Department of Computer Science
University of Bristol

COMS30030 - Image Processing and Computer Vision

Lecture 01

Intro + Image Acquisition & Representation

Majid Mirmehdi | majid@cs.bris.ac.uk

Staff on the unit

Prof. Majid Mirmehdi

Prof. Andrew Calway
UNIT DIRECTOR

Lectures: Mondays 2pm – PHYS BLDG G12 MOTT
Thursdays 11am – FRY BLDG LG02

Labs: **IF YOU ARE ON COMS30087 (20CP):** Thursdays 3-5pm – QUEENS BLDG 1.80

Delivery and Assessment: (Minor) COMS30081

COMS30081 Topics in CS (50%): EXAM ONLY 10CP

The unit aims at providing a first theoretical introduction to classical computational vision: theory, techniques, and algorithms.

Lectures introduce one or more topics and identify key theories, challenges and applications. Students are expected to follow with self study based on problem sheets, revision and further reading.

Problem sheets will be made available for self-study to help in understanding concepts.

Some **past exam papers** will be made available.

ASSESSMENT: 100% Exam

"Well, if the test is multiple choice I choose not to take it."

Delivery and Assessment: (Major) COMS30087

COMS30087: COURSEWORK + 20Cp

Lectures and Labs to give you a practical intro to classical computational vision to help you experience implementing such algorithms.

Implementations will be via OpenCV, which is open-source and freely available for most platforms and languages.

Choose to work on your platform in a language you are most fluent in (at your own risk!); we will support the QUEENS 1.80 Lab setup and the Python interface of OpenCV.

You will work individually during the labs, including parts that will be used for your coursework. You will submit both code and a report for the CW.

ASSESSMENT: 70% Coursework + 30% Mid-Term Exam

Some Suggestions for General Reading

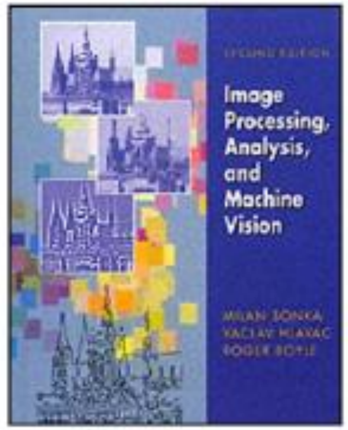
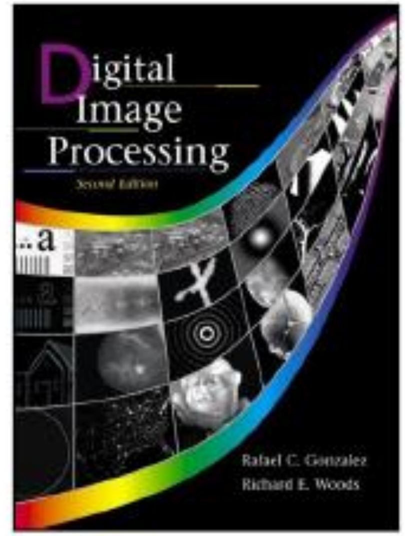
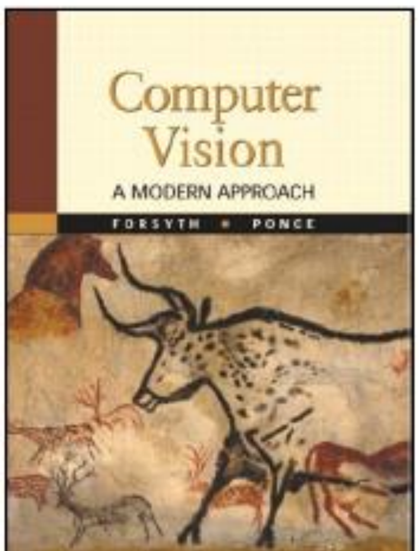


Image Processing Analysis and Machine Vision
by Sonka, Boyle and Hlavac

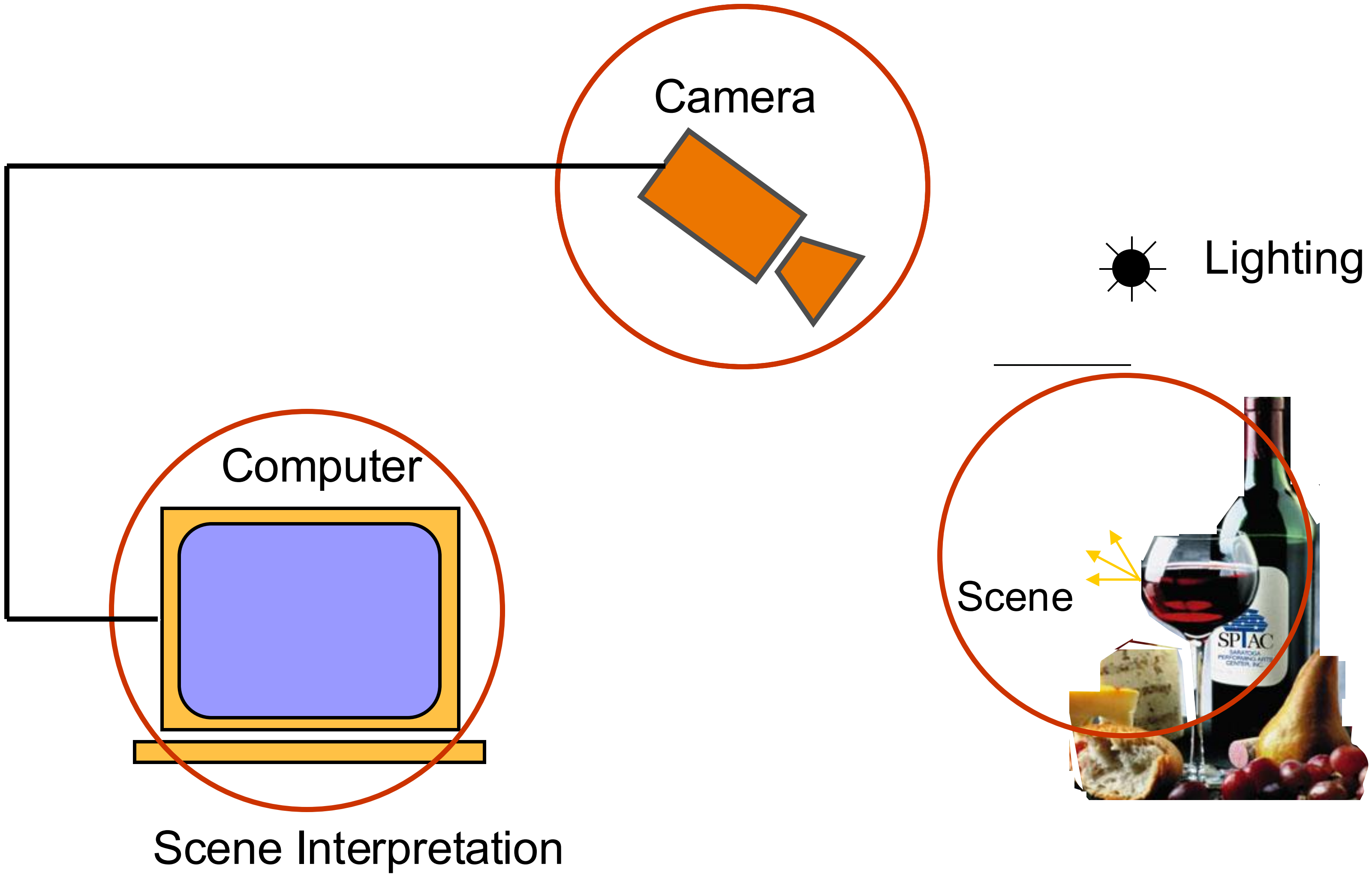


Digital Image Processing
by Gonzalez and Woods



Computer Vision:
A Modern Approach
by Forsyth and Ponce

Components of an imaging system



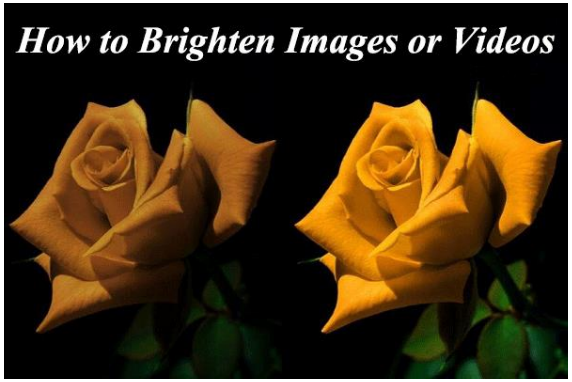
Source: Srinivasa Narasimhan

Images or Image Frames

A single image

A bunch of image frames that make up a video

What is Image Processing?



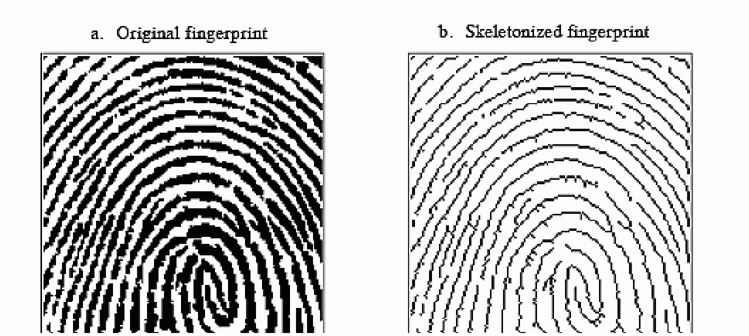
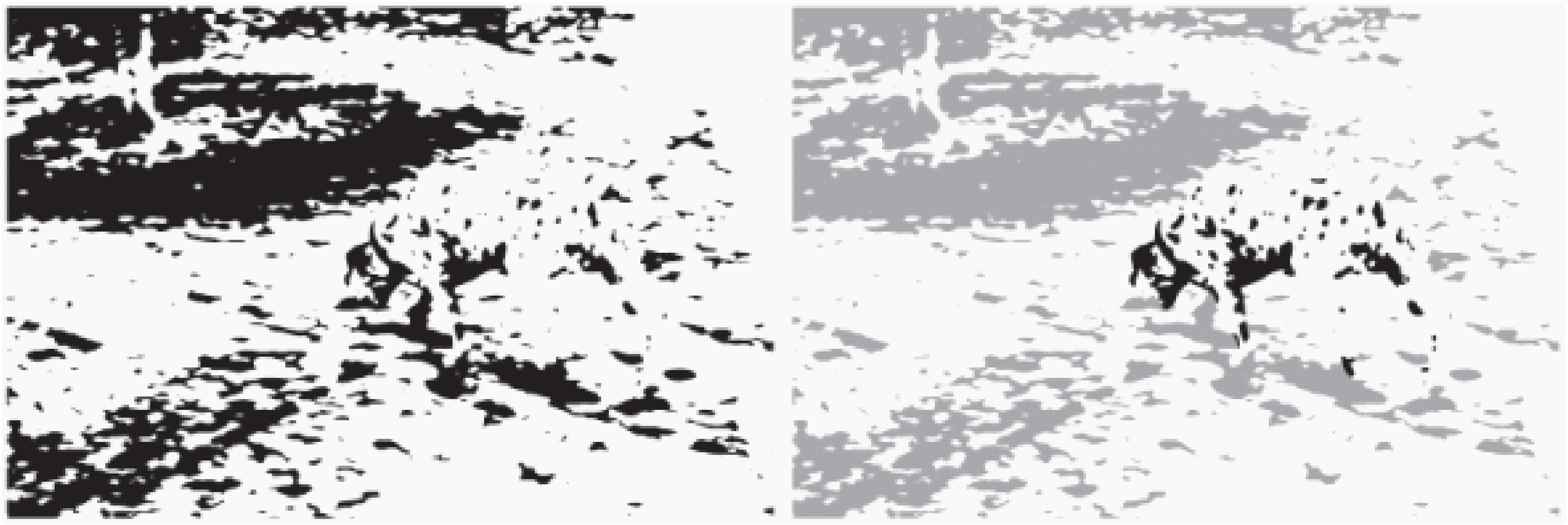


Image Processing... is the digital manipulation of an image to enhance it or extract some useful information from it for further processing.

Source image from: https://www.dspguide.com/ch25/4.htm

What is Computer Vision?



Computer Vision ... attempts to bridge the semantic gap between pixels and their meaning.

What is Computer Vision?

The goal is to perceive the story in the scene!

Computer Vision is hard!

