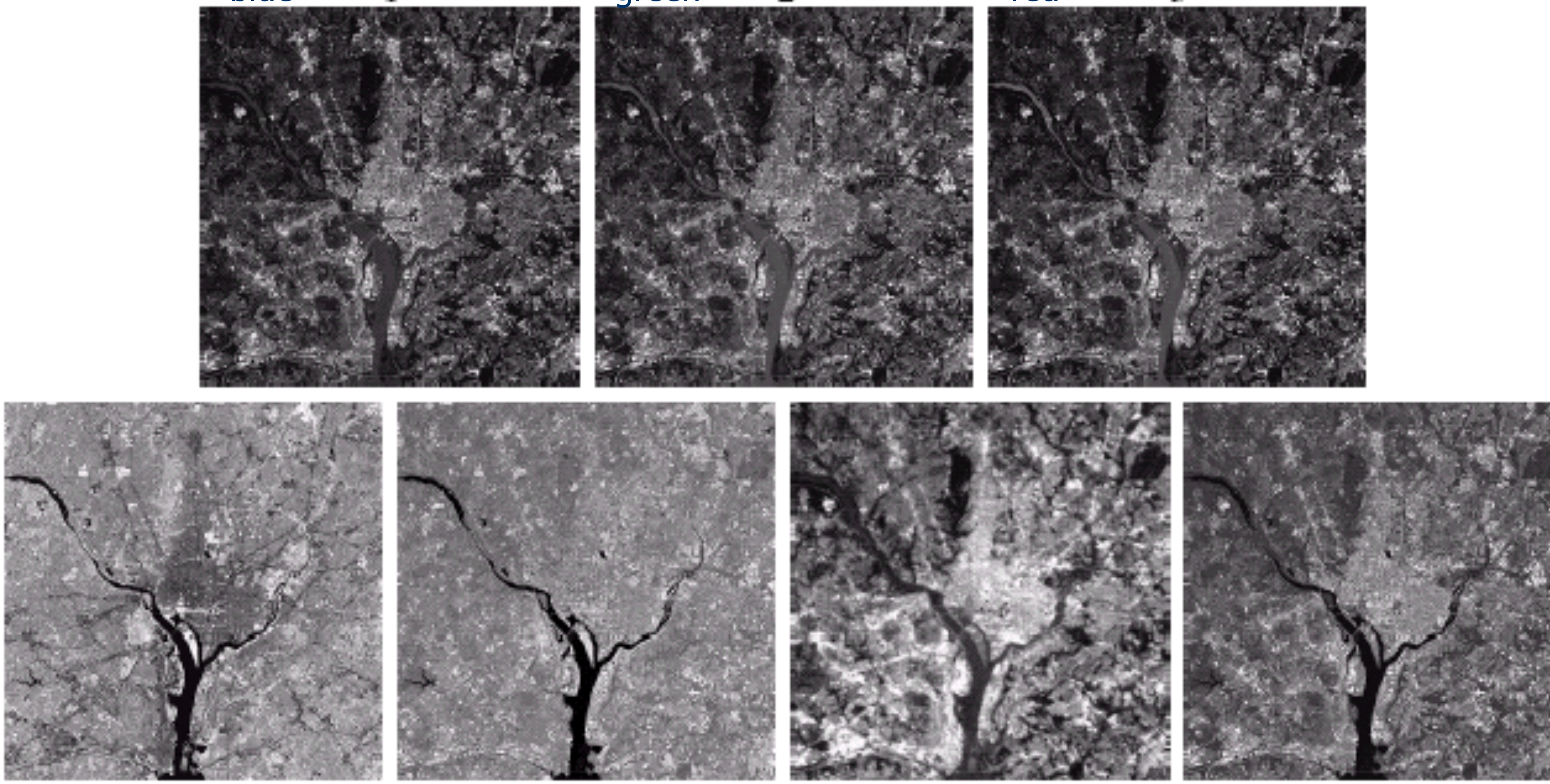
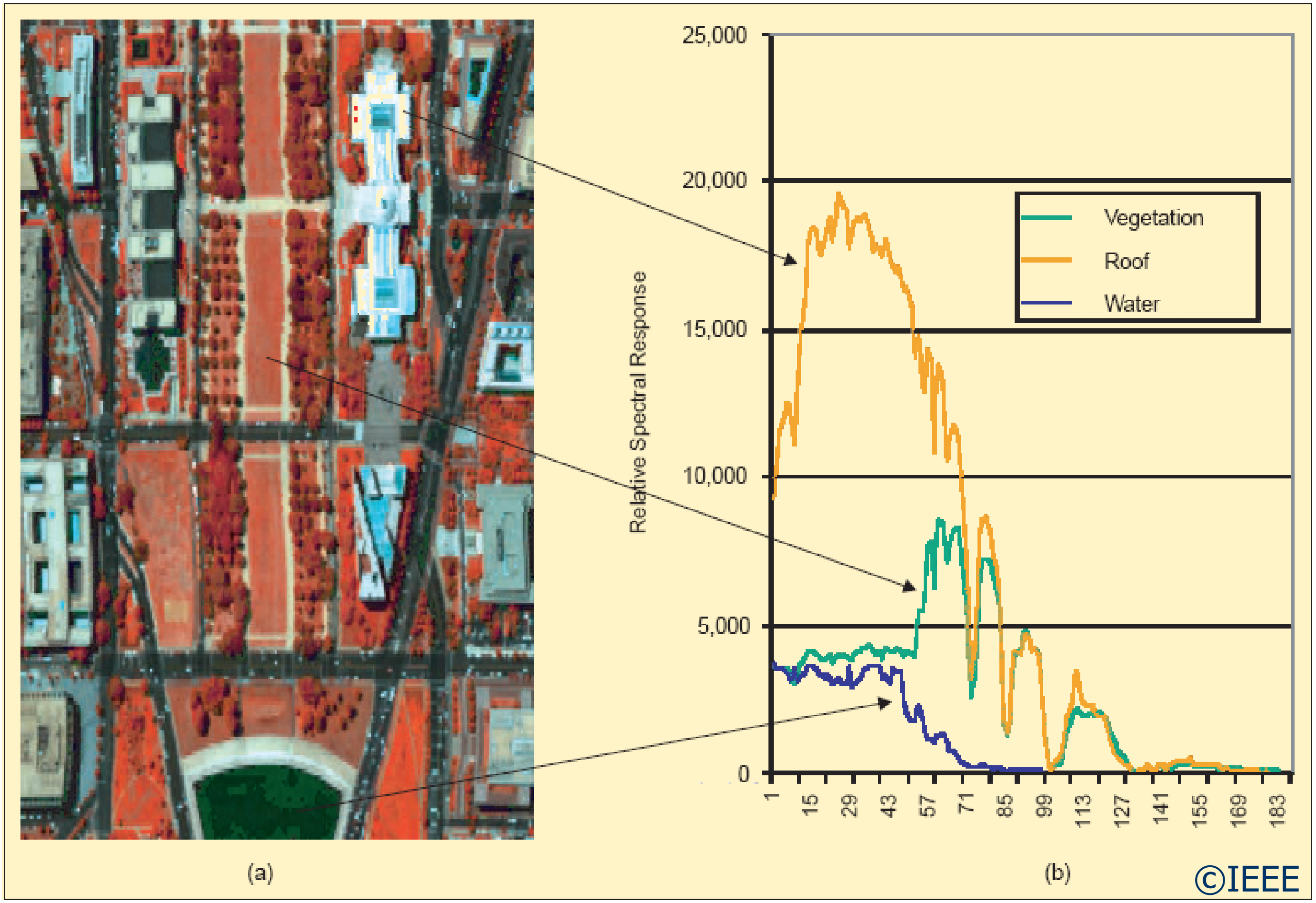


FIGURE 1.10 LANDSAT satellite images of the Washington, D.C. area. The numbers refer to the thematic bands in Table 1.1. (Images courtesy of NASA.)

Other types of sensors



1. (a) A simulated color IR image of an urban area, the Washington, D.C., mall. This image is made using three bands of the 210 bands collected by the sensor system, one band from the visible green, one from the visible red, and one from the near infrared. Such displays are referred to as displays in image space. (b) A display of the data of pixels of three materials as a function of wavelength by spectral band number. The bands in this case are approximately 10 nm wide over the range of 0.4-2.4 μm. This type of data display is referred to as a display in spectral space.

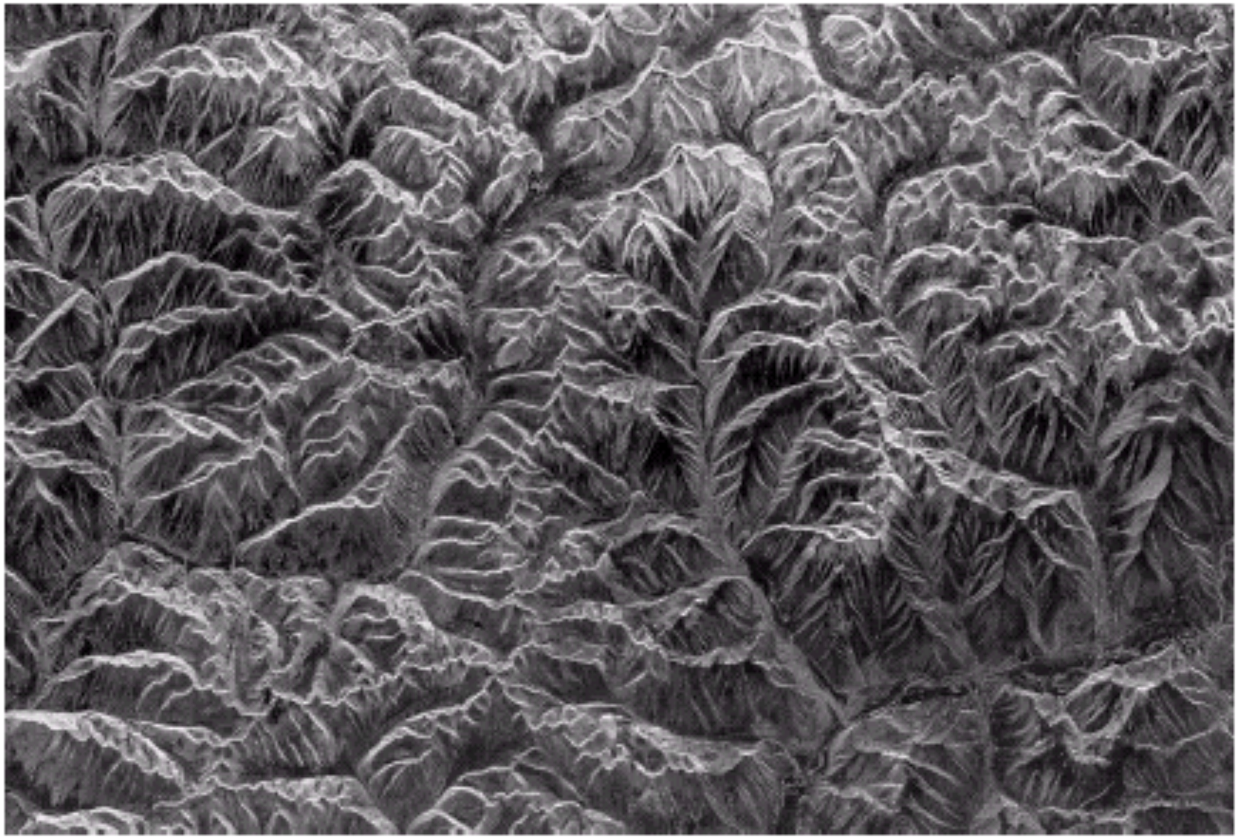
Other types of sensors

FIGURE 1.12 Infrared satellite images of the Americas. The small gray map is provided for reference. (Courtesy of NOAA.)

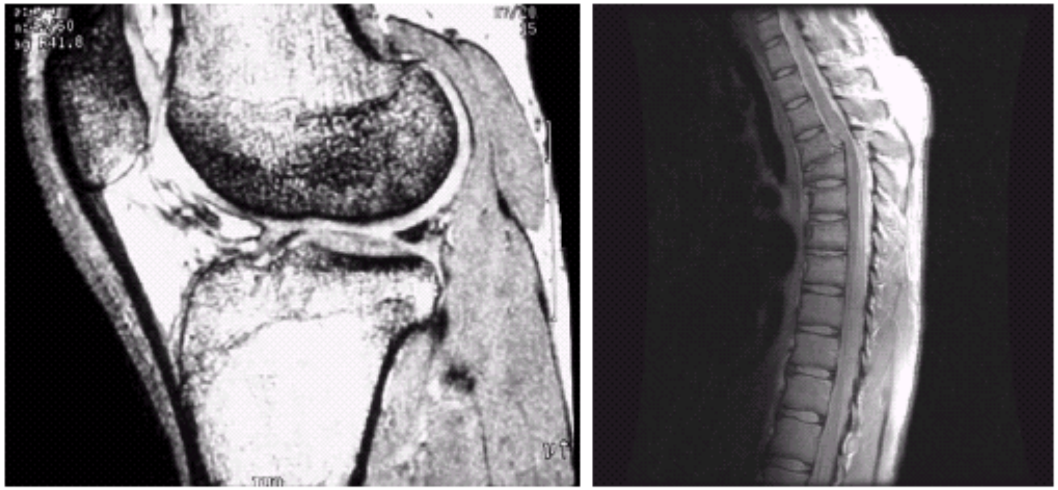
FIGURE 1.13 Infrared satellite images of the remaining populated part of the world. The small gray map is provided for reference. (Courtesy of NOAA.)

Other types of sensors

FIGURE 1.16 Spaceborne radar image of mountains in southeast Tibet. (Courtesy of NASA.)



Other types of sensors



a b

FIGURE 1.17 MRI images of a human (a) knee, and (b) spine. (Image (a) courtesy of Dr. Thomas R. Gest, Division of Anatomical Sciences, University of Michigan Medical School, and (b) Dr. David R. Pickens, Department of Radiology and Radiological Sciences, Vanderbilt University Medical Center.)

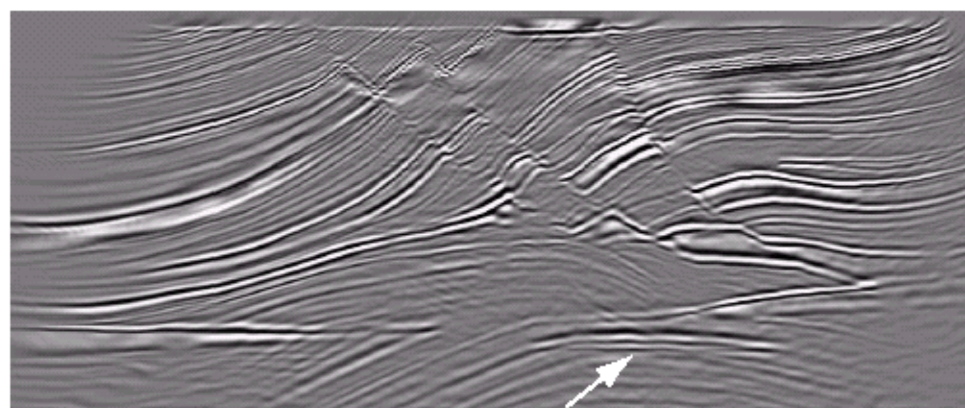
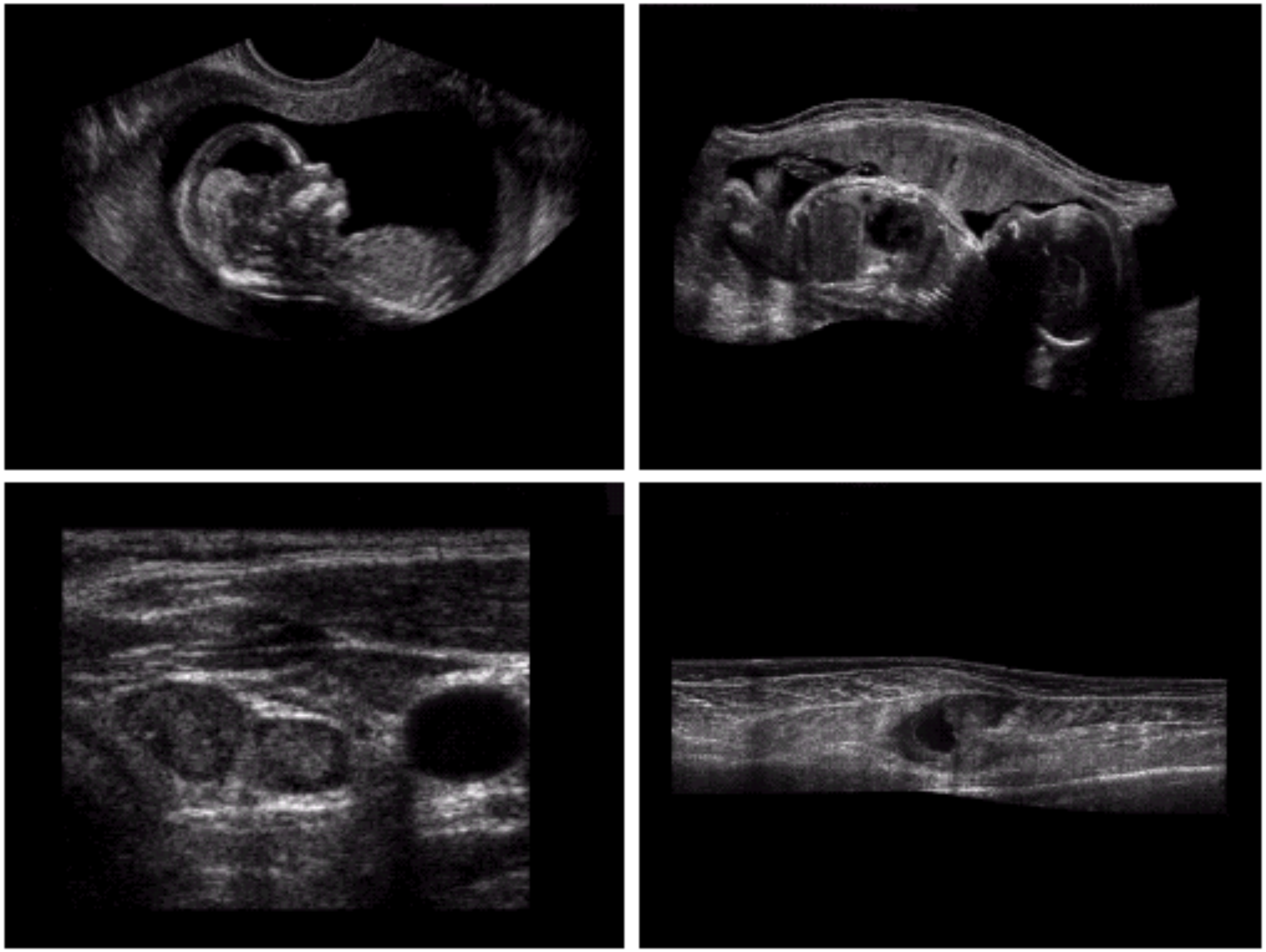