- Vision is an amazing feature of natural intelligence
  - More human brain devoted to vision than anything else

What we see

151	121	1	93	165	204	14	214	28	235
62	67	17	234	27	1	221	37	189	141
20	168	155	113	178	228	25	130	139	221
236	136	158	230	10	5	165	17	30	155
174	148	93	70	95	106	151	10	160	214
103	126	58	16	138	136	98	202	42	233
235	103	52	37	94	104	173	86	223	113
212	15	179	139	48	232	194	46	174	37
119	81	241	172	95	170	29	210	22	194
129	19	33	253	229	5	152	233	52	44
88	200	194	185	140	200	223	190	164	102
113	16	220	215	143	104	247	29	97	203
9	210	102	246	75	9	158	104	184	129
124	52	76	148	249	107	65	216	187	181
6	251	52	208	46	65	185	38	77	240
150	194	28	206	148	197	208	28	74	93
33	183	248	153	168	205	146	100	254	218
130	53	128	212	61	226	201	110	140	183
165	246	22	102	151	213	40	138	8	93
152	251	101	230	23	162	70	238	75	24
187	105	152	83	167	98	125	180	136	121
139	197	55	209	28	124	208	208	104	40
123	19	144	223	62	253	202	108	47	242
220	144	31	16	136	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

29	142	142	75	22	109	111	28	6	5
137	168	41	206	100	70	219	127	114	191
205	154	226	14	89	86	242	67	203	15
247	47	128	123	253	229	181	251	232	28
68	75	24	99	93	63	215	222	102	180
206	246	85	103	215	3	62	64	77	216
126	80	165	149	196	75	186	60	179	193
44	253	164	253	14	216	175	30	46	254
137	23	33	203	241	21	144	63	244	188
32	214	142	121	249	109	99	232	183	71
45	36	152	27	190	137	61	1	237	247
1	14	241	70	2	30	151	67	169	205
32	80	102	32	99	169	91	166	73	214
186	219	9	203	209	240	40	249	119	122
177	252	38	203	119	0	217	139	139	157
154	145	49	251	150	185	235	23	230	156
157	168	223	60	247	118	5	180	16	206
102	208	195	246	140	138	54	191	139	79
17	233	85	169	166	24	49	40	160	97
84	242	247	144	203	3	19	24	198	88
67	67	185	98	123	106	168	105	127	153
37	113	214	252	203	80	146	211	7	16
142	241	66	86	214	133	146	253	189	200
[illegible]	[illegible]	[illegible]	[illegible]	189	54	144	56	59	163

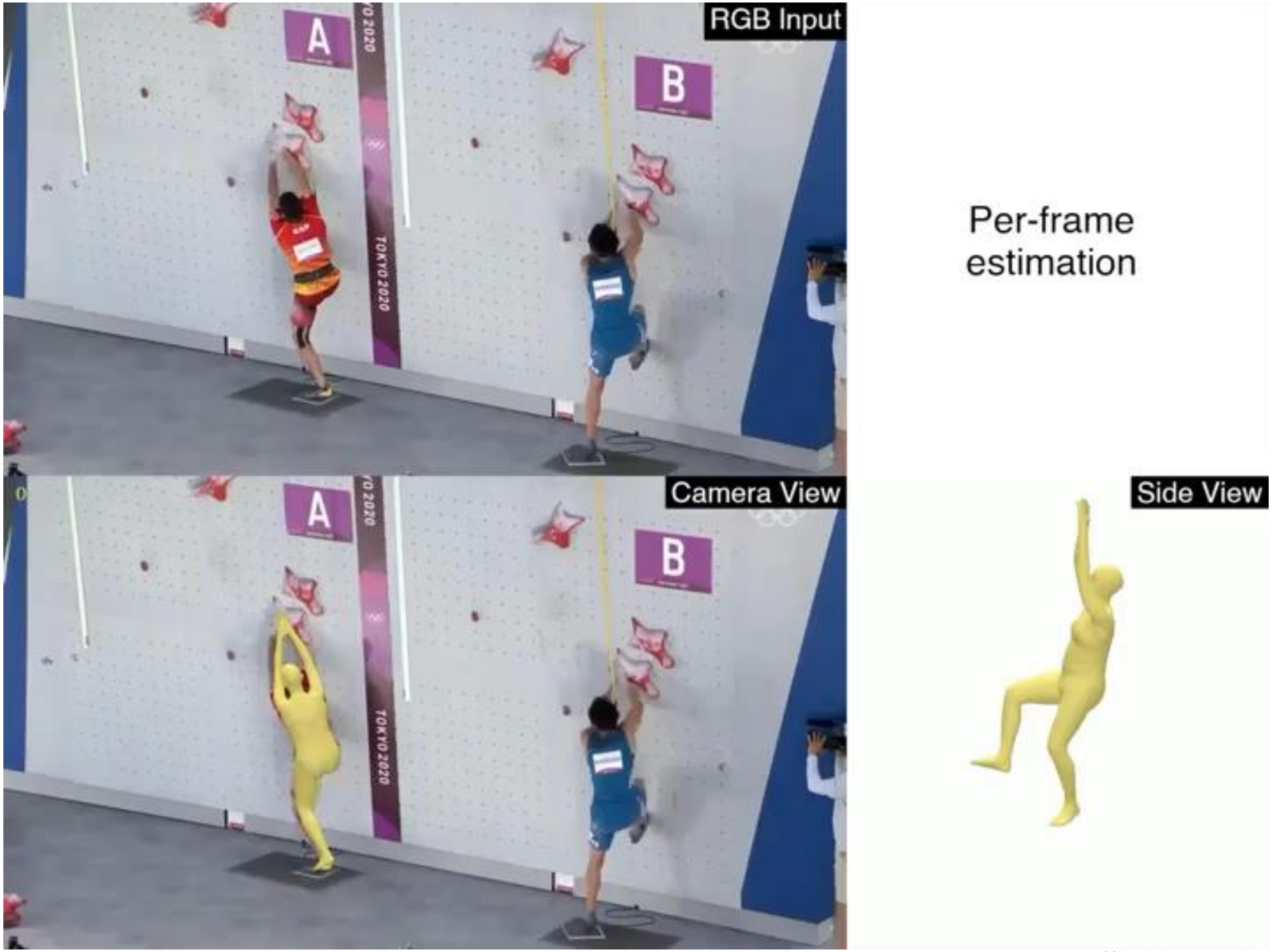
What a computer sees

What is Computer Vision?

Pixels | Features | Models | Meaning

Computer Vision ...
... concerns the study of the theory, engineering and application of artificial systems that extract semantic information from images or other structured, multidimensional data.

Latest Developments: Human Pose/Motion Modelling



https://shubham-goel.github.io/4dhumans/

COMS30030 Topics in a Nutshell

$\frac{1}{2}$

- Acquisition and Representation — Majid
- Image Transforms — Majid
- Edges and Hough Transforms — Majid
- Segmentation — Majid
- Object Detection — Majid and Amir

$\frac{1}{2}$

- Motion Analysis — Andrew
- Stereo Vision — Andrew

So what does the human visual system see?

Or what should an automatic *computer vision system* see?

slide by Fei Fei, Fergus & Torralba

Detection: are there cars?

slide by Fei Fei, Fergus & Torralba

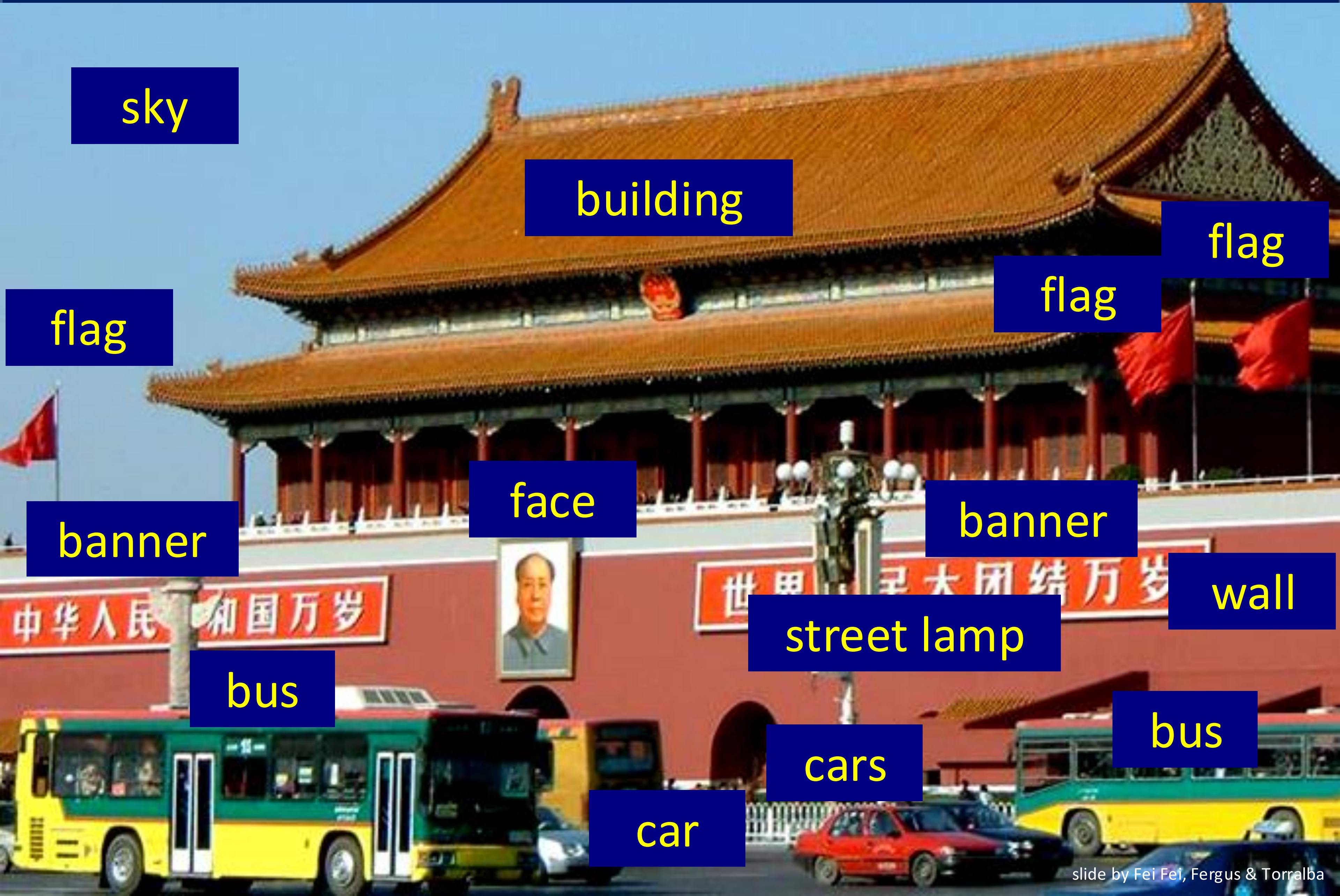
Verification: is that a bus?

slide by Fei Fei, Fergus & Torralba

Identification: is that a picture of Mao?

slide by Fei Fei, Fergus & Torralba

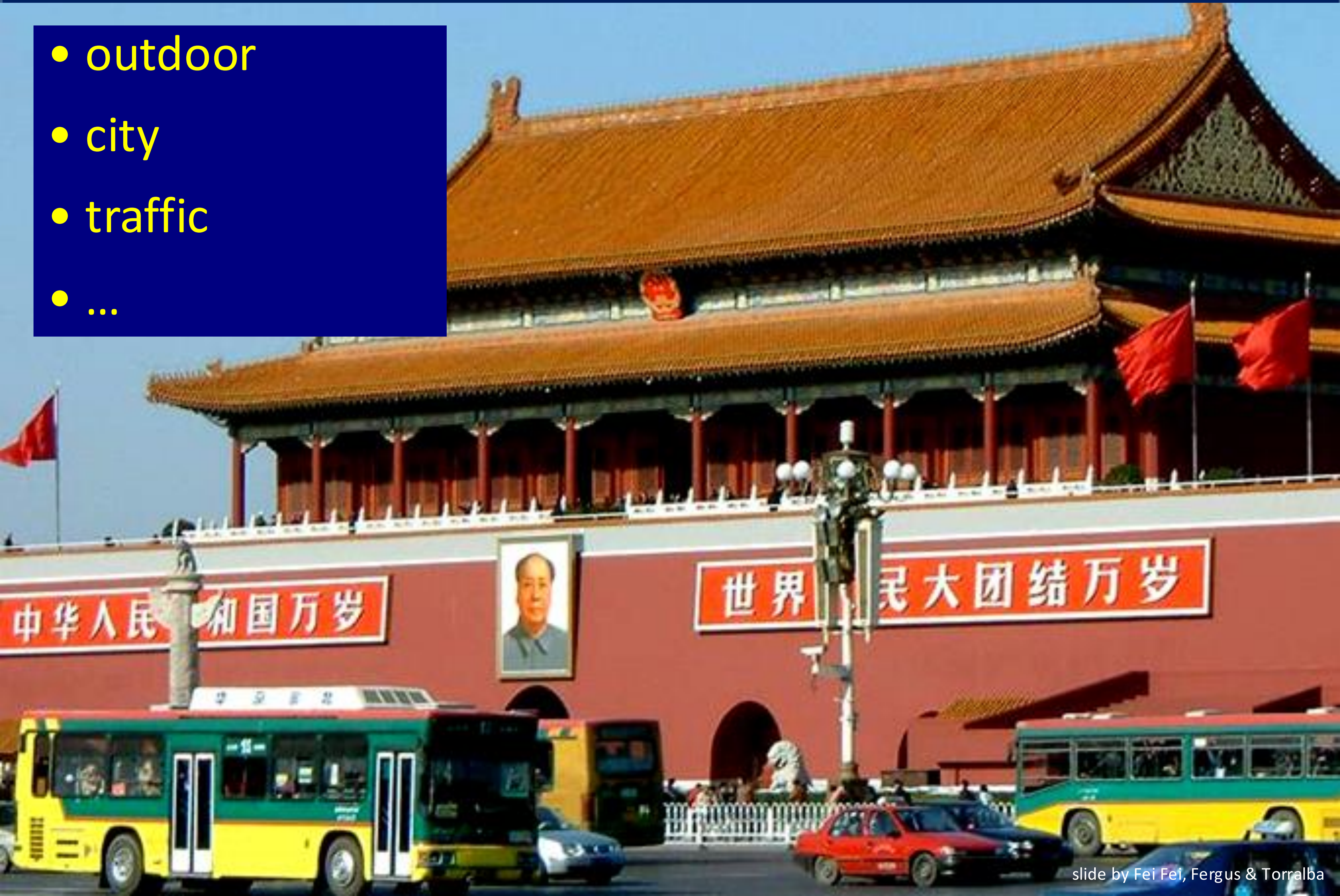
Object categorization



slide by Fei Fei, Fergus & Torralba

Scene and context categorization

- outdoor
- city
- traffic
- …



slide by Fei Fei, Fergus & Torralba

Rough 3D layout, depth ordering

slide by Fei Fei, Fergus & Torralba

Challenges

So what are the challenges faced by a *computer vision system*?

Challenge 1: view-point variation

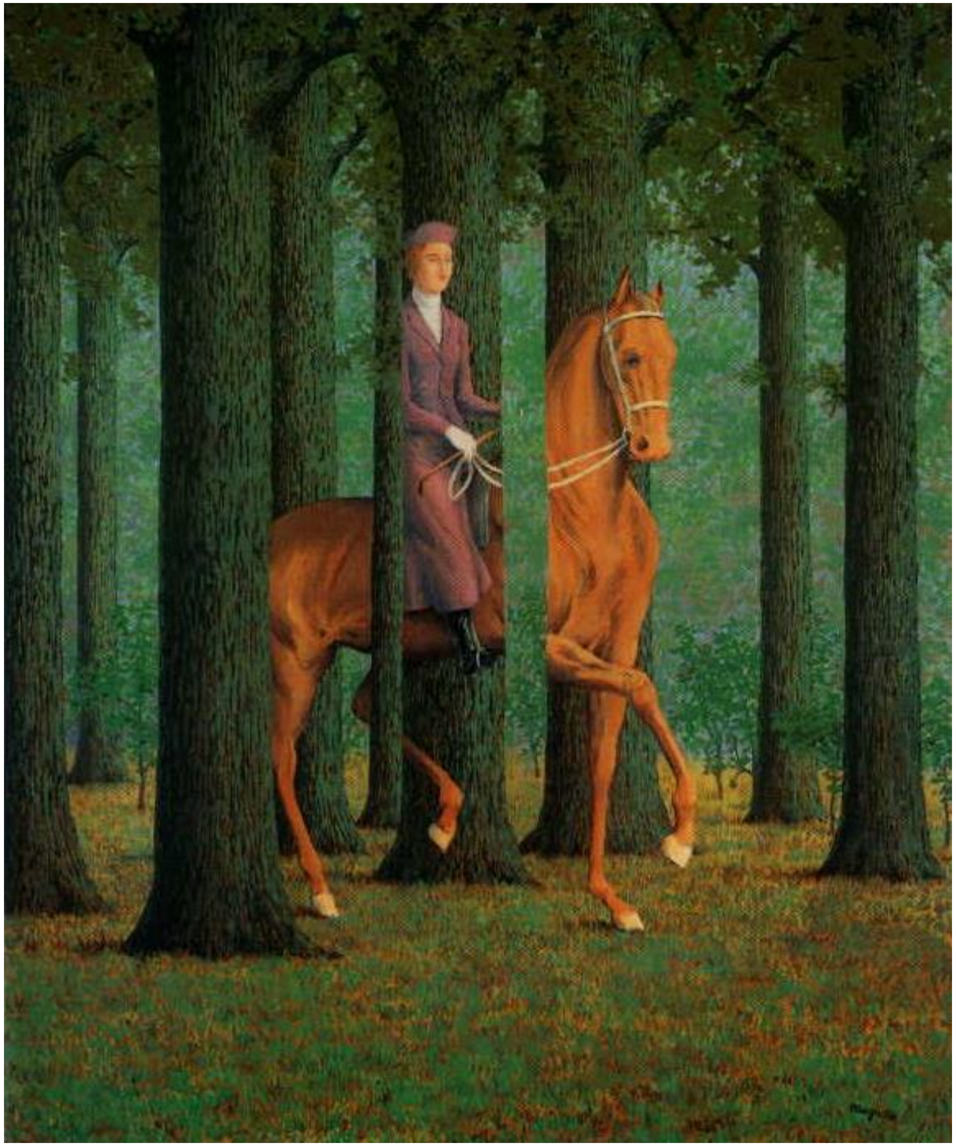
Michelangelo 1475-1564

original Challenge slides by Efros, Ullman, and others

Challenge 2: illumination

original Challenge slides by Efros, Ullman, and others

Challenge 3: occlusion



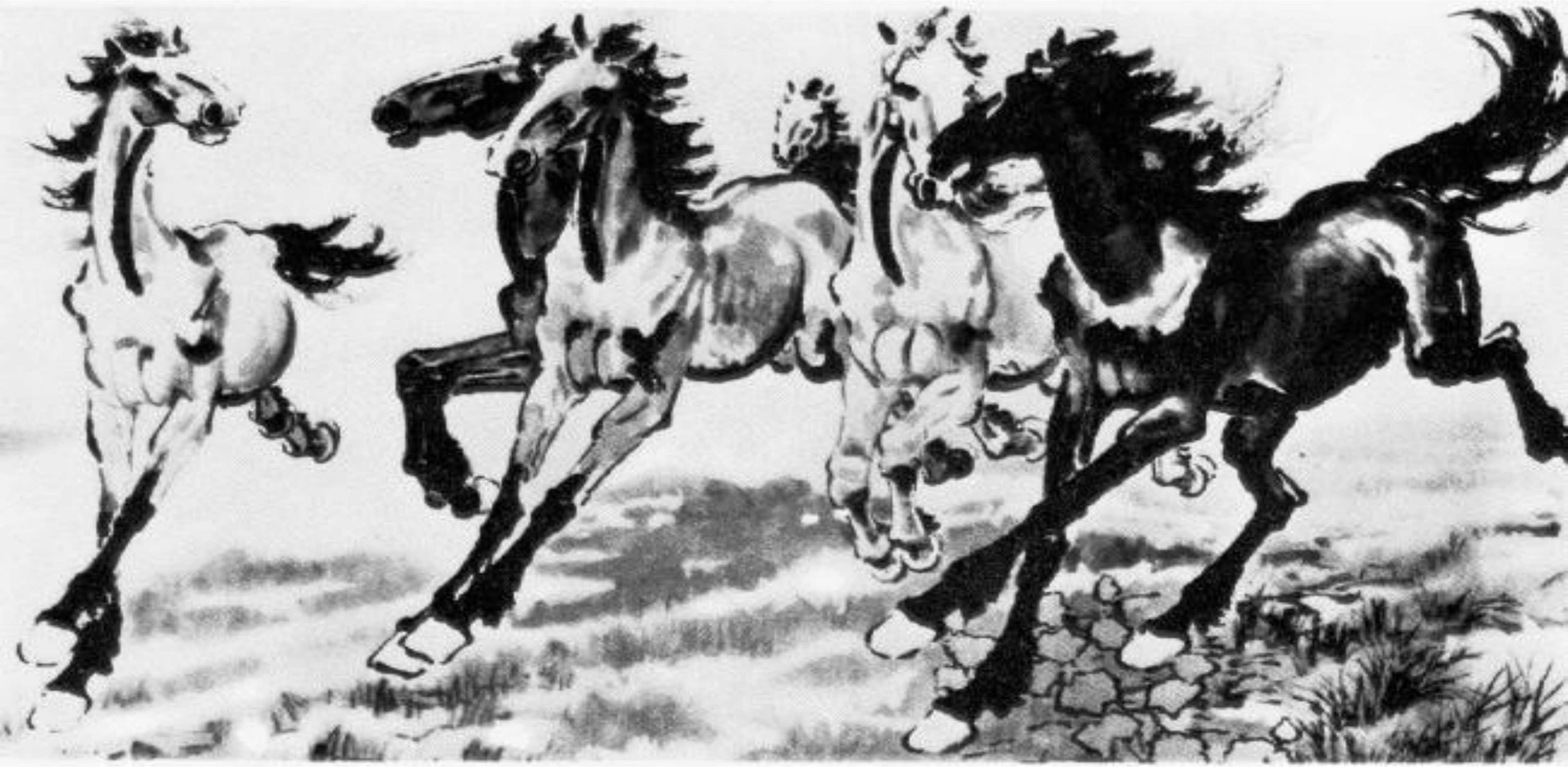
Rene Magritte, 1957

original Challenge slides by Efros, Ullman, and others

Challenge 4: scale

original Challenge slides by Efros, Ullman, and others

Challenge 5: deformation



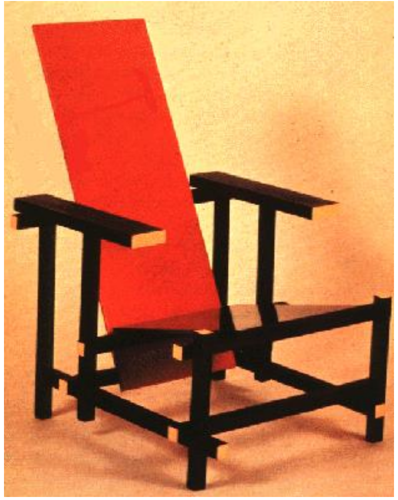
Xu, Beihong 1943

Challenge 6: background clutter

Klimt, 1913

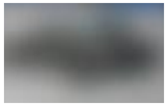
original Challenge slides by Efros, Ullman, and others

Challenge 7: object intra-class variation



original Challenge slides by Efros, Ullman, and others

Challenge 8: local ambiguity



original Challenge slides by Efros, Ullman, and others

The Basics of Image Acquisition and Representation

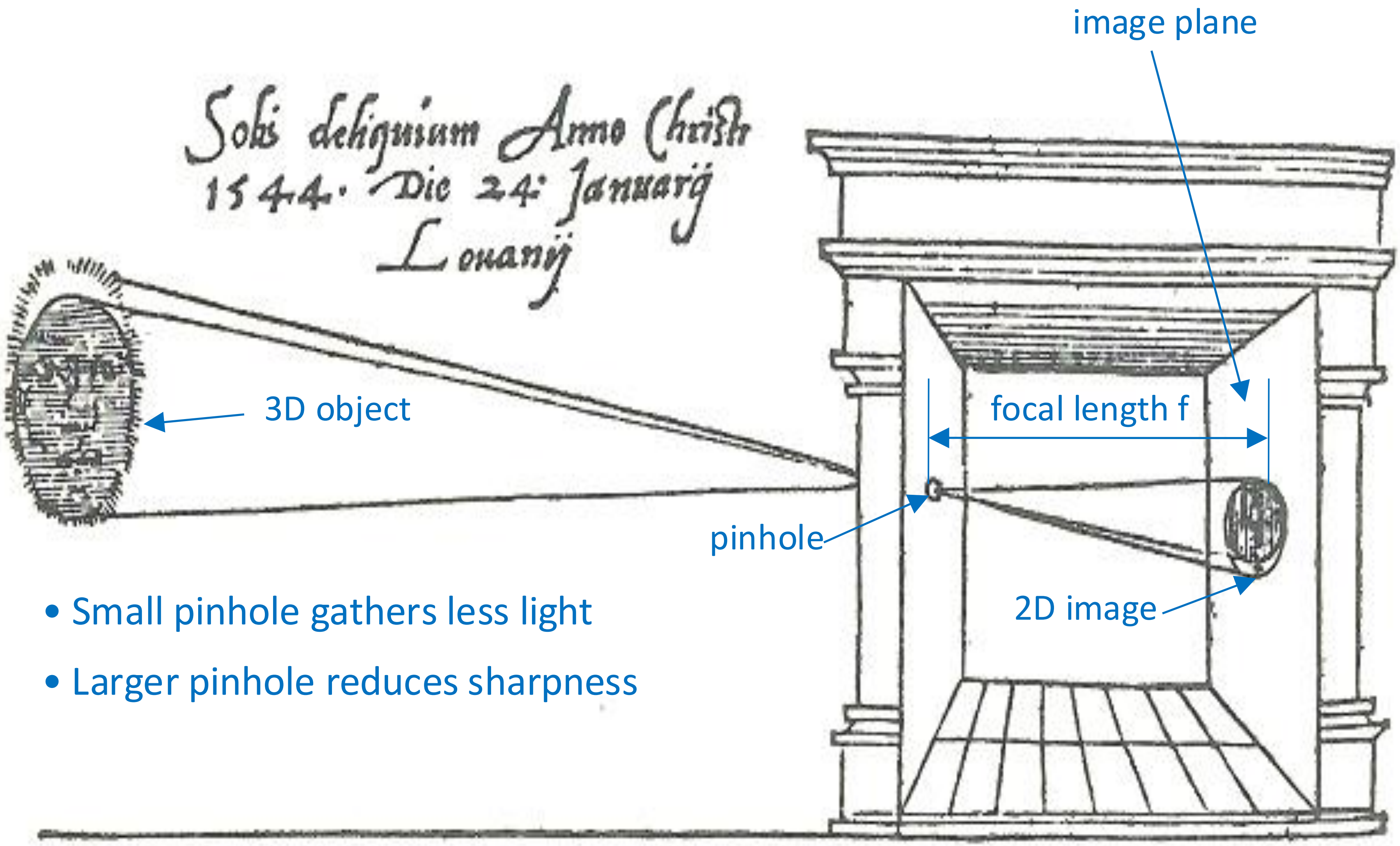
Images as Sensory Data

- How are images acquired?
- Which processes influence digital image formation?

Images as Structured Data

- How can digital images be represented?