FIGURE 1.19 Cross-sectional image of a seismic model. The arrow points to a hydrocarbon (oil and/or gas) trap. (Courtesy of Dr. Curtis Ober, Sandia National Laboratories.)

Other types of sensors



a b
c d

FIGURE 1.20 Examples of ultrasound imaging. (a) Baby. (2) Another view of baby. (c) Thyroids. (d) Muscle layers showing lesion. (Courtesy of Siemens Medical Systems, Inc., Ultrasound Group.)

Image formats

- Popular formats:
  - BMP Microsoft Windows bitmap image
  - EPS Adobe Encapsulated PostScript
  - GIF CompuServe graphics interchange format
  - JPEG Joint Photographic Experts Group
  - PBM Portable bitmap format (black and white)
  - PGM Portable graymap format (gray scale)
  - PPM Portable pixmap format (color)
  - PNG Portable Network Graphics
  - PS Adobe PostScript
  - TIFF Tagged Image File Format

Image formats

- ASCII or binary
- Number of bits per pixel (color depth)
- Number of bands
- Support for compression (lossless, lossy)
- Support for metadata
- Support for transparency (transparent gifs, etc.)
- …

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

Histogram

A graph!

A form of showing color-intensity distribution in a graph

At each color value, it shows how many pixels have that particuluar color value (a form of frequency)

Image

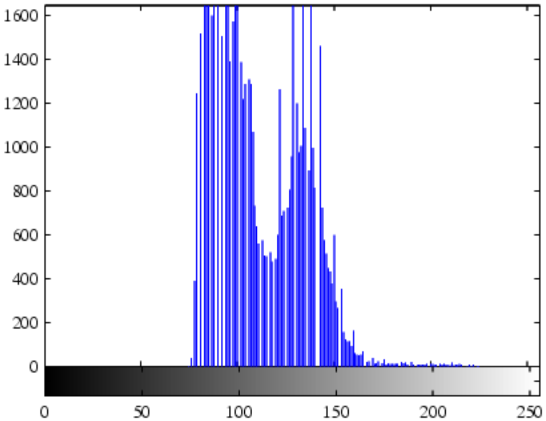


Image - histogram

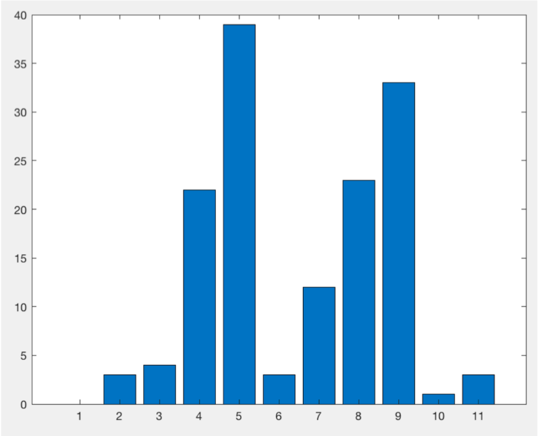
Source: http://www.ece.northwestern.edu/local-apps/matlabhelp/toolbox/images/getting2.html

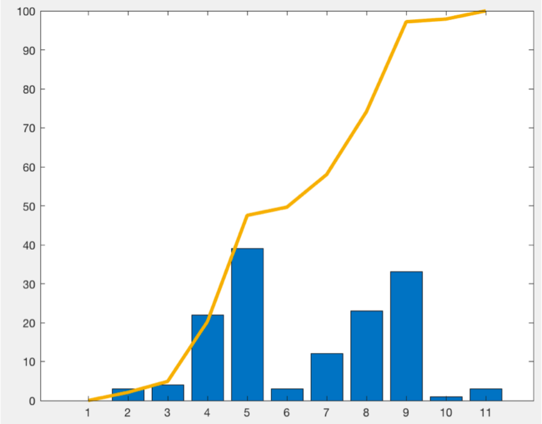
Cumulative Histogram

Bin	Frequency	Cumulative %
1	0	.00%
2	3	2.10%
3	4	4.90%
4	22	20.28%
5	39	47.55%
6	3	49.65%
7	12	58.04%
8	23	74.13%
9	33	97.20%
10	1	97.90%
11	3	100.00%