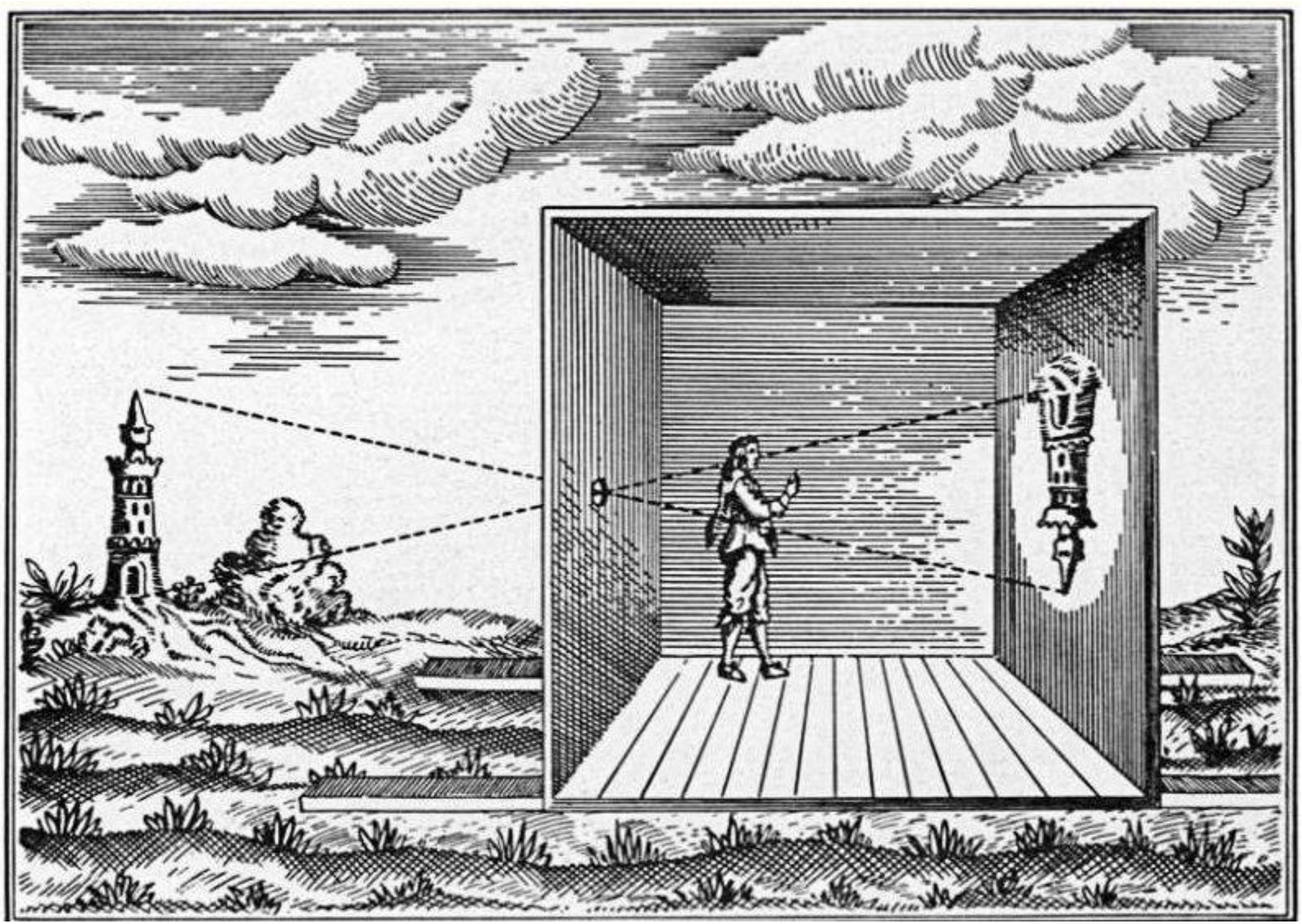
The Camera Obscura (Pinhole Camera)



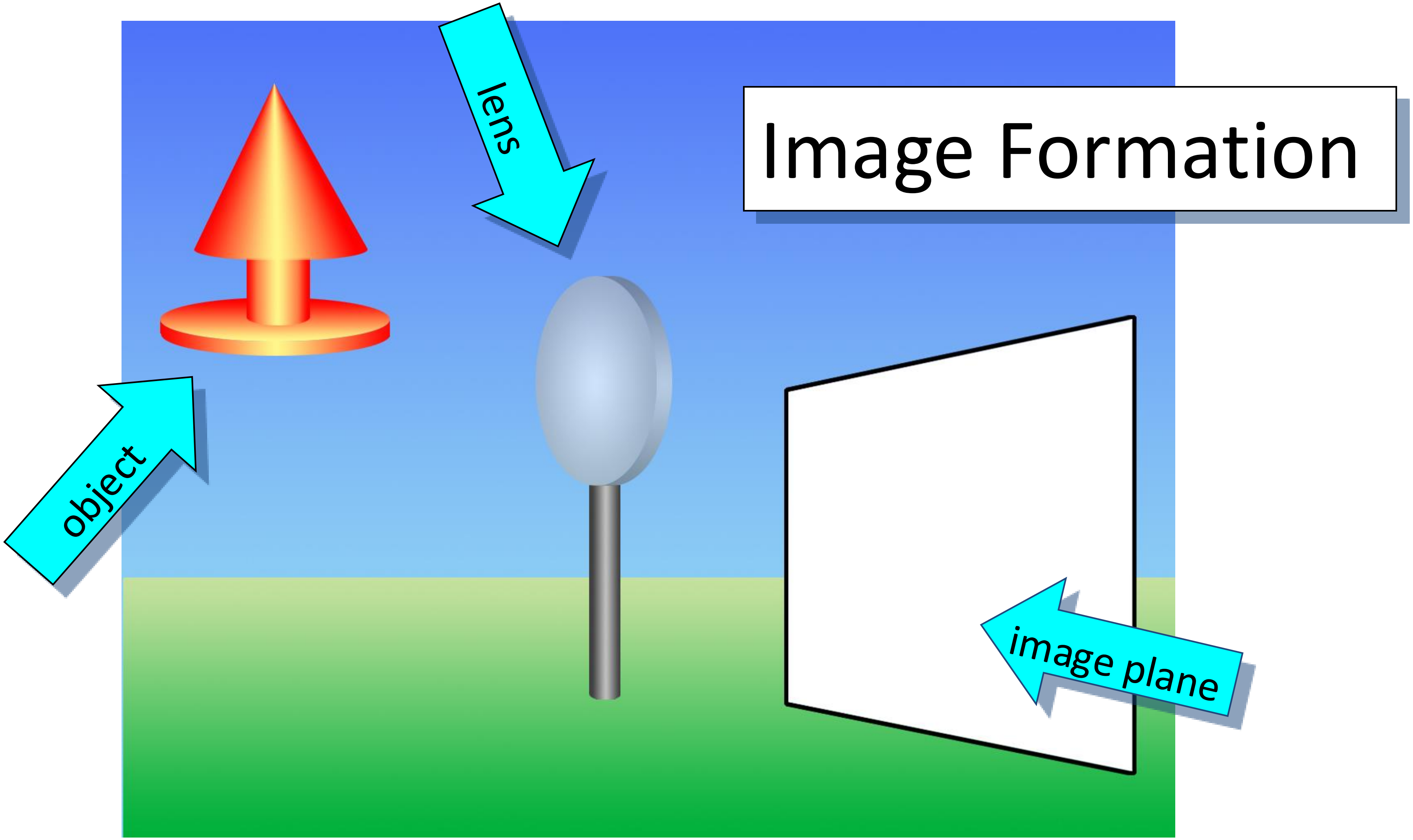
- Small pinhole gathers less light
- Larger pinhole reduces sharpness

First published picture of camera obscura in Gemma Frisius' 1545 book *De Radio Astronomica et Geometrica*

The Camera Obscura (Pinhole Camera)

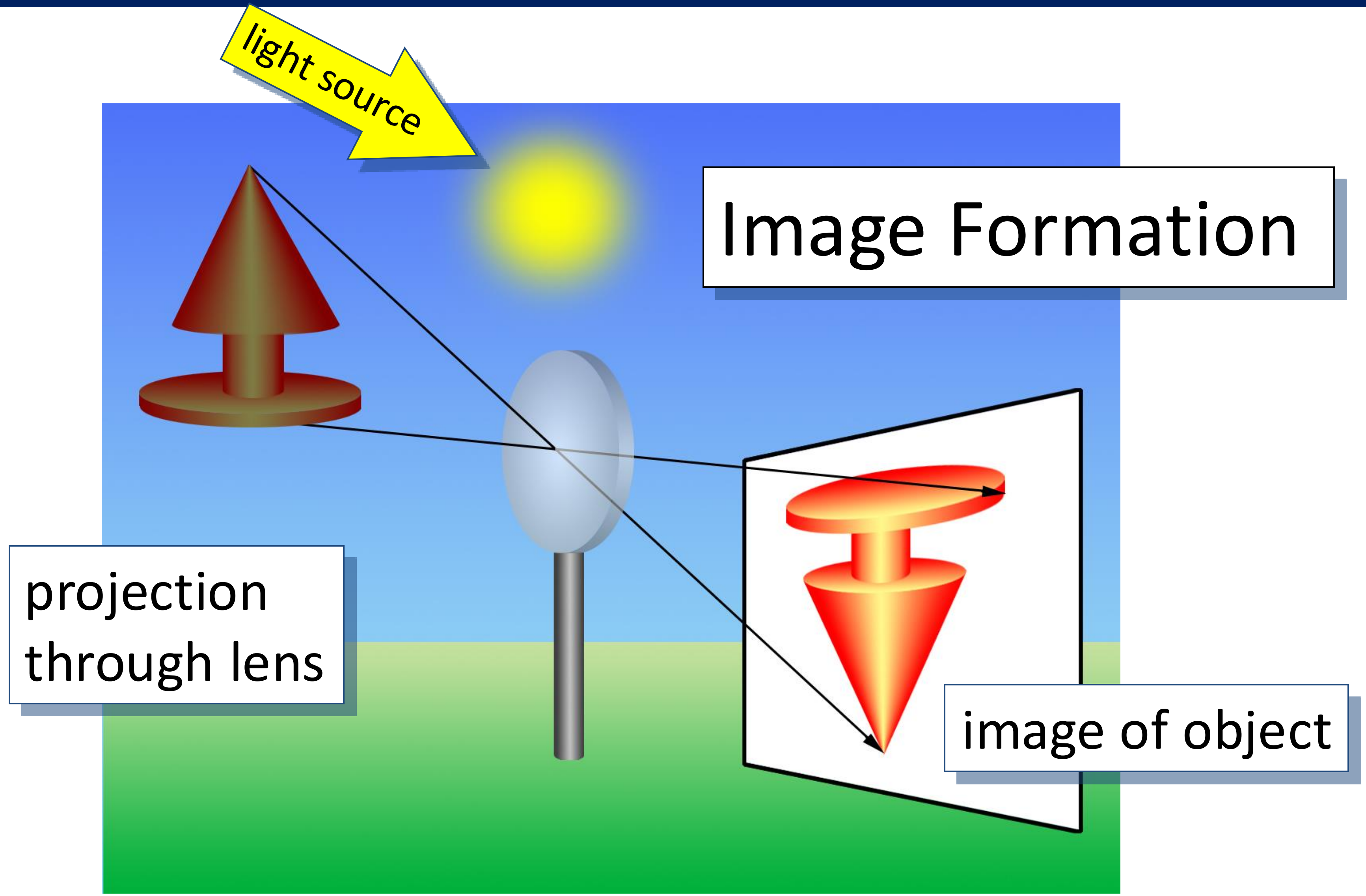


Digital Image Acquisition



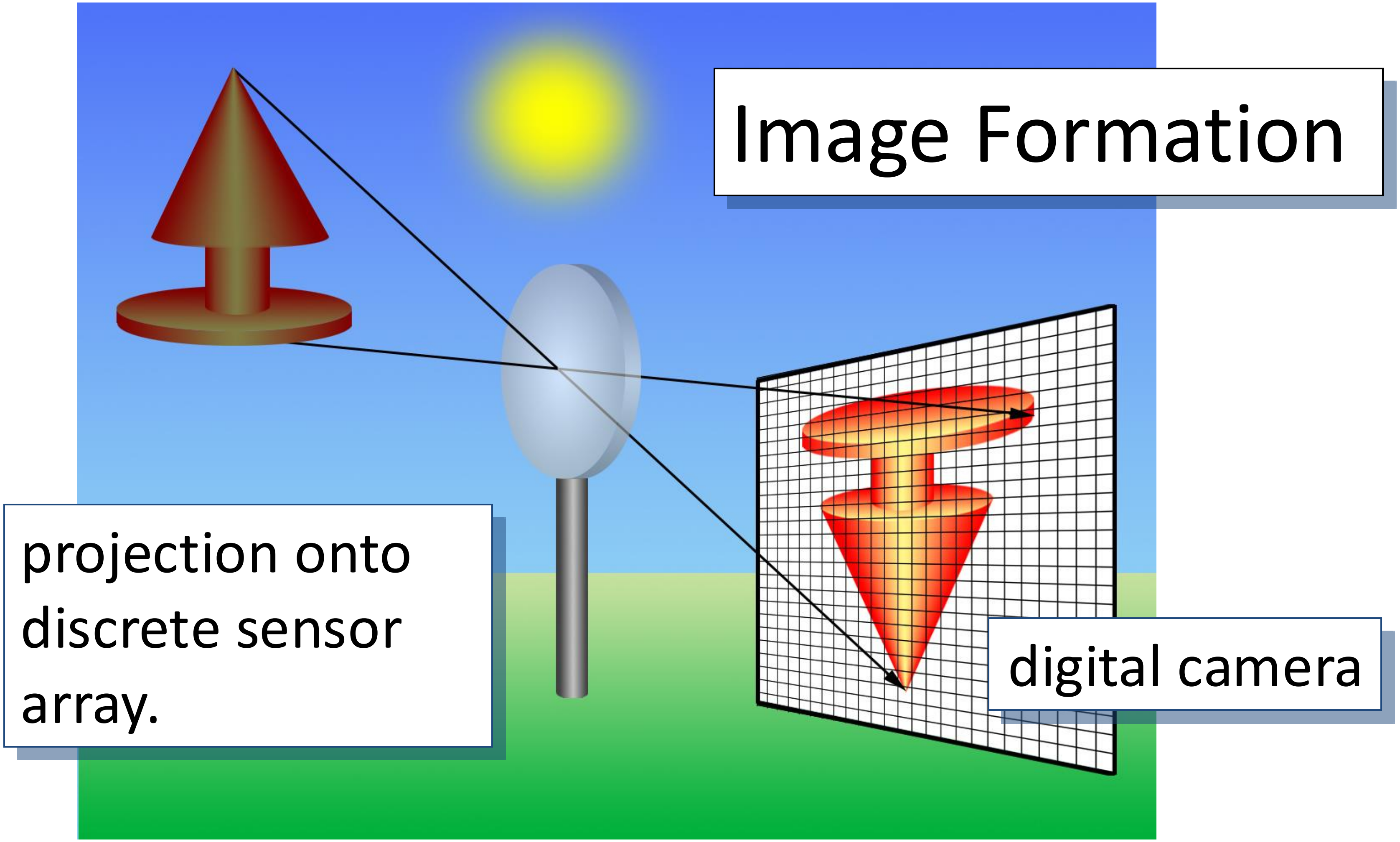
Slide by R A Peters

Digital Image Acquisition



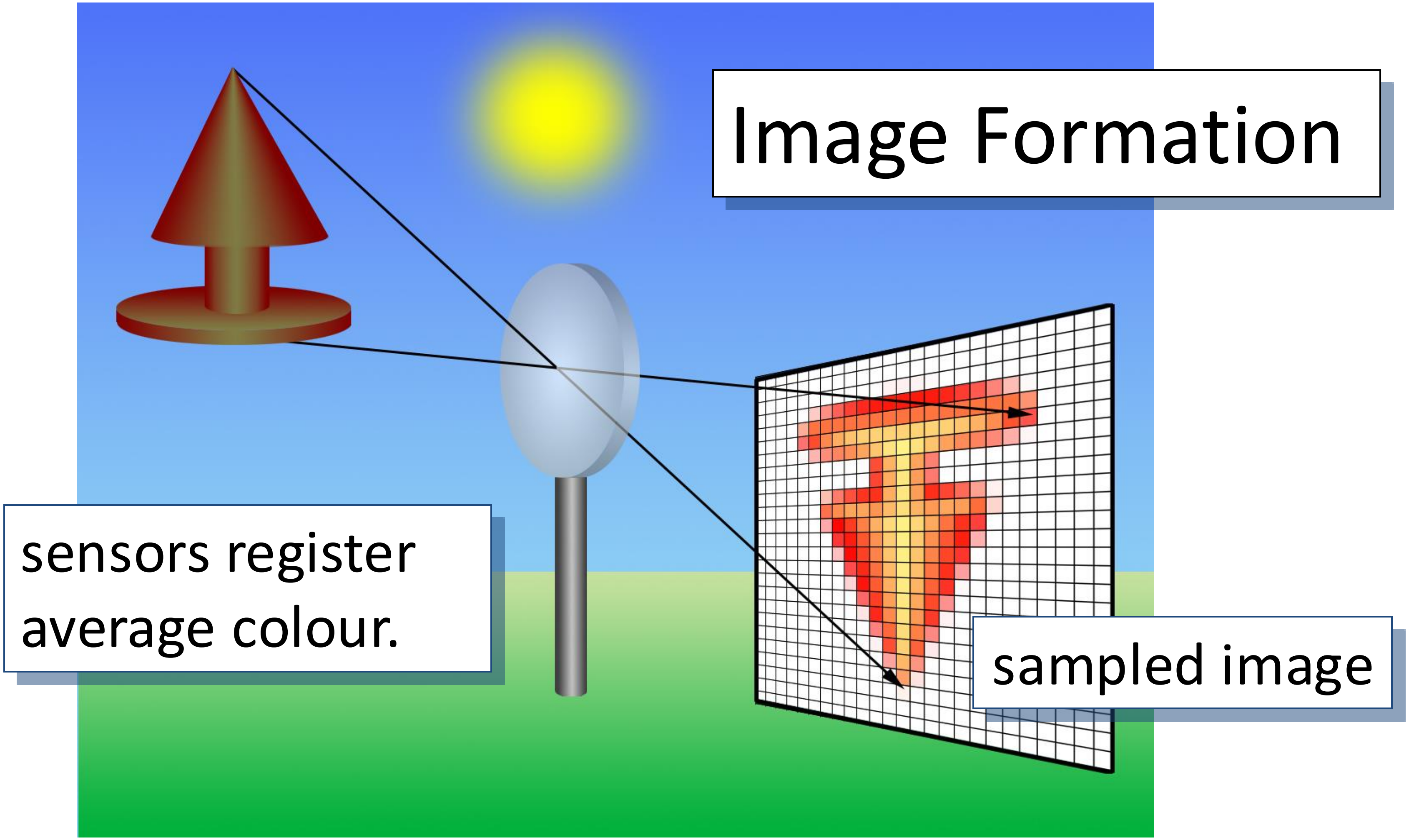
Slide by R A Peters

Digital Image Acquisition



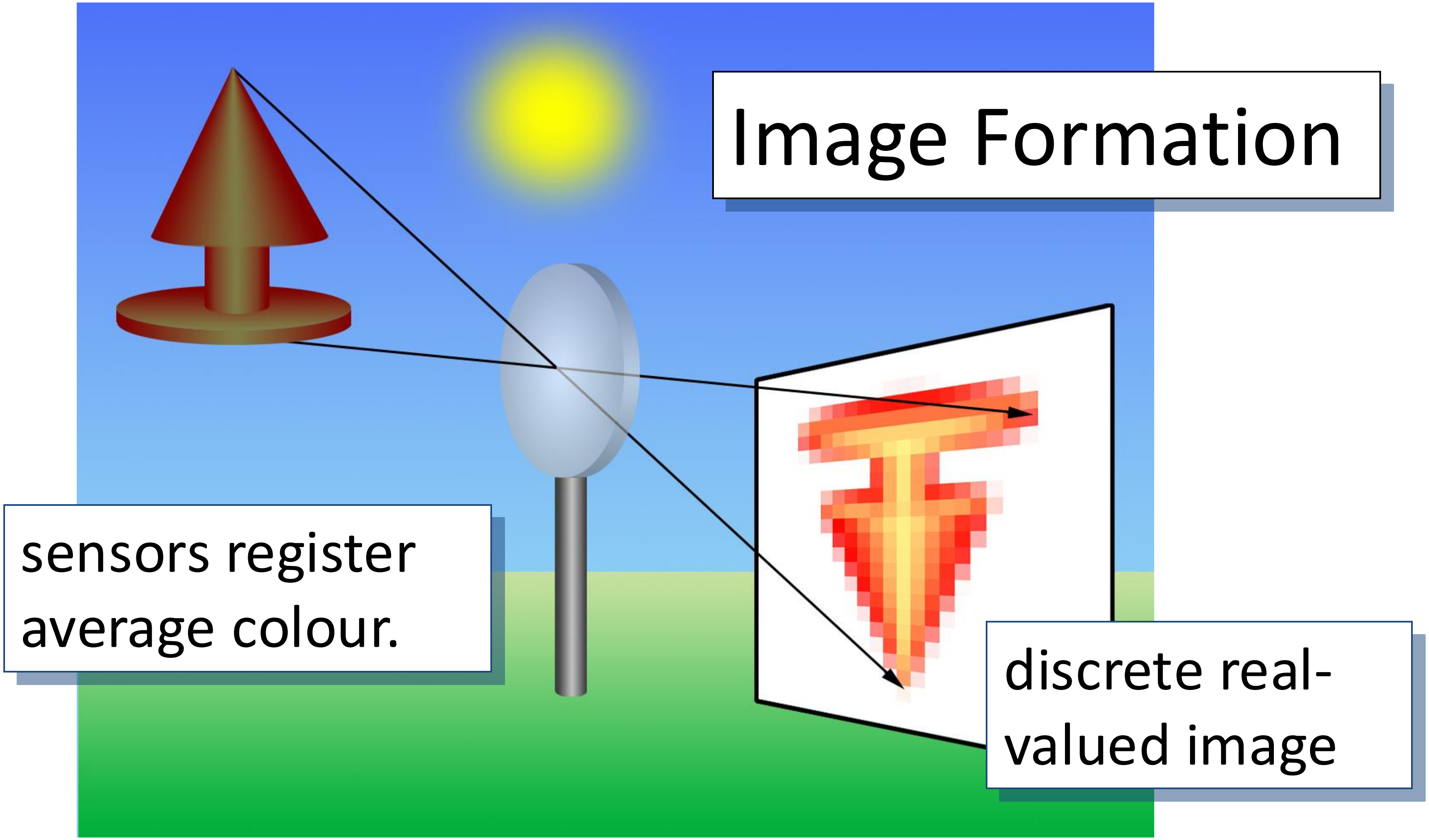
Slide by R A Peters

Digital Image Acquisition



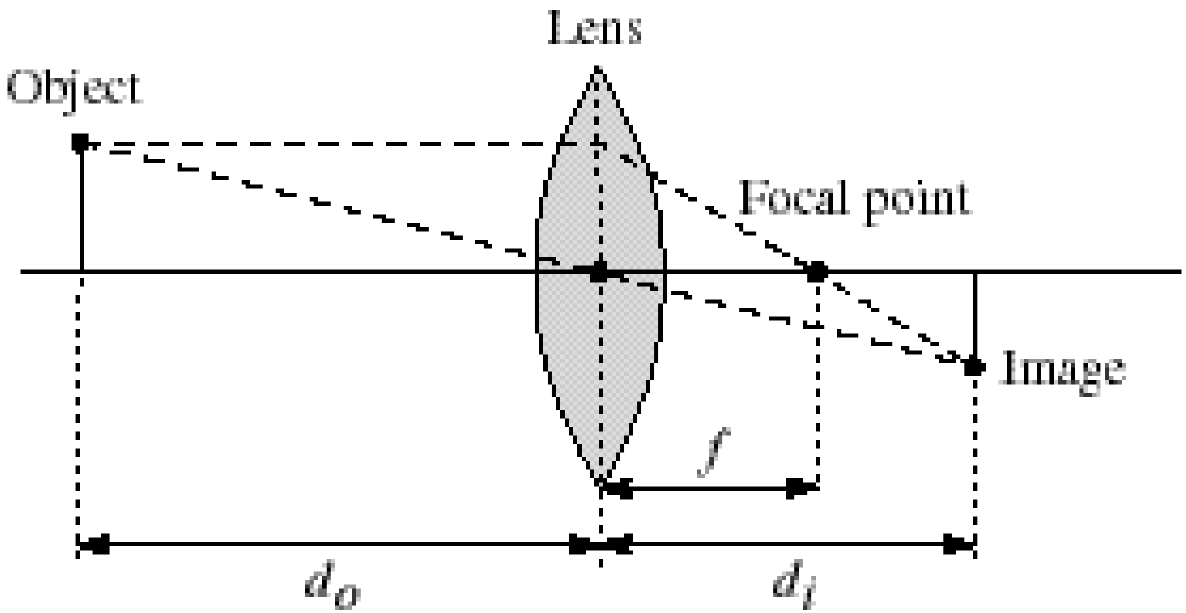
Slide by R A Peters

Digital Image Acquisition



Slide by R A Peters

Lens



Any object point satisfying the thin lens equation is in focus:

$$\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}$$

Sampling & Quantisation to obtain an Image

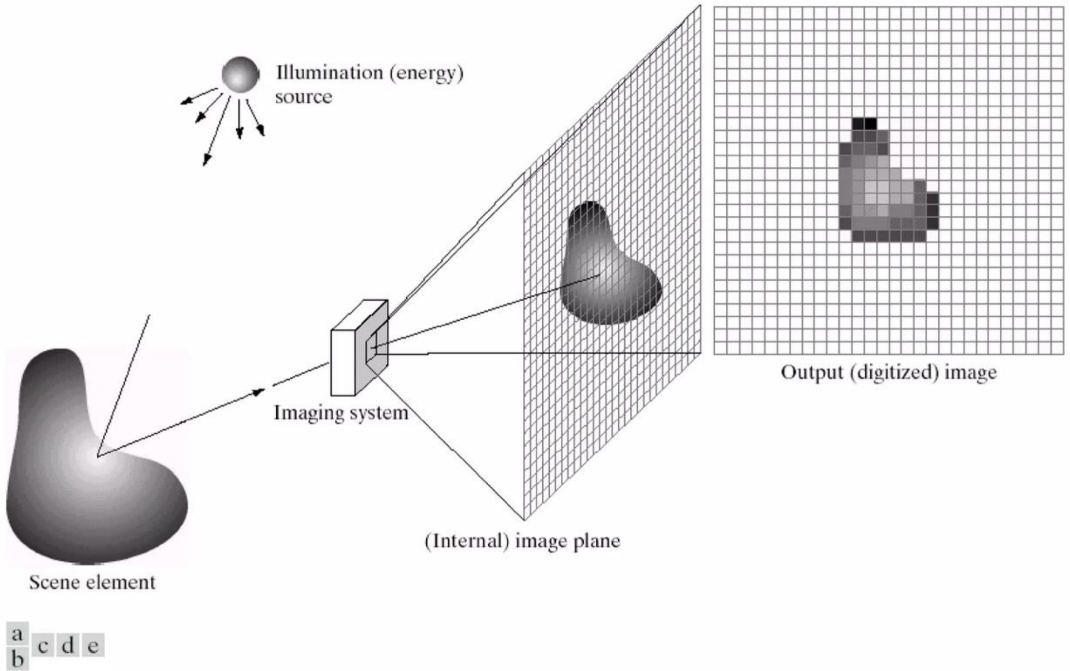


FIGURE 2.15 An example of the digital image acquisition process. (a) Energy ("illumination") source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