Example Cumulative Histogram Chart

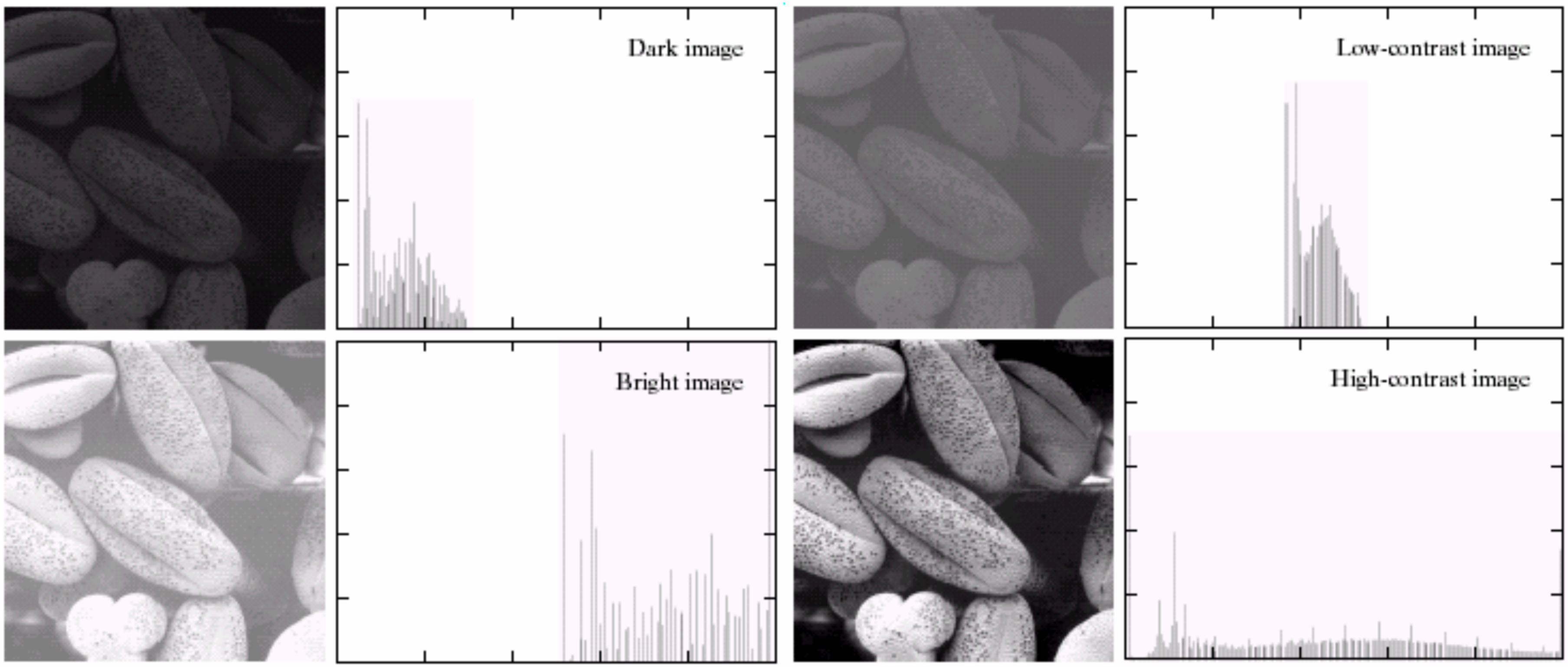
Sum = 143





Source: https://www.tulane.edu/~salem/Cumulative%20Histogram.html

Histogram processing

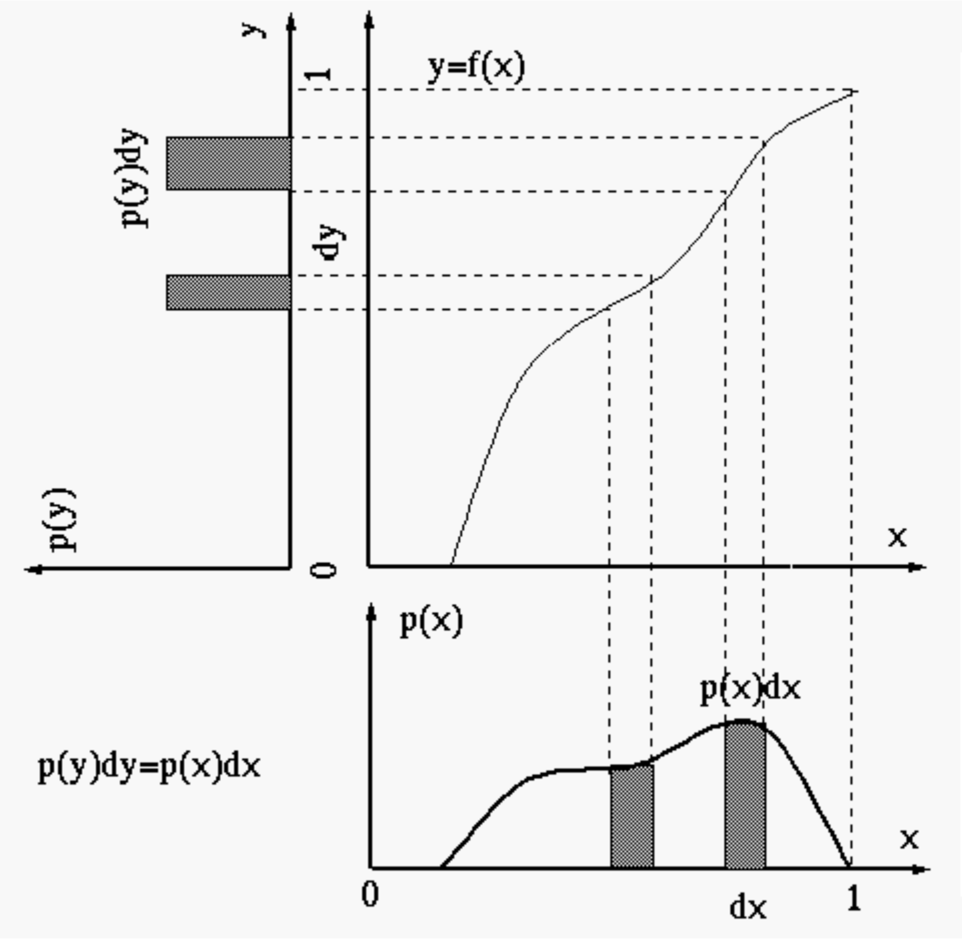


Histogram processing

- Intuitively, we expect that an image whose pixels
  - tend to occupy the entire range of possible gray levels,
  - tend to be distributed uniformly

  will have a high contrast and show a great deal of gray level detail.
- It is possible to develop a transformation function that can achieve this effect using histograms.

Histogram equalization



$p(x), 0 < x < 1$, is the pdf of the input image.
$p(y), 0 < y < 1$, is the pdf of the output image.
Number of pixels mapped from x to y is unchanged, so

$$p(y)dy = p(x)dx.$$

Let $p(y)$ be constant, i.e., $p(y) = 1, 0 < y < 1$. Then,

$$dy = p(x)dx$$

$$\frac{dy}{dx} = p(x)$$

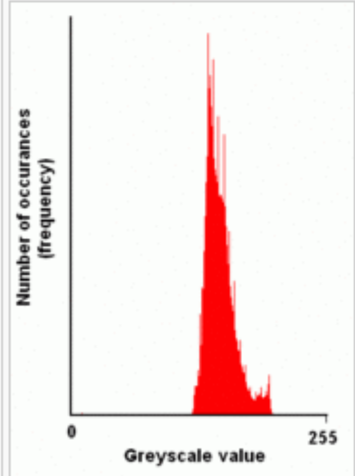
$$y = \int_0^x p(u)du = F(x) - F(0) = F(x)$$

where $F(x)$ is the cdf of the input image.

http://fourier.eng.hmc.edu/e161/lectures/contrast_transform/node3.html

Histogram equalization

An unequalized image



Corresponding histogram

Adapted from Wikipedia

Image enhancement

- The principal objective of enhancement is to process an image so that the result is more suitable than the original for a *specific* application.
- Enhancement can be done in
  - Spatial domain,
  - Frequency domain.
- Common reasons for enhancement include
  - Improving visual quality,
  - Improving machine recognition accuracy.

Image enhancement

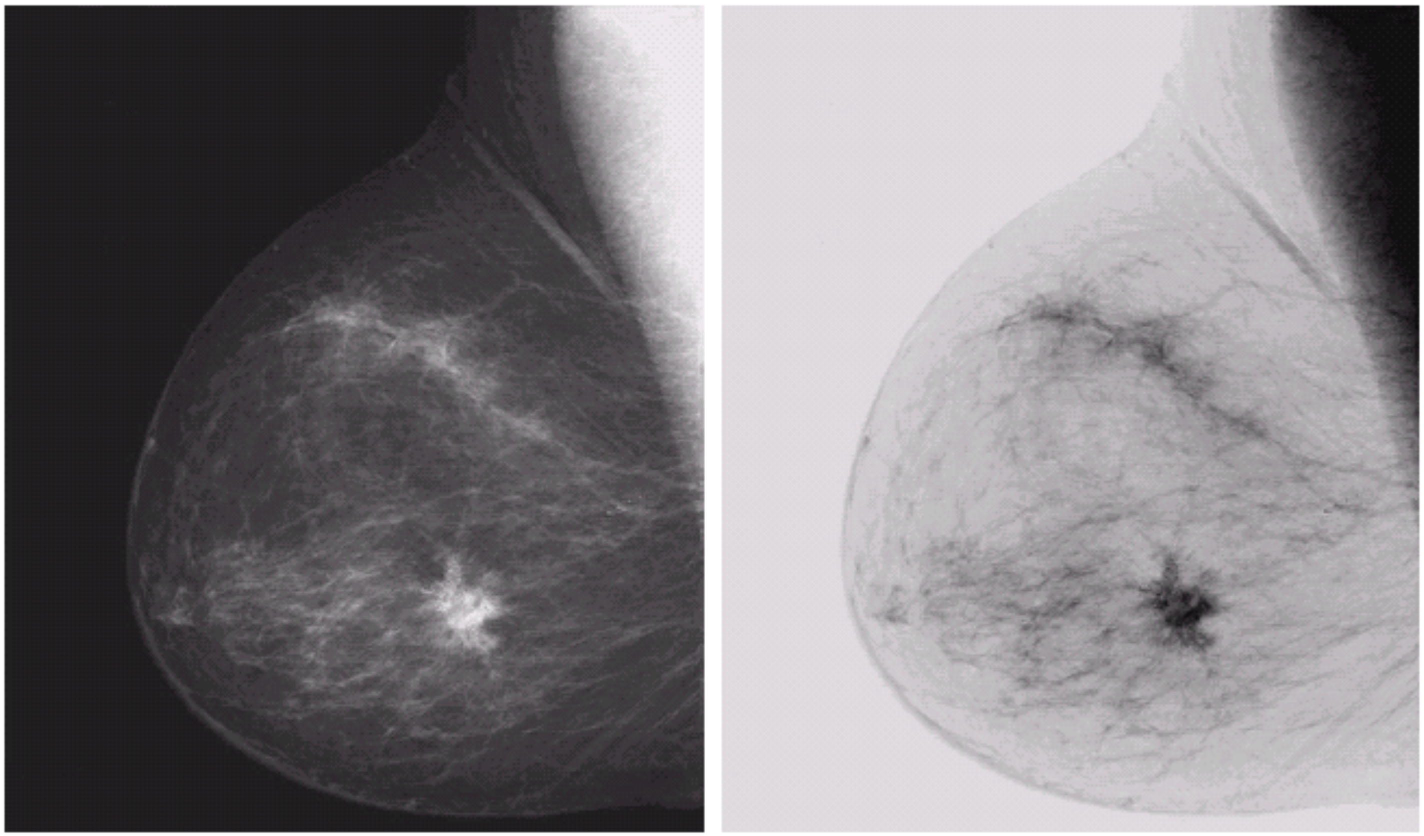


FIGURE 3.4
(a) Original digital mammogram.
(b) Negative image obtained using the negative transformation in Eq. (3.2-1).
(Courtesy of G.E. Medical Systems.)