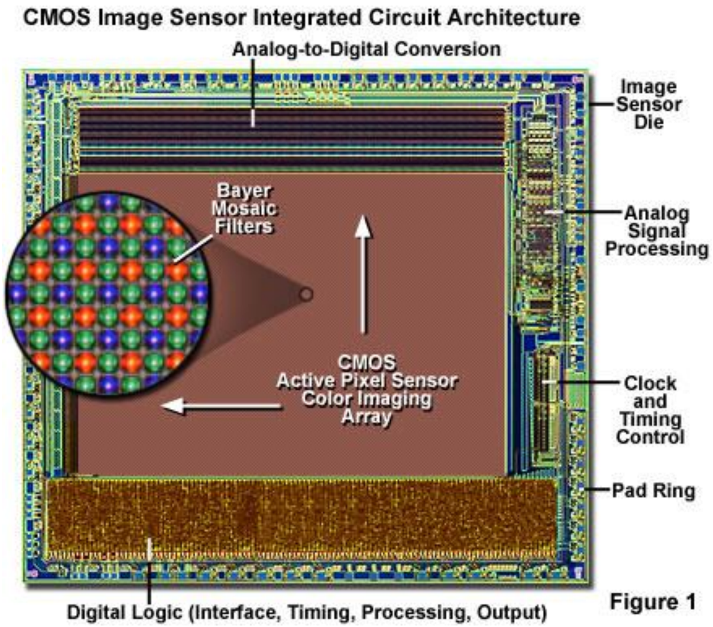


Figure 1

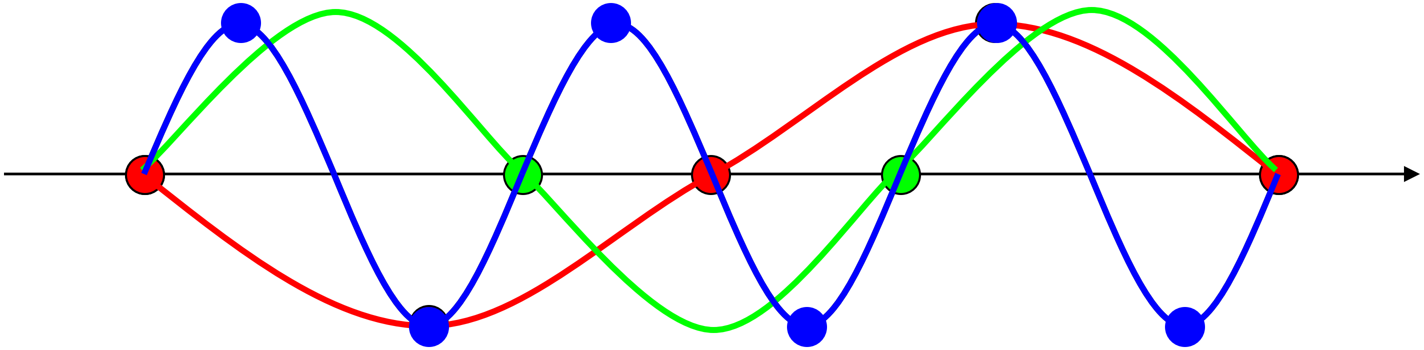
Pixel value corresponds to the electrical charges activated by light photons hitting the sensor array

Shannon's Sampling Theorem

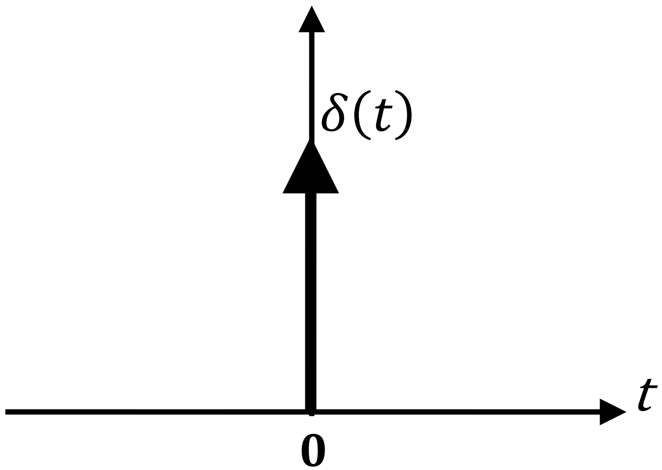
"An analogue signal containing components up to some maximum frequency ***u*** may be completely reconstructed by regularly spread samples, provided the sampling rate is above 2***u*** samples per second."

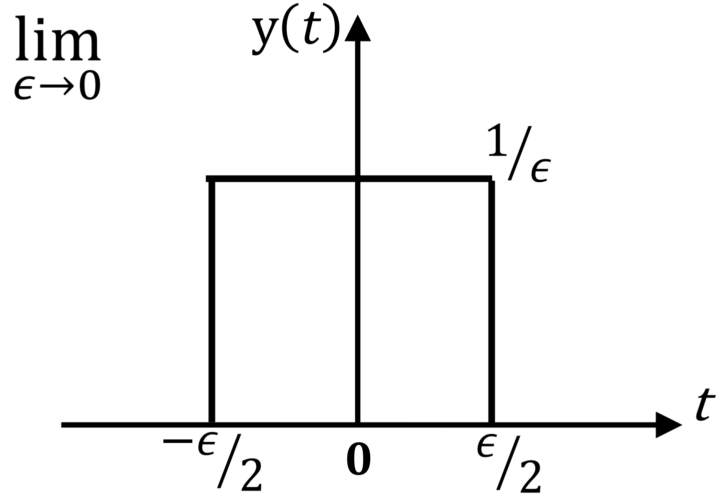
Claude Shannon

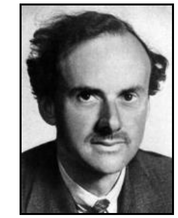
Also referred to as the Shannon-Nyquist criterion: Sampling **<u>must</u>** be performed above twice the highest (spatial) frequency of the signal to be lossless.



Modelling a Spatial Brightness Pulse - Dirac Delta-Function





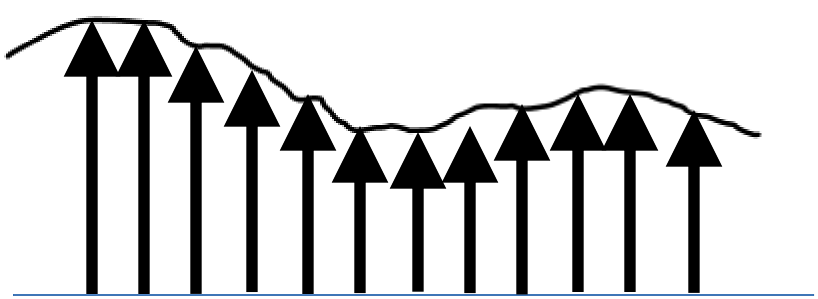

Paul Dirac

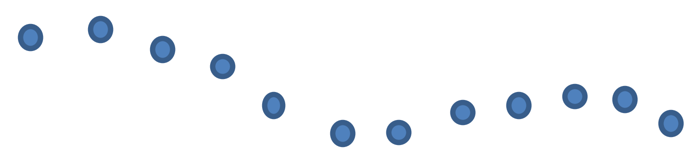
Definition:

$$\int_{-\infty}^{\infty} \delta(t)dt = 1$$

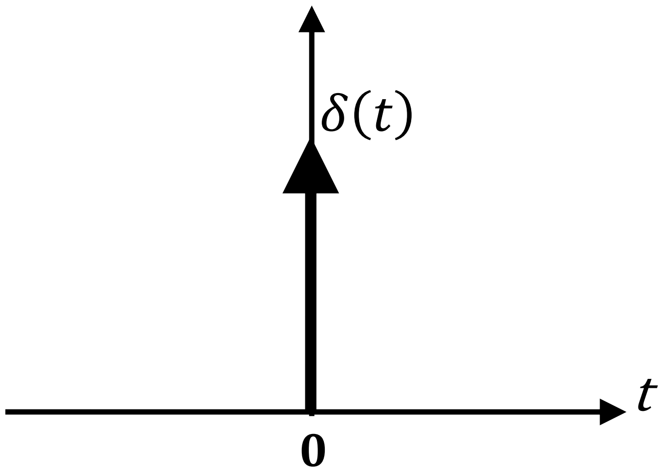
Intuitively:

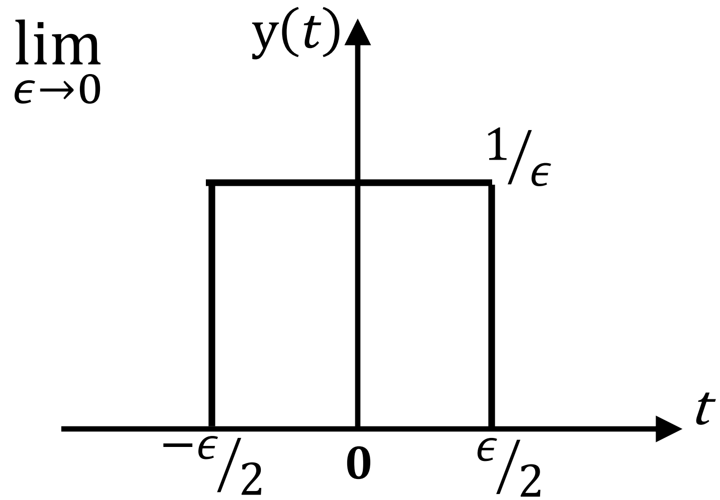
$$\delta(t) = \lim_{\epsilon \to 0}[y_\epsilon(t)]$$

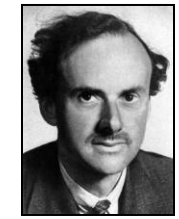




Modelling a Spatial Brightness Pulse - Dirac Delta-Function







Paul Dirac

Definition:

$$\int_{-\infty}^{\infty} \delta(t)dt = 1$$

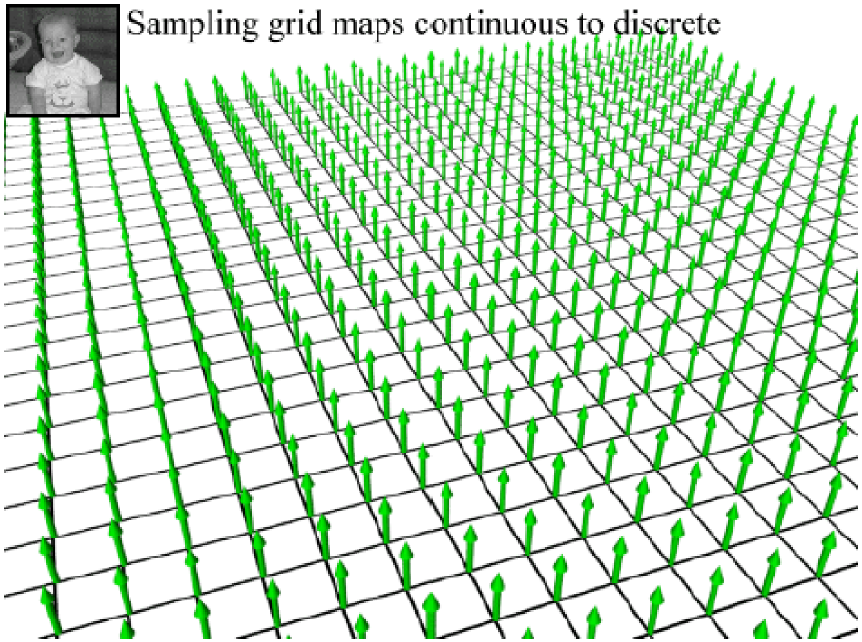
Intuitively:

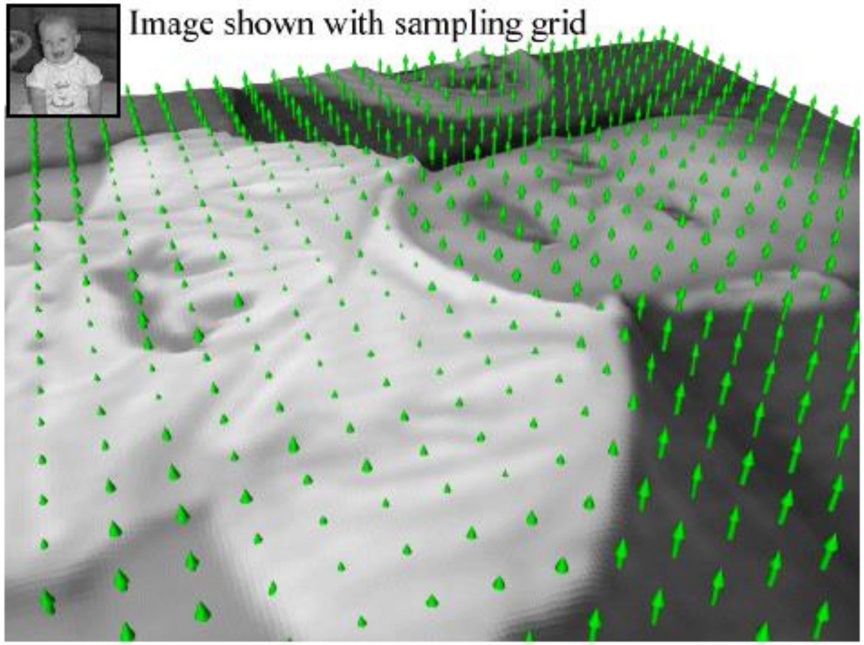
$$\delta(t) = \lim_{\epsilon \to 0}[y_\epsilon(t)]$$

Sifting Property:

$$\int_{-\infty}^{\infty} f(t)\delta(t)dt = f(0) \longrightarrow \int_{-\infty}^{\infty} f(t)\delta(t-\alpha)dt = f(\alpha)$$

Sampling in 2D To Obtain An Image



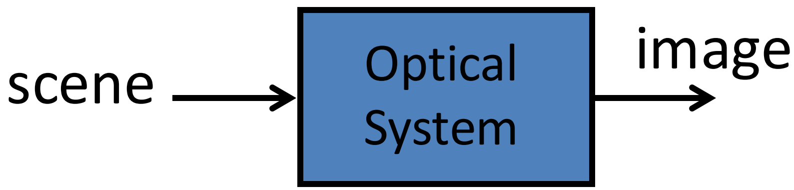


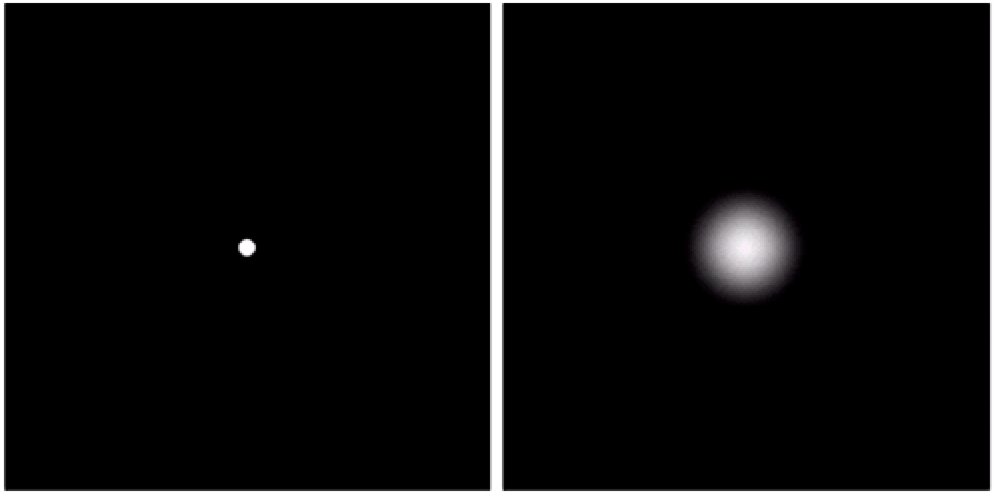
Images from Durand, Cutler and Hanrahan

The sifting property can be used to express a 2D 'image function' as a linear combination of 2D Dirac pulses located at points (a,b) that cover the whole image plane:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(a,b)\,\delta(a-x,b-y)\,da\,db = f(x,y)$$