Text book: Gonzalez & Woods

Image enhancement

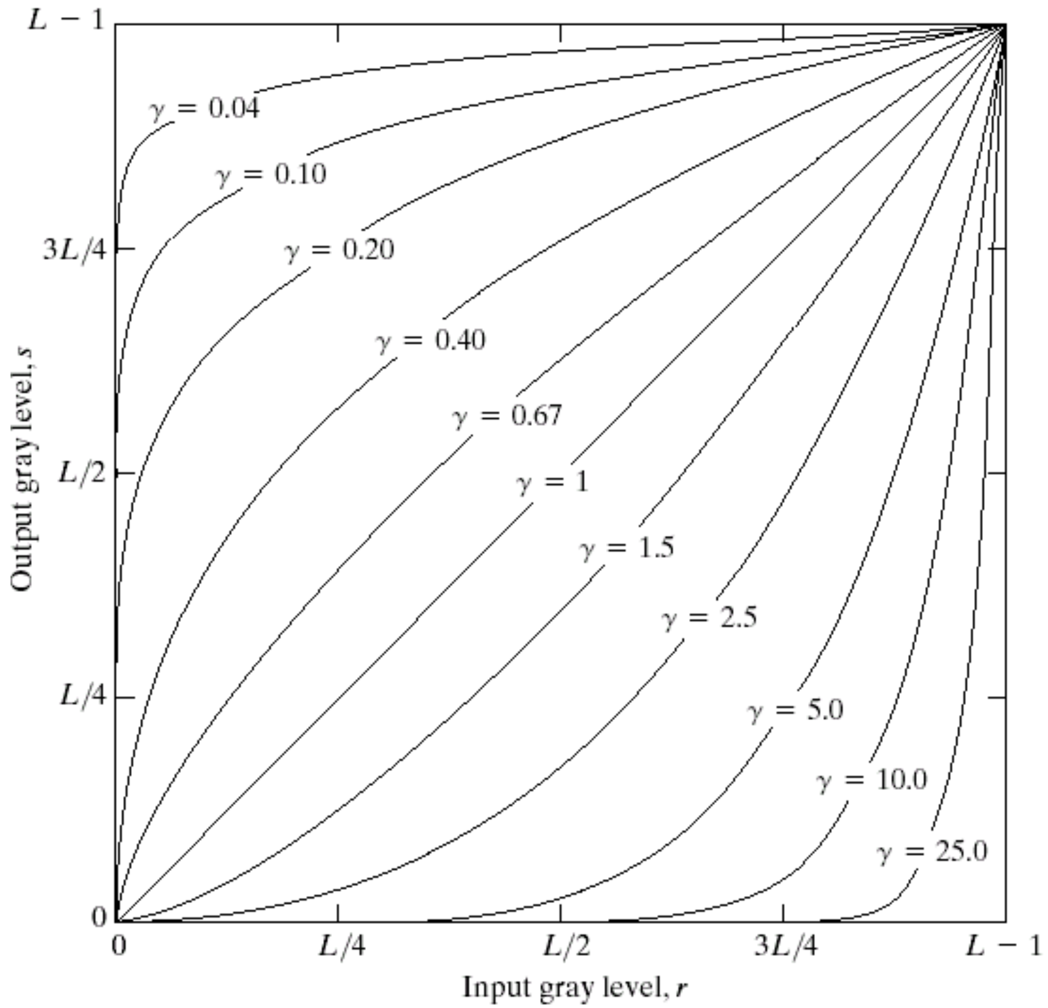
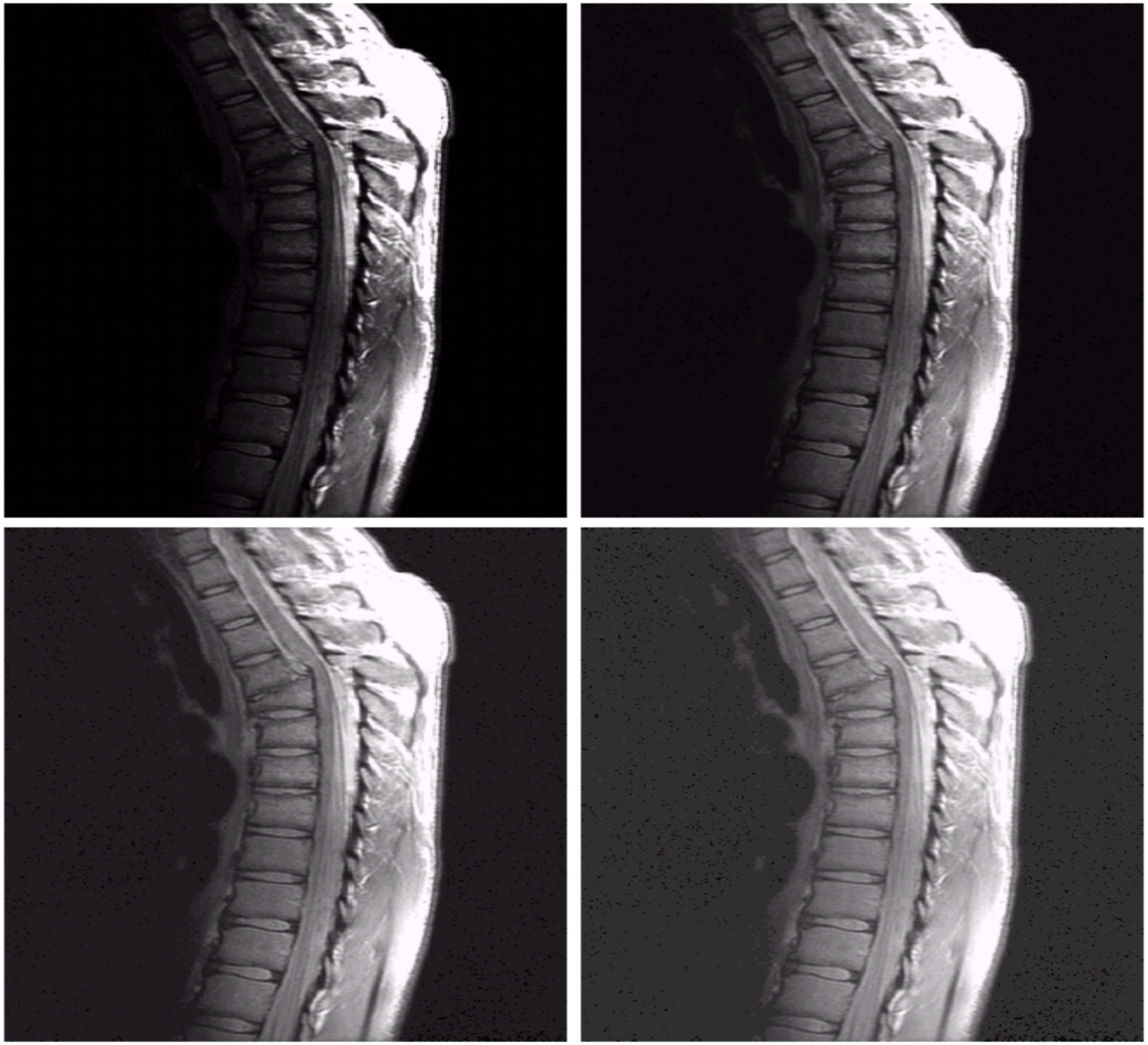


FIGURE 3.6 Plots of the equation $s = cr^{\gamma}$ for various values of γ ($c = 1$ in all cases).

Image enhancement



a b
c d

FIGURE 3.8 (a) Magnetic resonance (MR) image of a fractured human spine. (b)–(d) Results of applying the transformation in Eq. (3.2-3) with $c = 1$ and $\gamma = 0.6, 0.4$, and 0.3, respectively. (Original image for this example courtesy of Dr. David R. Pickens, Department of Radiology and Radiological Sciences, Vanderbilt University Medical Center.)

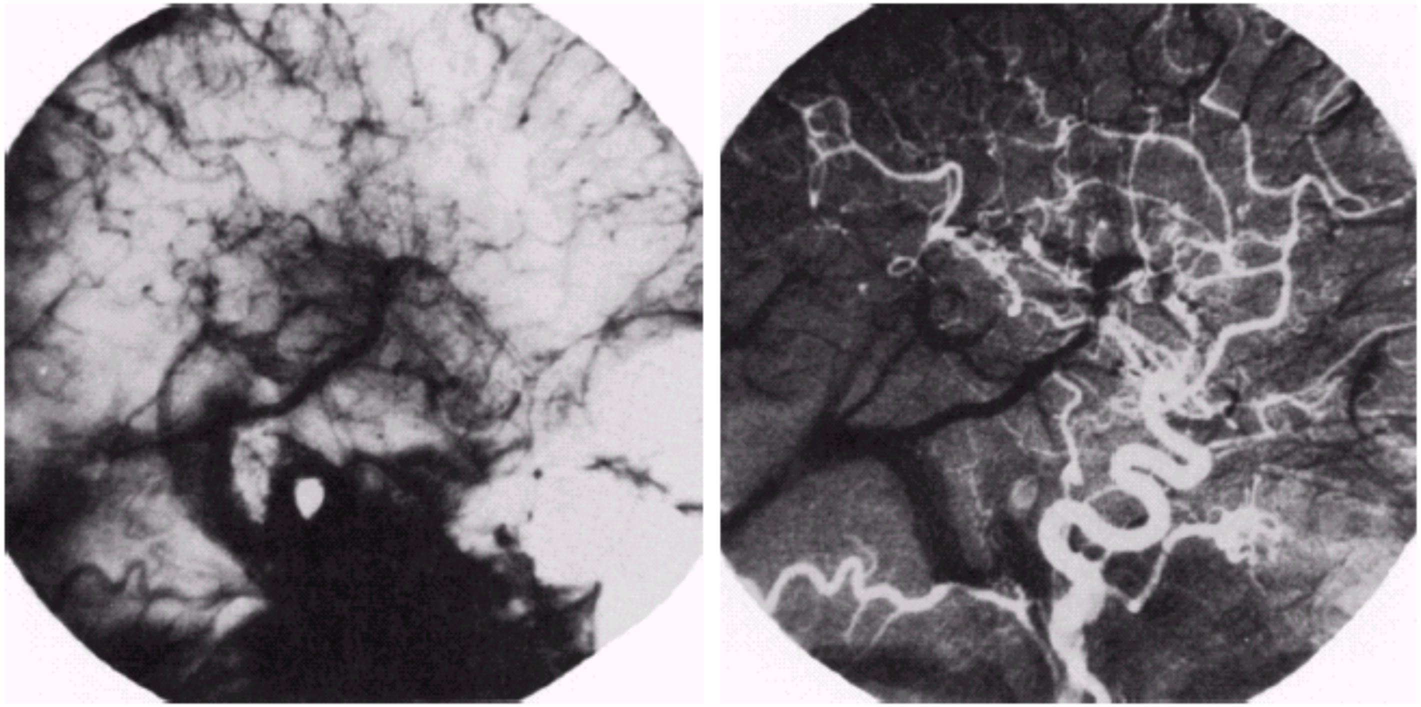
Image enhancement

- Contrast stretching:

$$I'[r,c] = \frac{I[r,c] - \min}{\max - \min}$$

$$I'[r,c] = \begin{cases} 0 & I[r,c] \leq \text{low} \\ \frac{I[r,c] - \text{low}}{\text{high} - \text{low}} & \text{low} < I[r,c] < \text{high} \\ 1 & I[r,c] \geq \text{high} \end{cases}$$

Enhancement using arithmetic operations

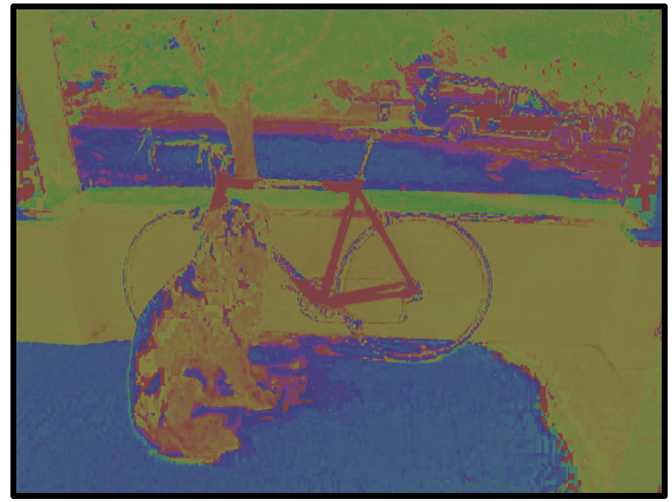


a b

FIGURE 3.29 Enhancement by image subtraction. (a) Mask image. (b) An image (taken after injection of a contrast medium into the bloodstream) with mask subtracted out.