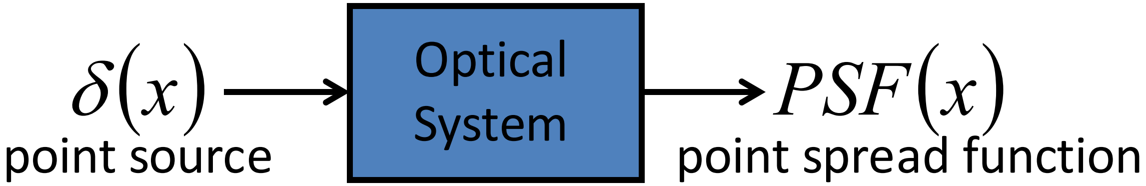
The Point Spread Function





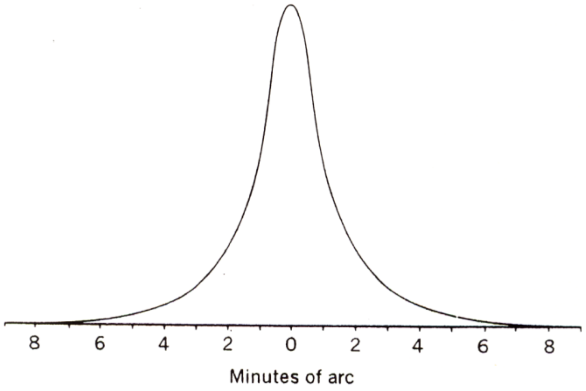
- Ideally, the optical system should be mapping point information to points again.

- However, optical systems are never ideal.



- Superposition Principle:
  An image is the sum of the PSF of all its points.

- Point spread function of Human Eyes

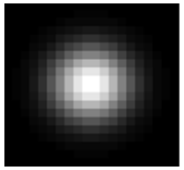


PSF Example

$$g(x,y) = f(x,y) * h(x,y)$$

$f(x,y)$ Original

$h(x,y)$



Blurry outcome $g(x,y)$

$h(x,y)$ is the PSF of the imaging device.

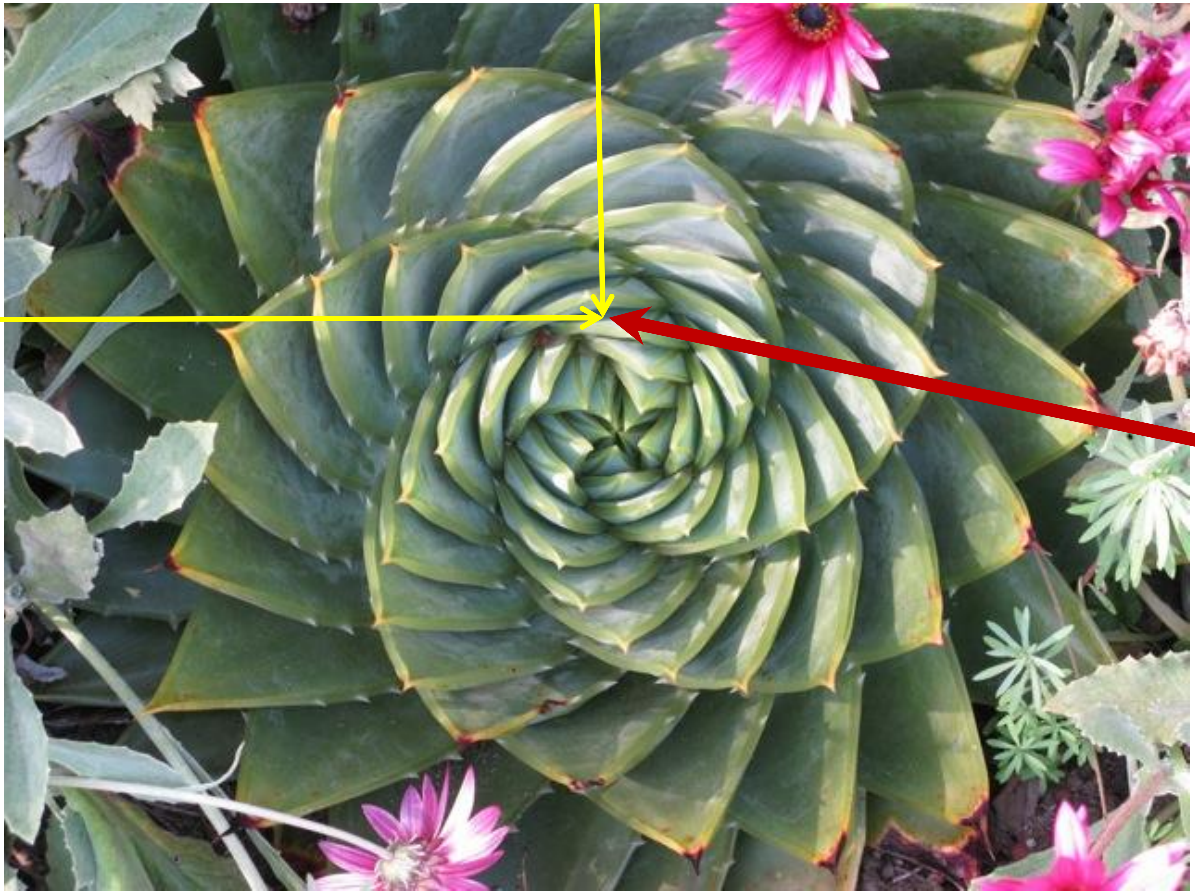
Adapted from a slide by A. Zisserman

How to model an image?

Addressing:

y_j

x_i



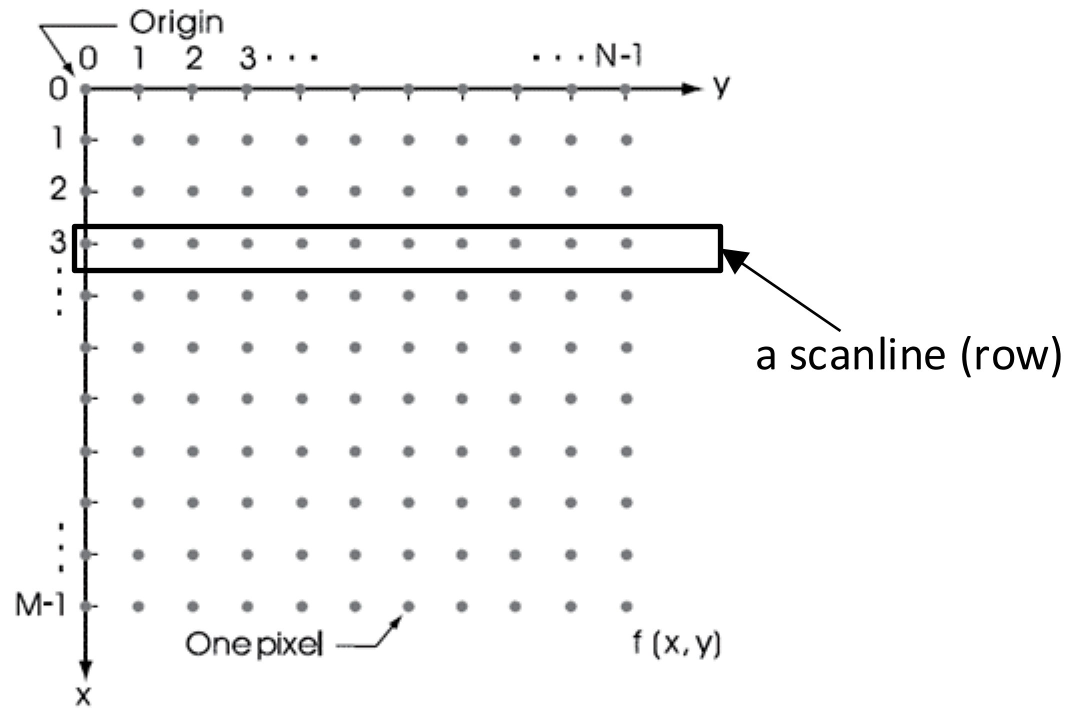
Colour

(c_R, c_G, c_B)

Storage: each colour value is 8-bits, hence 3x8=24 bits per pixel

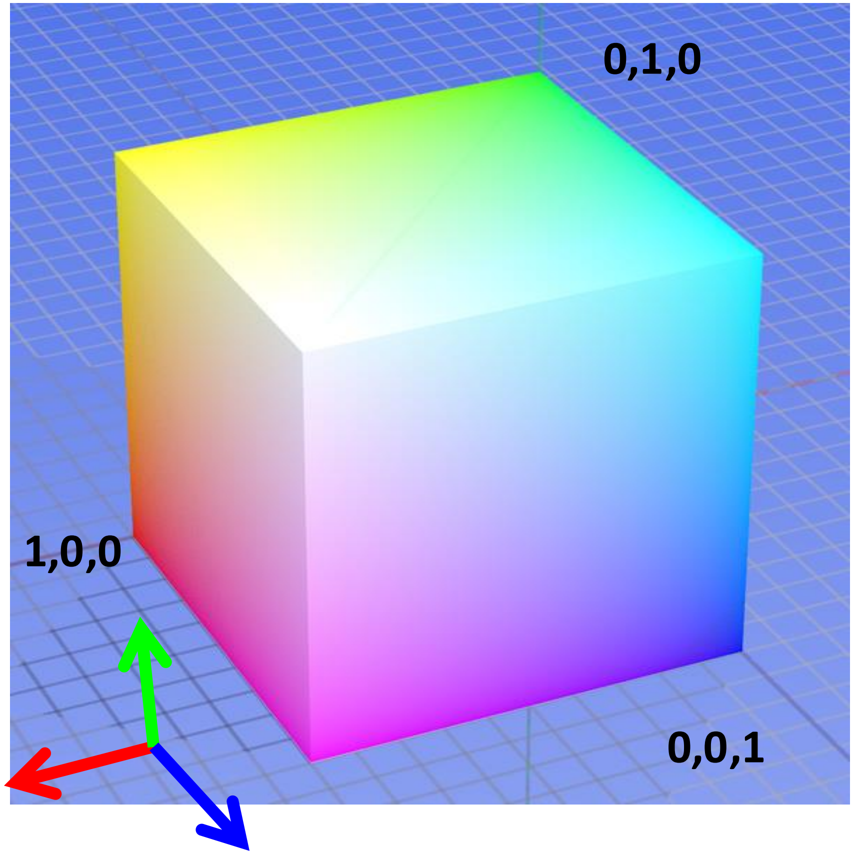
How to represent a digital image?

$$f[x,y] = \begin{bmatrix} f[0,0] & f[0,1] & \cdots & f[0,N-1] \\ f[1,0] & f[1,1] & \cdots & f[1,N-1] \\ \vdots & \vdots & \ddots & \vdots \\ f[M-1,0] & f[M-1,1] & \cdots & f[M-1,N-1] \end{bmatrix}$$



Colour spaces: RGB

Default colour space



R
(G=0,B=0)

G
(R=0,B=0)

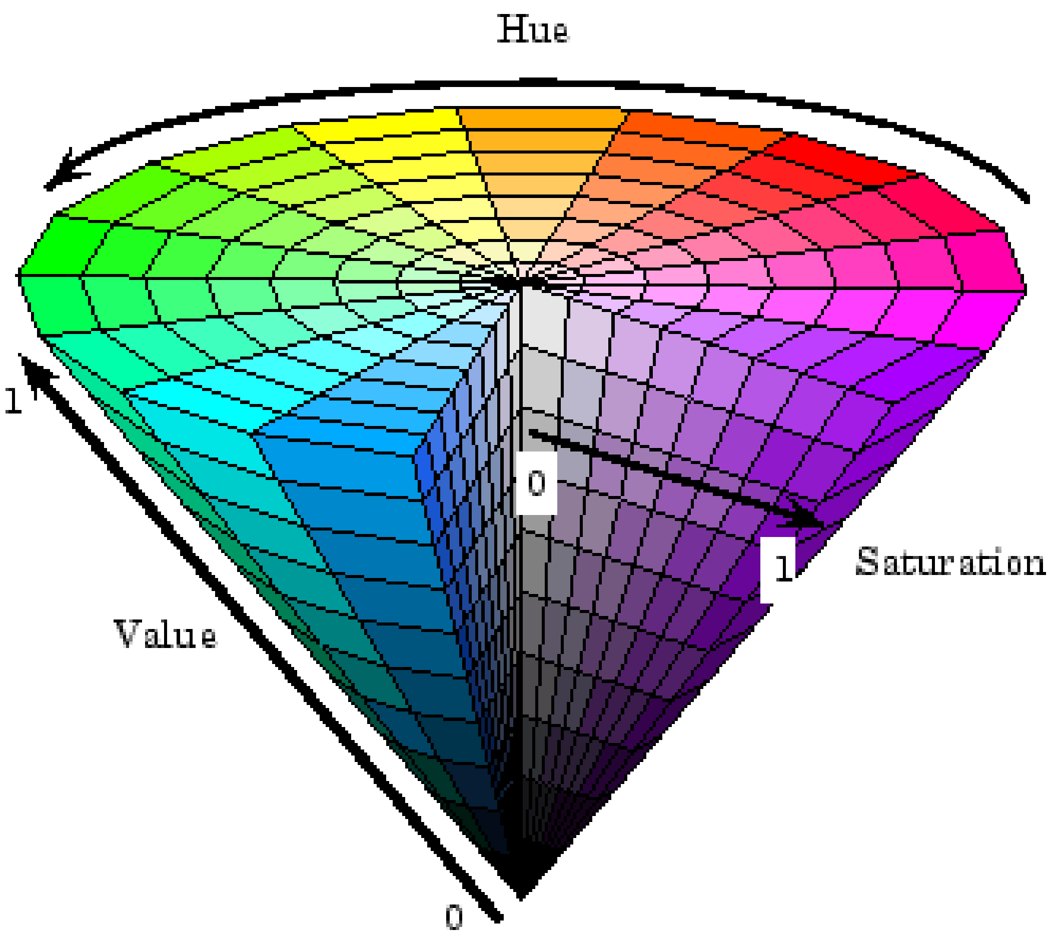
B
(R=0,G=0)