Enhancement using color channels

RGB



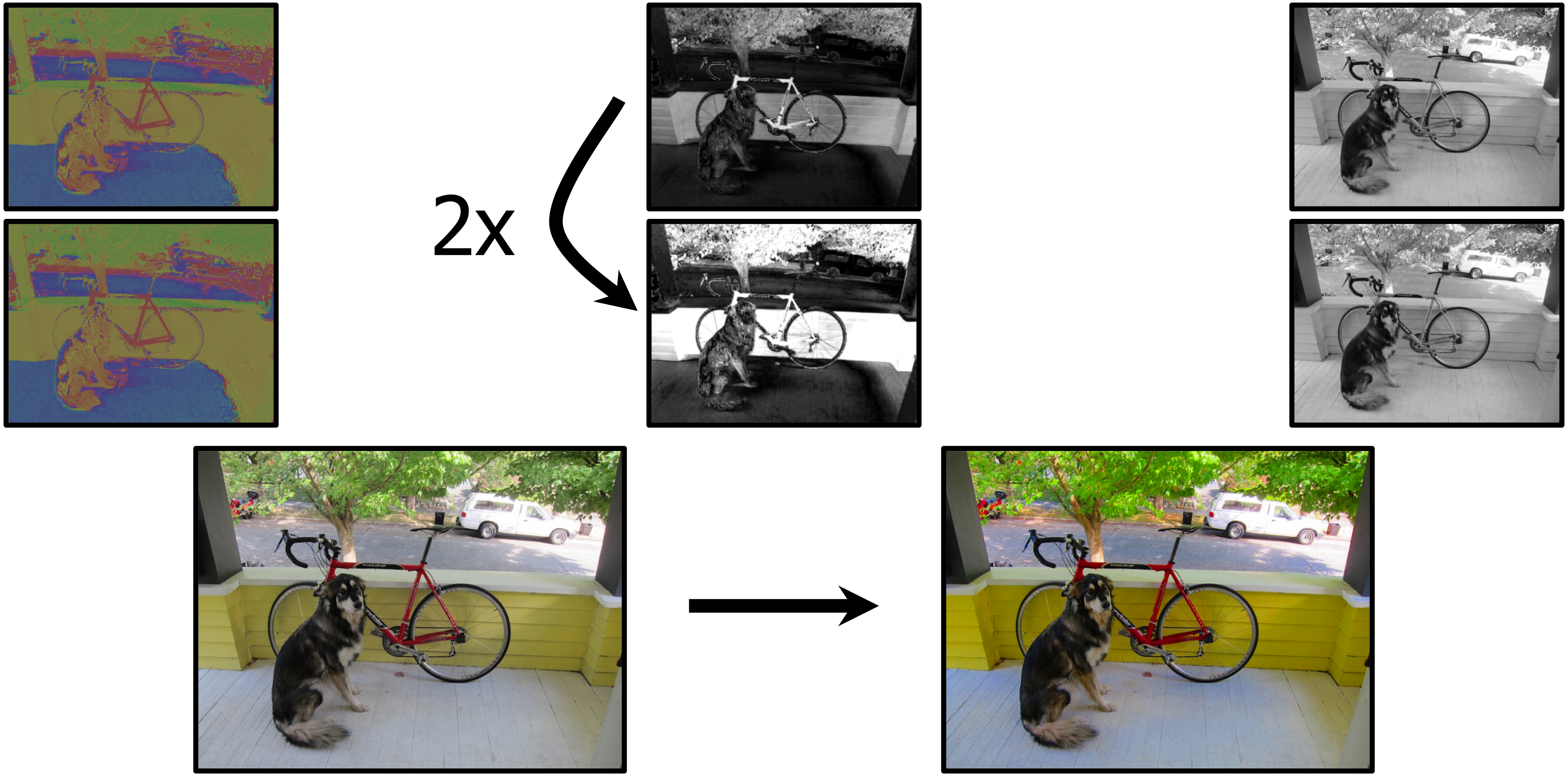
Hue

Saturation

Value

Adapted from Joseph Redmon, U of Washington

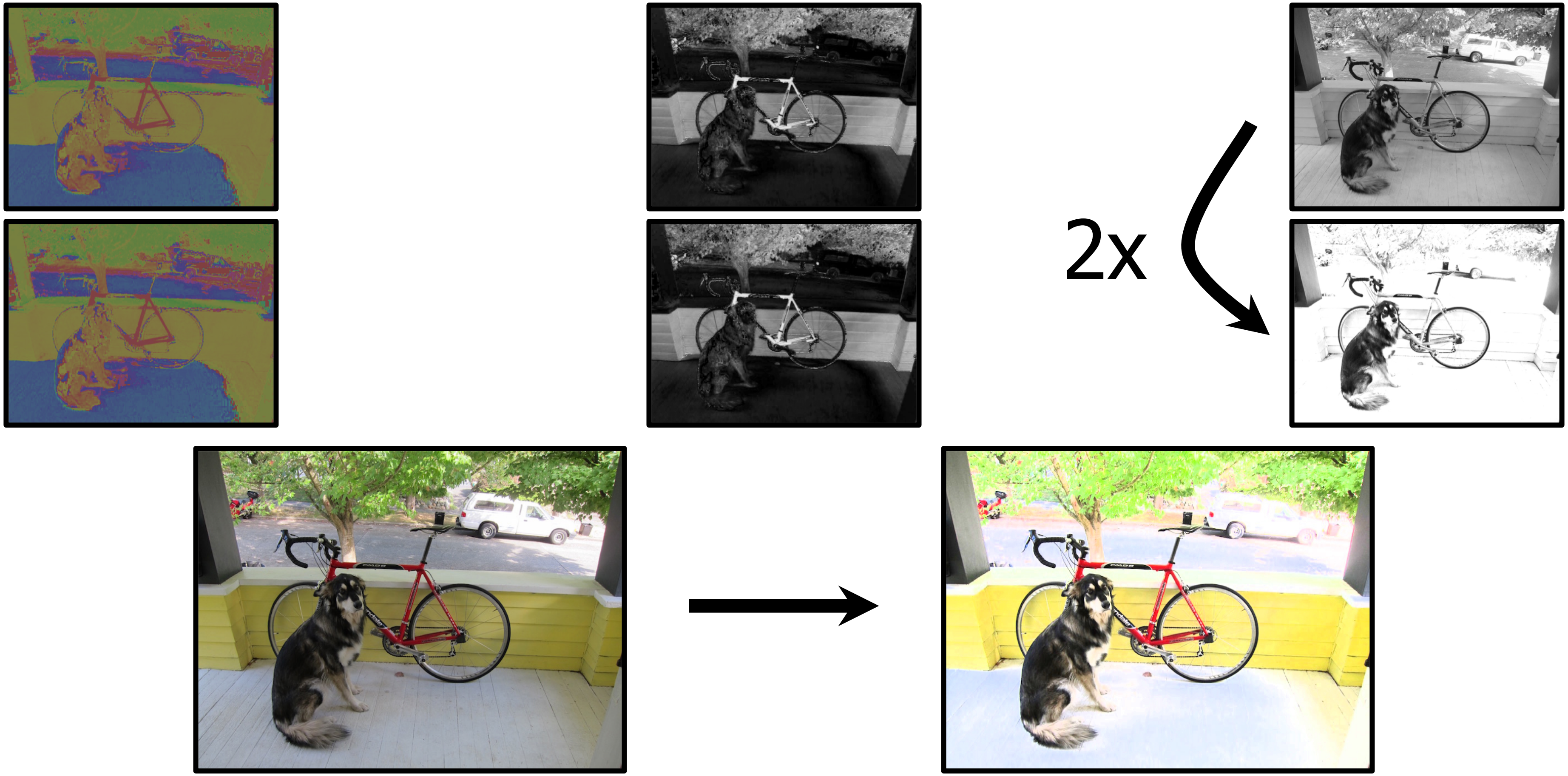
Enhancement using color channels



More saturation -> intense colors

Adapted from Joseph Redmon, U of Washington

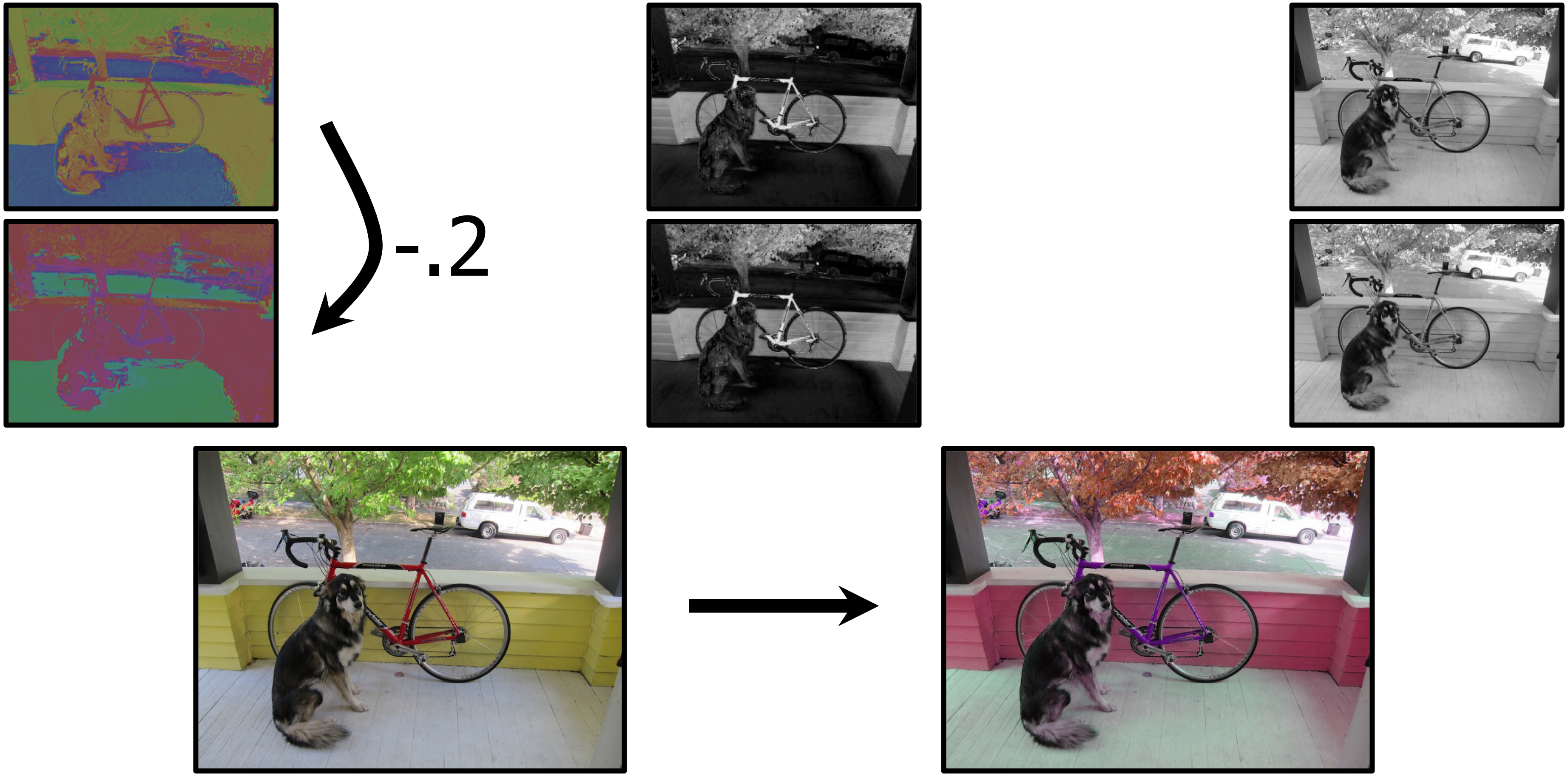
Enhancement using color channels



More value -> lighter image

Adapted from Joseph Redmon, U of Washington

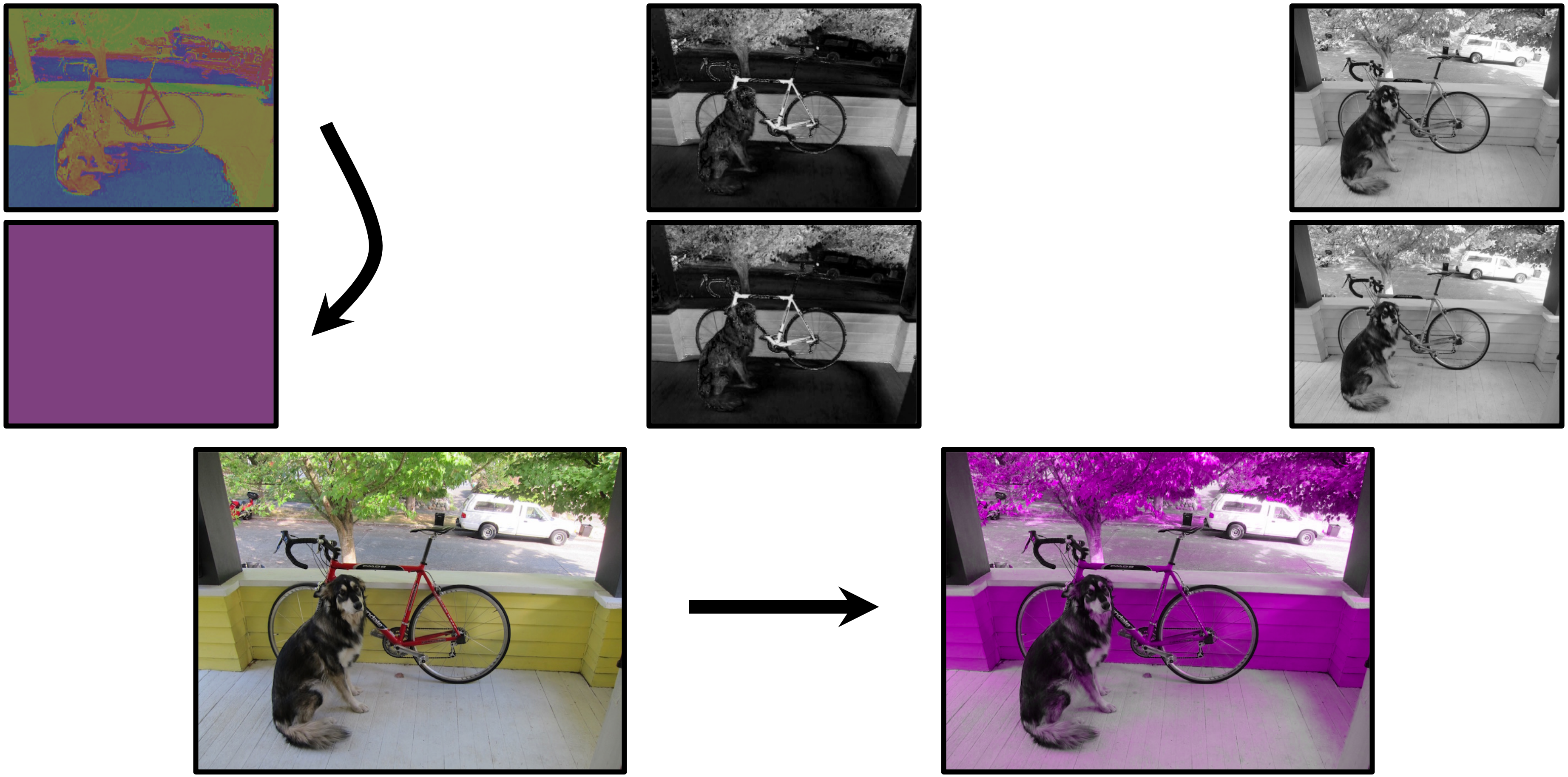
Enhancement using color channels



Shift hue -> shift colors

Adapted from Joseph Redmon, U of Washington

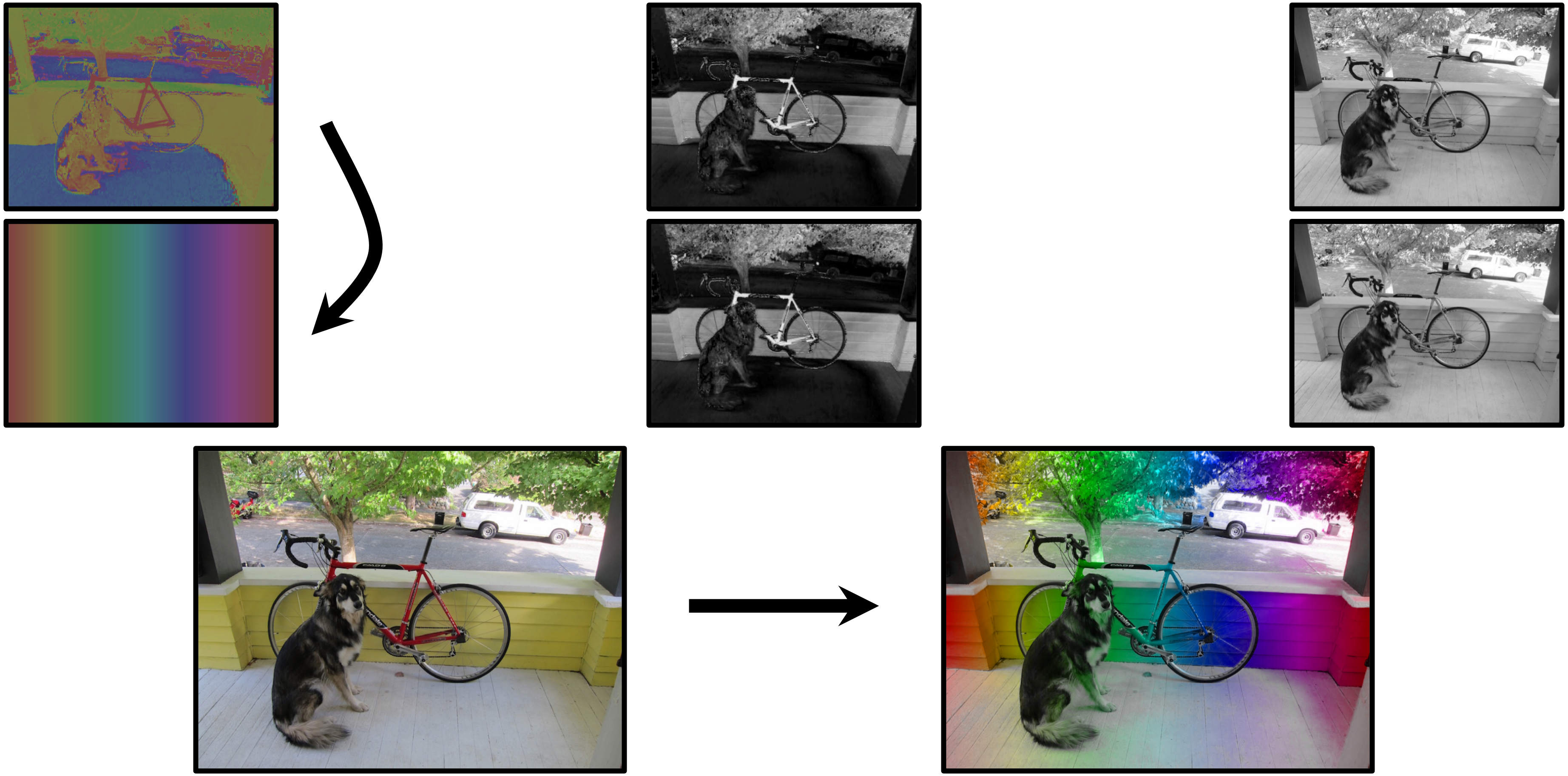
Enhancement using color channels



Set hue to your favorite color

Adapted from Joseph Redmon, U of Washington

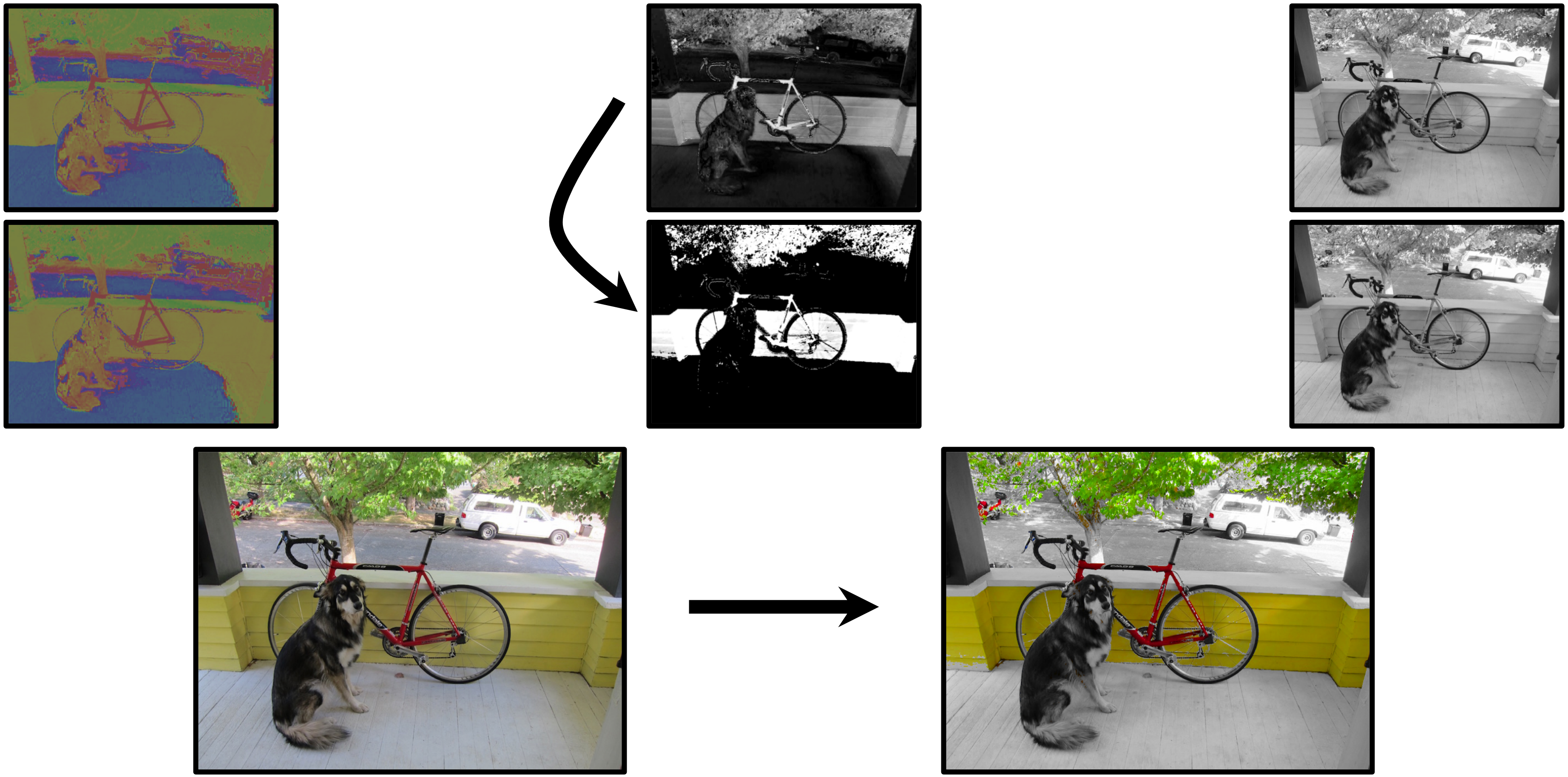
Enhancement using color channels



… or patterns

Adapted from Joseph Redmon, U of Washington

Enhancement using color channels



Increase and threshold saturation

Adapted from Joseph Redmon, U of Washington