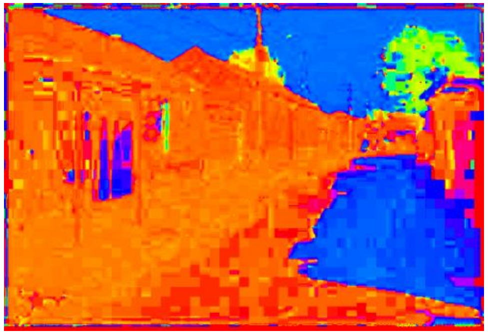
Some drawbacks
- Strongly correlated channels
- Non-perceptual

Image from: http://en.wikipedia.org/wiki/File:RGB_color_solid_cube.png

Colour spaces: HSV

Intuitive color space





H

(S=1,V=1)

S

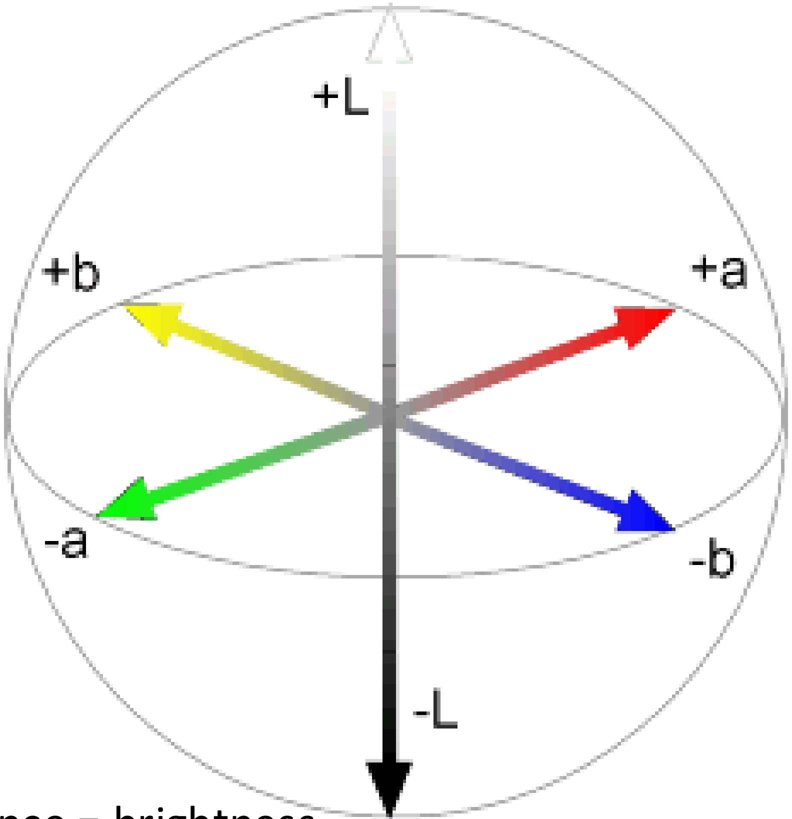
(H=1,V=1)

V

(H=1,S=0)

Colour spaces: CIE L*a*b*

"Perceptually uniform" colour space



L
(a=0,b=0)

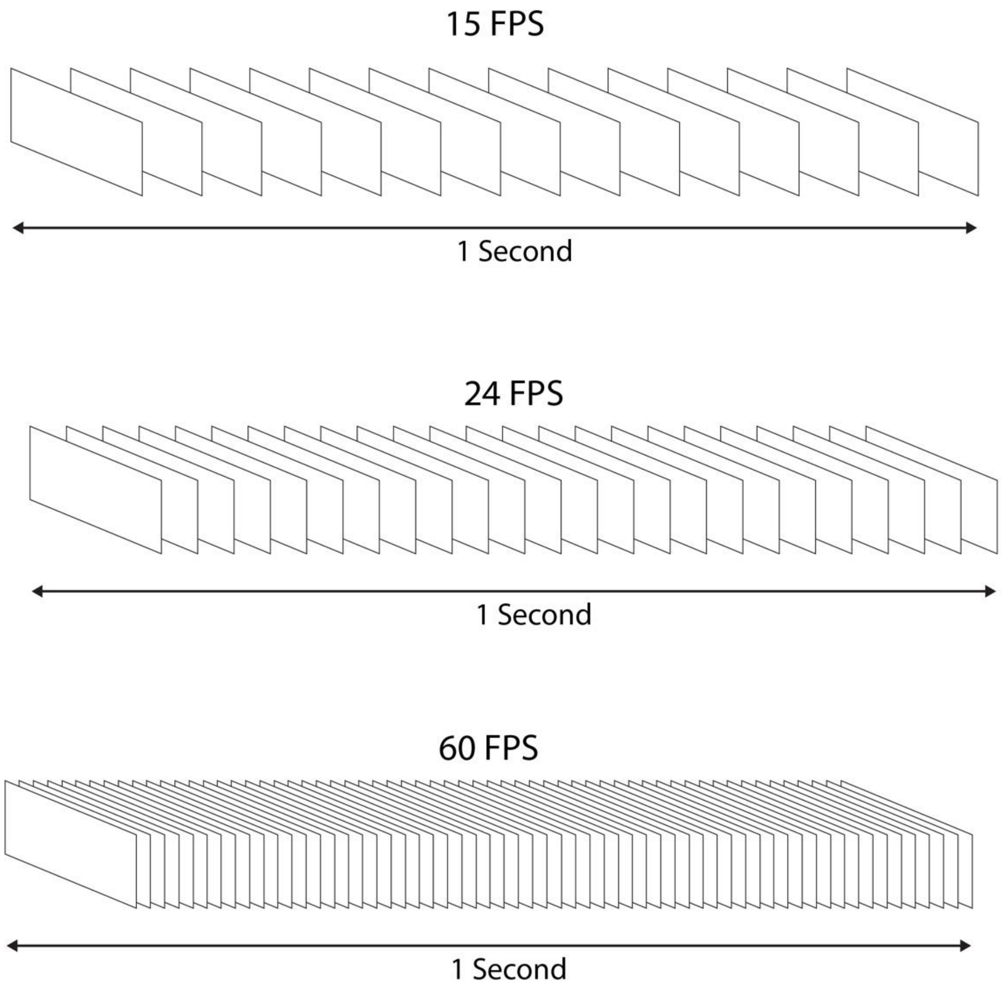
a
(L=65,b=0)

b
(L=65,a=0)

Luminance = brightness
Chrominance = colour
Can use Euclidean distances to represent colour differences!

How to represent a video?

$$f[x, y, t] = (R, G, B)$$



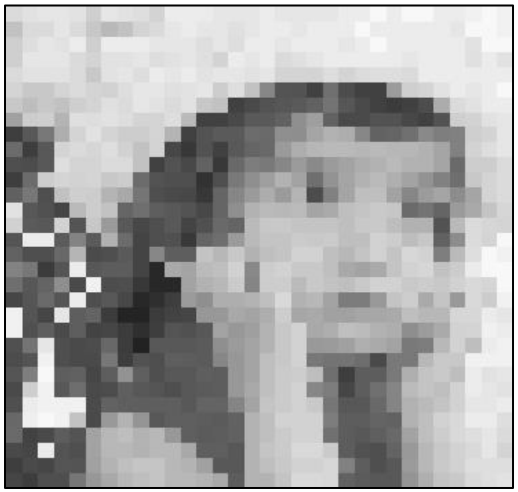
From quora.com

Summary: Spatial Sampling in Practice

The effect of very sparse sampling … is often ALIASING

256 x256

64x64



32x32

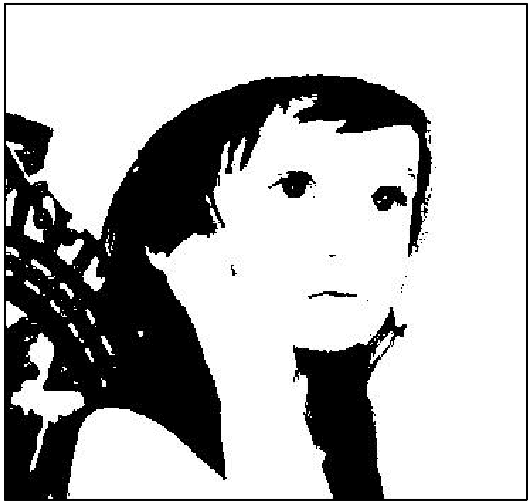
Anti-aliasing can be achieved by removing spatial frequencies above a critical limit (so-called Shannon-Nyquist Limit).

Summary: Quantization of the Image Function

Representing a continuously varying single channel image function $f(x, y)$ with a discrete one using quantization levels:

16 levels

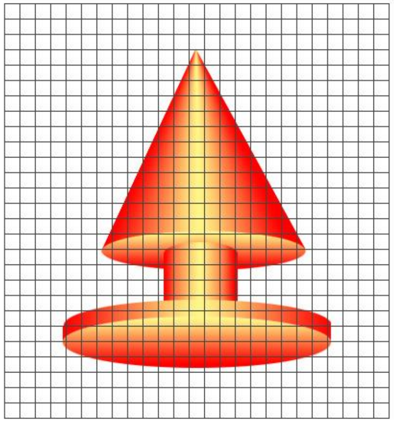
6 levels



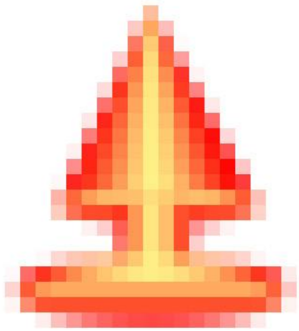
2 levels

Summary: Sampling and Quantization

pixel grid

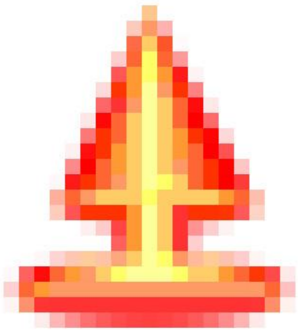


real image



sampled

quantized

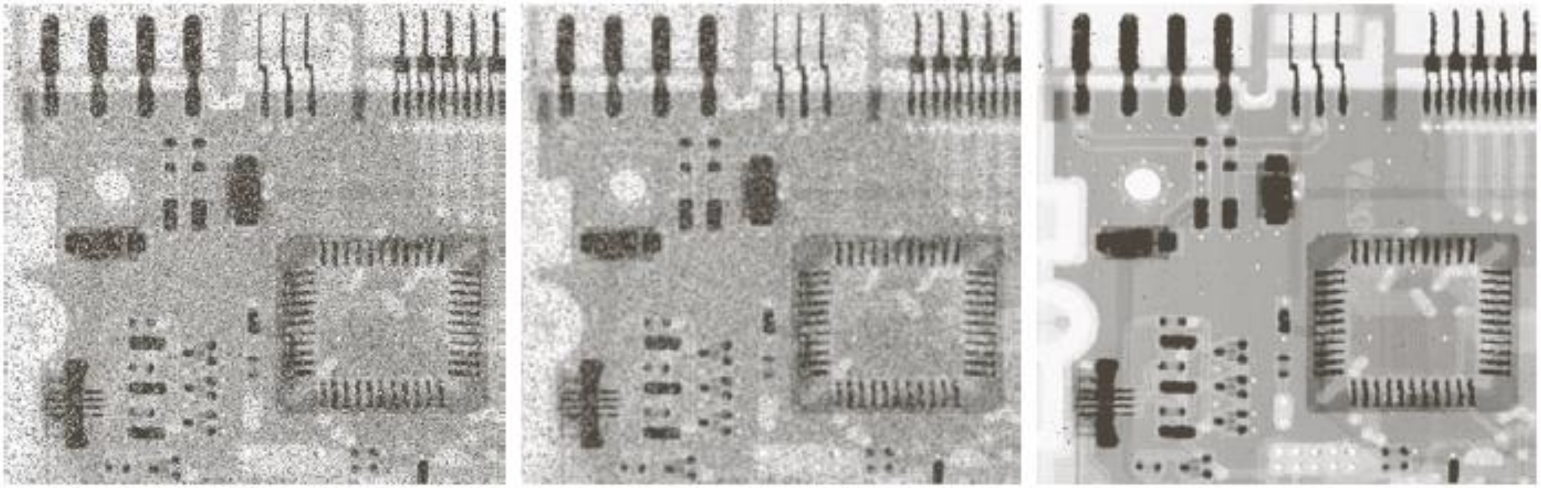


sampled & quantized

Slide by R A Peters

Next Lecture

Filtering



a b c

FIGURE 3.35 (a) X-ray image of circuit board corrupted by salt-and-pepper noise. (b) Noise reduction with a 3 × 3 averaging mask. (c) Noise reduction with a 3 × 3 median filter. (Original image courtesy of Mr. Joseph E. Pascente, Lixi, Inc.)